Scandinavian Journal of Forest Research

ISSN: 0282-7581 (Print) 1651-1891 (Online) Journal homepage: http://www.tandfonline.com/loi/sfor20

A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models

Hailemariam Temesgen, David Affleck, Krishna Poudel, Andrew Gray & John Sessions

Scandinavian Journal of Forest Research, 2015
Vol. 30, No. 4, 326–335, http://dx.doi.org/10.1080/02827581.2015.1012114

Taylor & Francis
Taylor & Francis Group

REVIEW ARTICLE

A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models

Hailemariam Temesgen[a*], David Affleck[b], Krishna Poudel[a], Andrew Gray[c] and John Sessions[a]

[a]*Department of Forest Engineering, Resources and Management, Oregon State University, Corvallis, OR 97333, USA;* [b]*College of Forestry and Conservation, The University of Montana, 32 Campus Drive, Missoula, MT 59812, USA;* [c]*Vegetation Monitoring Science & Applications Team, Resource Monitoring and Assessment Program, USDA Forest Service PNW Research Station, 3200 SW Jefferson Way, Corvallis, OR 97331, USA*

(*Received 5 January 2014; accepted 21 January 2015*)

Accurate biomass measurements and analyses are critical components in quantifying carbon stocks and sequestration rates, assessing potential impacts due to climate change, locating bio-energy processing plants, and mapping and planning fuel treatments. To this end, biomass equations will remain a key component of future carbon measurements and estimation. As researchers in biomass and carbon estimation, we review the present scenario of aboveground biomass estimation, focusing particularly on estimation using tree-level models and identify some cautionary points that we believe will improve the accuracy of biomass and carbon estimates to meet societal needs. In addition, we discuss the critical challenges in developing or calibrating tree biomass models and opportunities for improved biomass. Some of the opportunities to improve biomass estimate include integration of taper and other attributes and combining different data sources. Biomass estimation is a complex process, when possible, we should make use of already available resources such as wood density and forest inventory databases. Combining different data-sets for model development and using independent data-sets for model verification will offer opportunities to improve biomass estimation. Focus should also be made on belowground biomass estimation to accurately estimate the full forest contribution to carbon sequestration. In addition, we suggest developing comprehensive biomass estimation methods that account for differences in site and stand density and improve forest biomass modeling and validation at a range of spatial scales.

Keywords: biomass allometries; sampling error; uncertainity

Introduction

Forest management faces new and evolving challenges as society assesses ways to mitigate or adapt to climate change and reconsiders the balance of its interests between wood production and the provision of ecosystem services. Whatever paths this process may take, sound and broad-based decisions will continue to require accurate and defensible biomass and carbon estimates of past, current, and future forest conditions under different management scenarios. In addition, accurate biomass measurements and analyses are critical components in quantifying carbon stocks and sequestration rates, assessing potential impacts due to climate change, locating bio-energy processing plants, and mapping and planning fuel treatments. To this end, biomass equations will remain a key component of future carbon measurements and estimation. However, in many cases our ability to estimate forest biomass accurately is either unknown or severely limited. In this paper, we identify key points that we believe will improve the accuracy of biomass and carbon estimates from tree-level biomass equations to better meet societal needs.

The purpose of this paper is to review the present scenario of aboveground biomass estimation, focusing particularly on estimation using tree-level models. We highlight the sources of errors in these models and how they may propagate in time and space upon application. In addition, we discuss the practical challenges faced in the measurement of aboveground forest biomass and its components in the field. Since the challenges and opportunities for improvement discussed here cover a broader aspect of biomass estimation rather than a comparison of performance of certain techniques or approaches, we believe that the ideas are applicable to different species and across many different regions.

In the following sections, we first outline different types of biomass estimation methods presently in use and classify the sources of errors in estimating forest biomass using tree-level models. We then identify what we feel are the most critical challenges in developing

*Corresponding author. Email: Hailemariam.Temesgen@oregonstate.edu

biomass and carbon models and discuss selected strategies to develop sound and defensible methods. The fourth section presents an overview of how selected strategies can be integrated with existing databases and knowledge and includes examples and comments on the role accurate biomass equations play in meeting the above-described societal challenges. The final section summarizes our main points and provides the concluding remarks.

Background

Estimating forest biomass is important for quantifying the roles of forests as carbon sources or sinks and for supporting sustainable forest management. The knowledge of carbon stocks and fluxes is essential to understand current states and future courses of the carbon cycle in response to changing land uses and climatic conditions (Hollinger 2008). Growing carbon trade and the desire to mitigate climate change has spawned a number of policies, programs, and legislative actions. For example, forest carbon stocks for many developed and developing nations are reported as part of the overall carbon accounting under the United Nations Framework Convention on Climate Change. Also, Section 1605(b) of the Energy Policy Act for the USA also allows a voluntary greenhouse gas reporting program. Under this program, organizations may report their overall emission budgets and carbon sequestration rates.

The importance of forest biomass inventories is further highlighted by the necessity of improving our understanding of carbon fluxes within ecosystems and between ecosystems and the atmosphere. To that end, the amount of biomass existing as living vegetation or dead wood and debris in forest systems is an important factor in determining how forestry functions to regulate atmospheric carbon. In particular, accurate forest biomass estimates are crucial for the growing number of emissions cap and trade systems designed to reduce emissions of CO_2 and other greenhouse gases (GHG). Given that these systems provide offsets or credits for carbon sequestration, the need for accurate biomass estimation is greater than ever before. This need is further compounded by the increasing number of climate change agreements and action plans at varying scales. For example, in the 2007 legislative session, the Oregon State Legislature passed House Bill 3543 which declared that it is the state's policy to reduce GHG to 10% below the 1990 levels by 2020 and to further reduce GHG to 25% below the 1990 levels by 2050.

Tree biomass estimates are the basis for US forest carbon inventories and most international negotiations. The emergence of biomass as a critical variable in assessing sequestration of atmospheric carbon and in providing critical information to forest resource management and policy decision-making has focused attention on its accuracy (Heath et al. 2008). Local prediction accuracy of biomass and carbon maps developed for regional analyses (>4,000,000 ha) has been questioned by land managers, decision-makers, and research scientists (Ver Hoef & Temesgen 2013). It is reasonable to seek information to add confidence in biomass estimates and map products. In this application, studies on improving the predictive ability of biomass estimation methods and models are crucial.

Currently, many state and regional forest biomass estimates for the USA are obtained from tree diameter measurements using equations developed for national-level application by Jenkins et al. (2003). The lack of species-specific parameters and site-specific attributes in these models allows for a wider domain of application and greater consistency across administrative zones but has clear implications for accuracy at smaller scales. The performance of the national biomass equations for the Pacific Northwest forests in particular has recently been questioned. Zhou and Hemstrom (2009) reported that the equations of Jenkins et al. (2003) resulted in merchantable biomass estimates that were 17% higher than those based on regional biomass equations. While the accuracy of the regional biomass models that were derived from local volume equations is unknown, a 17% difference between the estimates is disquieting.

Biomass estimation methods used in forestry

Although forest biomass used generally includes all live and dead material in all forms of vegetation (trees, shrubs, vines, etc.), most of the research on biomass estimation has focused on the above-ground components in live trees because of the prominence of this fraction and the difficulty in collecting below-ground data (Lu 2006). Moreover, estimations of this fraction in live trees can typically take advantage of strong allometric relationships between biomass and measured (directly or indirectly) tree characteristics. Quite different methodologies are generally needed to estimate biomass in standing dead trees, biomass in understory vegetation, biomass in downed wood, and biomass in the forest floor (Domke et al. 2011).

Based on the resolution of prediction, biomass equations for aboveground live tree biomass estimation can be categorized into three levels of increasing specificity, demanding correspondingly more detailed inventory data inputs:

(1) Regional biomass conversion factors
(2) Stand-level biomass equations
(3) Tree-level biomass equations

Commonly, tree-level biomass equations are derived through the use of destructive sampling and linear or nonlinear regressions (Baldwin 1987; Parresol 2001). Such equations estimate total and component biomass for individual trees based on empirical allometric relationships with more easily measured inventory attributes – most commonly diameter at breast height. Applying these equations to obtain stand or regional biomass estimates requires that tree-level inventory data are collected across the population of interest. In contrast, stand-level biomass equations are derived by relating aggregate tree biomass characteristics to other stand-level attributes such as stand basal area and tree density. Stand or regional biomass estimates can then be formed using strictly stand-level inventory data (e.g. aggregate basal area per unit area, tree density). Finally, regional biomass conversion factors identify biomass density levels (determined from direct or indirect biomass measurements) associated with qualitative stand types common to a particular geographic area (Fang & Wang 2001). This approach then requires only forest area and forest-type classifications in order to obtain regional or national biomass estimates.

Relative to the use of tree-level models, application of stand-level biomass equations or regional conversion factors requires less detailed information about the populations of interest. In particular, only stand-level characteristics are needed. Yet clearly this can also be a disadvantage from an accuracy standpoint: variation among stands with, for example, the same forest type or basal area stocking, cannot be recognized in biomass or carbon estimation. Application of tree-level models not only demands more detailed input data, but also allows one to recognize variation observable at the most basic unit of forest inventory. As noted below, the development of tree-level models also allows one to consider a sizeable body of scientific knowledge concerning the structural and hydraulic constraints on tree form and mass distribution. For these reasons, we henceforth focus primarily on tree-level biomass models.

Sources of error in estimating forest biomass

Sampling errors or uncertainty in the inputs needed to estimate forest biomass leads to error and uncertainty in biomass estimates, irrespective of the biomass equations or models applied. As well, errors in the form or calibration of the models themselves are also relevant because these errors will propagate through to the application stage and thus may be amplified over time and space. Errors in forest biomass models can be traced back to aspects of sample design and model calibration, beginning with the selection of sample plots and sample trees through model development. Moreover, errors incurred at one stage are generally carried through to or may otherwise impact later stages.

A large number of field measurements are a prerequisite for developing aboveground biomass estimation models and for evaluating the aboveground biomass estimation results (Lu 2006). However, early biomass models or equations were generally focused on narrowly defined populations, sometimes as small as a single stand or few stands (e.g. Pollard 1972). Most of the applications envisioned today are for large ownerships or for regional/state/national scales. At these scales it is problematic to identify, for example, all silvicultural treatments in use or potentially in use in the near future that might affect tree allometry and biomass distribution. Yet, the construction of biomass equations requires a sample from the entire range of tree sizes, ages, sites, and silvicultural treatments, represented in the population of interest (e.g. Bockheim & Lee 1984; Cochran et al. 1984; Barclay et al. 1986).

The predictive accuracy of a biomass model depends, in part, on the scope and extent of data used in development, on the variability in biomass within the population, and on the methods used to formulate and calibrate the model. Thus, sources of errors in biomass estimation can be considered as arising from three main phases of model development: sampling, measurement, and model specification. The following points highlight some of the errors involved in these phases, as well as the potential for error propagation to occur from phase-to-phase in application:

Sampling errors

Errors arise from plot and tree selection because of intrinsic variability in tree attributes such as wood density and crown architecture. That is, across most tree populations there is appreciable variation in biomass levels (and allocations to wood, foliage, etc.) within species, size, age, and productivity classes, and there is variation also in the strength or form of the relationships between the latter attributes and biomass. Direct biomass measurement is an expensive and destructive process so most models are developed with relatively small samples of trees. Therefore, sampling error in model form and fit is important elements of later biomass estimate uncertainty, and typically only errors in model fit are estimated.

Different studies have shown a varying range of sampling error in biomass estimation. Additionally, the percent sampling error for the total biomass of trees in the model-development sample differs from the percentage sampling error for individual trees. Using two-stage sampling to estimate individual tree biomass, Ozcelik and Eraslan (2011) showed the sampling error ranged from 2.51% to 22.63% per tree (and 2.65% of total biomass). Their method of subsampling involved the first stage of randomized branch sampling (RBS) and the second stage of importance sampling (IS). Williams (1989) also tested the RBS and IS on loblolly pine (*Pinus taeda*) data and found per-tree sampling error ranging from 5.3% to 28.9% whereas sampling error for total biomass was just 3.3%. Based on the field test with eight trees from a mixed oak stand, Valentine et al. (1984) found that the sampling error ranged from 2.6% to 14.4% of the actual fresh weights of the trees and 4.9% of the total biomass.

In addition, if sample selection has been restricted to trees with symmetric and undamaged crowns, to fully stocked stands, or to the most accessible portions of the landscape, then the sampling error estimated in model development may not correspond to the error incurred when the model is extrapolated to the full population of trees or stands.

Measurement errors

As noted, tree biomass is difficult to measure directly, as to a lesser extent are tree taper/volume and wood density attributes that may be used to predict biomass. Biomass measurements are destructive and with many opportunities to lose material (e.g. branch breakage in felling, saw kerf in stem dissection) and a few to gain material (entanglement of branches from other trees) measurement error is inevitable and not necessarily symmetrically distributed. Measurement errors also arise from irregularities in tree form such as out-of-round boles (leading to variable diameters or cross-sectional areas) and instrument errors associated with weighing plant materials in the field. Subsampling procedures used to estimate crown or stem mass can mitigate certain nonsampling errors but add another level of within-tree uncertainty. The latter may be possible to identify and quantify from the subsampling procedure itself (e.g. if a probability sampling strategy is used), yet this additional uncertainty around individual tree biomass can accentuate the difficulty of identifying the form and strength of relationships between biomass and other tree or stand factors. The presence of measurement errors results in the biased and inconsistent estimates of model parameters and leads to the erroneous conclusions (Canavan & Hann 2004).

Model misspecifications

These arise from the methods of model identification and calibration and are influenced primarily by the size and scope of sample data. In general terms, larger samples facilitate the identification and quantification of biomass allometries. Yet it is important to recognize that biomass samples are often hierarchically structured. Components of interest in above-ground biomass estimation are generally foliage, branches, and bole wood and bark; measurements of these components are taken on the same trees and these trees are often nested within plots that are in turn nested in stands. These forms of data clustering must be accounted for to properly understand variation in the allometric relationships across a species' range. Also, biomass model parameters are often estimated on the log-transformed scale. These transformations are made to stabilize variation or so that the assumptions of parametric tests are satisfied. However, nonlinear transformation of variables fundamentally alters the meaning of model parameters complicating inference regarding the original allometric parameters of interest. Finally, most biomass models are developed in a parametric framework and errors will arise from the goodness-of-fit of the parametric approximations. Other model specification errors include:

Omitted variables. Only a few biomass models account for tree height variation within diameter classes. Past work has demonstrated that diameter has the strongest associative relationship with tree biomass, but decades of research on bole volume estimation makes clear that height:diameter ratios are critical determinants of cubic volume and by extension stem biomass. On the other hand, uncertainty in height estimation is another source of error in estimating forest biomass from tree-level inputs. Similar arguments could be made for stem taper, crown length, and site quality, particularly for individual components such as stem biomass or crown biomass.

Data integration and fusion. Indiscriminate pooling of data across a wide range of sites varying in age, site quality, and stand density, leads to high variability in biomass within DBH classes. In turn, this complicates the identification of multivariate biomass allometries or contributes to a lack of precision in diameter-based biomass estimation. Similarly, tree species and certain species subpopulations (e.g. interior [*Pseudotsuga menziesii* var. *glauca*] vs. coastal Douglas-fir [*P. menziesii* var. *menziesii*], or shore pine [*Pinus contorta* ssp. *contorta*] vs. lodgepole pine [*P. contorta* ssp. *latifolia*]) may differ in their growth, form, and mass characteristics (see Standish et al. 1985). Equation systems that provide biomass estimates only by species groups, or apply one species' equation to multiple "similar" species, can thus result in biased estimates (Zhou & Hemstrom 2009).

In partial summary, sampling and measurement errors in tree biomass, along with measurement errors in potential predictor variables (tree height, taper, and site index), lead to errors in biomass model form and fit. These errors will be compounded by sampling errors accrued in the inventory of target stands and by the application of fitted biomass models to the tree data in those stands especially those that are outside the scope of the models (e.g. species, sizes, and form classes not considered in model development). However, the magnitudes and impacts of these forms of error propagation when equations are applied over hundreds to thousands of hectares are unknown.

Critical challenges in developing or calibrating tree biomass models

Accurate forest biomass and carbon estimation is a complex endeavor that requires sound statistical formulations and rational biological considerations. In light of the sources of errors described in the previous section there is always room for improving forest biomass and carbon estimation. In this section, we describe the critical challenges and emerging informational needs currently encountered in forest biomass and carbon estimation.

Variation in biomass allometries

The use of allometric biomass equations is inevitable because the weighing of trees and their components for direct biomass determination is destructive and prohibitively expensive. However, allometric relations are not necessarily stationary across a species' range or even across size classes within a species. Therefore, it is important to note that different regions and species link easily measurable tree and stand attributes to biomass through different functions such as logarithmic vs. linear or quadratic forms. There are decades of research documenting the effects of stand density or local competition on height–diameter ratios and on crown characteristics including biomass (Grigal & Kernick 1984; Barclay et al. 1986) and wood density (Gonzalez 1990). For regional or small-scale biomass estimation, it is important that biomass models allow for calibration against relatively inexpensive stand or tree metrics such as stand density or crown length. Also, geographic variation in moisture stress has been associated with differences in biomass allometries (e.g. Callaway et al. 1994), and adaptations to moisture stress linked with differences in wood density (Bouffier et al. 2003). Identifying such variations is a major challenge given the high costs of direct biomass measurements. Nevertheless describing these forms of variation is critical to making efficient use of relatively small numbers of direct biomass measurements from felled-tree studies, and ultimately for predicting biomass over a broad geographic range and differential climatic conditions.

Similarly, most biomass equations are static in nature, deriving mass from current stand attributes. We are aware of no efforts to model tree or stand growth directly in terms of biomass. However, silviculture – and density control in particular – can affect biomass accumulation rates. The literature on the impacts of silviculture on diameters and basal areas is voluminous. Despite early work in the 1980s (e.g. Bockheim & Lee 1984; Cochran et al. 1984; Barclay et al. 1986), there is a lack of understanding on how silvicultural treatments such as thinning and fertilization affect allometric relationships and biomass components. Using the data from 79 destructively sampled *Pinus ponderosa* trees in Northeastern California, Ritchie et al. (2013) found that foliage biomass relationships varied substantially between thinned and unthinned units, although branch wood and bole biomass estimates were more stable. Similar research carried out in different locations and for different species would help determine the type of models and amount of data to be collected to develop an efficient allometric equation.

As an alternative to the use of allometric biomass equations, tree-level volume equations are often used in conjunction with biomass expansion factors (BEFs) to obtain biomass estimates. BEFs are used to convert tree volume to mass in several ways, including conversion of not only gross volume to overall stem mass but also gross volume to merchantable biomass. The use of BEFs allows for the application of existing volume equations that have been developed from substantial sampling efforts and that have been in use in commercial and inventory settings for long periods of time. Nevertheless careful consideration should be given in using the expansion factor approach because volume:biomass ratios can be expected to vary with stand age, stand density, site index, and other factors. Also, BEFs are multiplicative functions so the BEF's errors could be much greater than the additive functions that sum the different component biomass (Tobin & Nieuwenhuis 2007). Therefore, it is necessary to consider different volume-to-biomass conversion factors for small and large trees while making a choice of a BEF.

Support to quantify carbon sequestration

Emerging carbon markets demand detailed long-term information on forest carbon stocks. These include not only standing merchantable tree stocks, but also carbon stocks in small trees and noncommercial species. Likewise included are stocks in standing dead trees and downed woody debris. Thus a comprehensive forest biomass estimation strategy must extend beyond the traditionally merchantable segment of the tree population. Small live trees and minor species might be approached using traditional methods, although diameter at breast height obviously will not be a useful scaling factor for trees below breast height. More importantly, accounting for the wide variety of forms and decay in dead wood will require departures from simpler allometric methods used for sound, mature trees.

Belowground biomass and carbon (including stocks in soils, root systems, and the forest floor) as well as biomass in understory vegetation are other components of the forest carbon pool that demand increased attention. These components are not tracked in many forest inventories, not only in part because of the expense of belowground measurements but also in part because techniques for biomass/carbon determination are lacking or poorly developed relative to tree stocks. The very high spatial variation in belowground carbon (see Nave et al. 2010) also complicates estimation.

Improving the quality of biomass and carbon estimates and their validation at a range of spatial scales is important to provide confidence in carbon markets and in quantifying the amount of sequestration. Deterministic analysis cannot portray the variability of biomass and carbon that is a fundamental characteristic and critical for most emerging carbon accounting and sequestration policies. For example, while the Intergovernmental Panel on Climate Change (IPCC 2007) asserts that carbon in trees is approximately 50% of dry biomass, it is unknown if that ratio is correct when estimating coarse root biomass. In addition, none of the biomass equations in IPCC (2007) report account for soil biomass.

Studies on biomass and carbon attributes at small and large scales over short- and long-time periods are needed to mimic temporal and spatial variability and answer large-scale carbon sequestration inference problems. Quantifying and incorporating biomass and carbon variability and error distributions would expand the scope of inference, validate results (or projections), and support carbon sequestration endeavors.

Support to quantify bio-energy projects

Aboveground biomass is a substantial source of renewable energy. Increasingly, biomass of traditionally noncommercial components such as broken tops, standing dead trees, bark and branches, and downed wood is becoming important substitutes for fossil fuel. The mass and composition of such forest harvest residues are affected by many factors including age and size of the trees harvested, species composition, topography, market conditions, history and structure of the stand, stumpage prices, utilization standards, and climatic conditions (Hakkila 1989). Included in the noncommercial components may be tree sections with sweep and butt swell that are removed during log manufacturing at the landing or at the stump. This suggests the importance of taper equations that allow for estimation of log utilization under alternative utilization standards. For example, our experience in the Pacific Northwest USA indicates that logs being cut for export generally yield greater forest harvest residues than logs cut for domestic grades. Also, dead branches usually will not reach the landing during yarding or skidding; depending on pulp markets and logging system, many tops will not make it to the landing either. This also supports the use of taper equations so that stem biomass to different top diameters can be estimated.

Biomass yield is usually could be expressed in terms of kilograms per unit area. Alternative metrics are also useful. For example, development of biomass yield expressed in units of commercial volume removed, e.g. dry tons of biomass per thousand cubic feet or thousand board feet timber removed. Mass tables to supplement volume tables are also needed.

Besides biomass, bio-energy projects often require estimates of related physical or chemical properties. For example, for biomass utilization, specific gravity is an important variable. Specific gravity varies by species and by position in the tree (Bergman et al. 2010). For liquid fuel, the lignin percentage of tree components is important. Moisture content is another variable of interest. Moisture content affects transportation cost and, in the case of combustion, recoverable energy. The moisture content of biomass varies between species, between trees, within a tree, and during the season (Hakkila 1989).

The development of short-rotation woody crops such as willow (*Salix* sp.) as a source of bio-energy and bioproducts is growing in the USA and Europe; however, accurate biomass equations and estimation methods for clones and coppices are lacking (Tuskan & Rensema 1992). This is particularly true for very high density energy plantations, which often constitute very different tree populations. For example, mechanized cutting and chipping of poplars (*Populus* sp.) may be limited to trees less than 10 cm DBH.

Support for forest fuel analyses

In many regions, including the inland and dry forest types of the western USA, forest management is increasingly being shaped by wildfire and fuel management considerations. Wildfire and fuel concerns in these regions have highlighted the importance of crown biomass, its components, and its spatial distribution. Although crown characteristics are also relevant to wood quality considerations and growth modeling, there has been substantially less biomass data collected for crown components (branches, tops, foliage) relative to stem components. This is in part because crown components traditionally lacked commercial value and in part because the merchantable stem often accounts for the majority of aboveground tree biomass.

Central to the modeling of wildfire behavior and intensity is the disaggregation of crown biomass into size classes and the identification of connections among crowns. Size distributions (e.g. the 1-, 10-, or 100-hour fuel time lag classes used in the western USA) are important determinants of the sensitivity of fuel to changing moisture conditions. However, similar component size classes have not been used outside of the fire science domain and, as with crown characteristics in general, differences in their absolute or relative magnitudes are not always strongly associated with variations in tree diameter. Also important in the emerging generation of fluid dynamics-based fire models is the vertical distribution of fuel in the finer fuel classes, as this determines fuel proximity to the ground, fuel density within different portions of the crown, and the connectivity of fuel throughout the canopy.

Opportunities for improved biomass estimates

To address emerging biomass and carbon information needs it will be necessary to collect new sets and forms of biomass data, with an aim to developing a framework to integrate existing data and equations. Anticipating changes in future climate regimes, it is not a viable strategy to simply update or develop new biomass equations for individual geographic areas based on those areas' historic growing conditions. Instead, it will be necessary to study variations in the allometric forms of a species across broad climatic and edaphic gradients and to account for these variations in integrated biomass equations that are also sensitive to a greater number of measurable tree dimensions.

Advance understanding of biomass allometries

A considerable body of empirical data and descriptive information on tree biomass has been developed. For instance, Ter-Mikaelian and Korzukhin (1997) detail 803 diameter-based component equations collectively derived from thousands of trees. Many of these data were collected for the estimation of local weight tables, or to uncover regionally relevant empirical trends. This situation is similar to that found in the field of tree volume estimation. Although substantially more research and data have accumulated on volumetric analyses, these efforts have been aimed primarily at calibrating empirical models to account for local differences in regional height–diameter relationships, stem form quotients, or utilization/merchantability standards. Aside from the important and extensively documented impacts of stand density, relatively little research has been aimed at the fundamental questions of why and how height–diameter or form quotients vary among species or across species' ranges. Further research on the types and magnitudes of climatic controls on height–diameter ratios, stem taper, wood density, and other biomass characteristics is required to advance our ability to accurately estimate tree biomass within and across broad geographic areas. For example, the edaphic conditioning of allometric relationships documented by Lines et al. (2012) and Callaway et al. (1994) has two clear implications. First, new tree biomass data should be collected, in part, to assist in identifying the intraspecific variations in biomass along climatic gradients. Second, an emerging understanding of these variations should be utilized to improve the collection of additional data and to guide the integration of existing biomass data.

Accuracy evaluation and improvements

Because the performance of biomass estimation models varies, evaluations of their predictive abilities in large-scale biomass modeling and estimation are warranted. For estimating biomass and carbon sequestration, models' suitability and predictive abilities should be considered during the data collection phase (Lu 2006). A precise model requires a sound statistical formulation. Hence, it is desirable to have two independent data-sets; one for developing prediction equations (the fit data-set) and the other for evaluation of the methods (the validation data-set).

Given the high costs of collecting biomass data, statistical modeling and estimation techniques that allow one to borrow strength from related units or variables should always be considered. Common, though not unique, to tree-level biomass models is the multivariate nature of the regressand; often models are sought not only for total biomass but for crown biomass, stem biomass, etc. Moving beyond simple linear models and ordinary least squares algorithms to capture cross-correlations among these attributes is important for improving parametric and predictive accuracy. As well, it is often desired that the biomass components should add up to the total biomass. Parresol (2001), Sabatia et al. (2008), and Dong et al. (2014) describe models for ensuring additivity and multi-stage least squares estimation routines for calibrating them. However, the latter can restrict the class of models that can be entertained, and there is a need to advance these methodologies to allow for more flexible (nonlinear or nonparametric) model forms. In addition to modeling and capitalizing on cross-correlations among biomass component variables, similarities in model forms and coefficients across species or regions could be considered. For instance, Zapata-Cuartas et al. (2011) have used Bayesian methods to take advantage of commonalities across species and regularize individual species model coefficients.

The lack of local biomass equations and the uncertainty of estimates obtained from existing regional or global equations necessitate improved stochastic modeling and validation at a range of spatial scales (Temesgen et al. 2007). Uncertainty could vary by tree sizes. Chave et al. (2004) reported that the uncertainty involved in aboveground biomass estimation of a single tree ≥10 cm diameter is 47% of the estimated aboveground biomass: 31% from choice of allometric model and 16% due to measurement uncertainty, but this error averaged out at the stand level.

Integration of taper and other tree attributes

Tree volume is related to biomass which in turn is related to carbon. Compatible taper functions segmented polynomial models and other simpler taper equations have been used very successfully in the past to estimate stem volume for commercial species. Oderwald and Rayamajhi (1991) found that the taper-equation-based estimates of volume produced less than one-half of the mean-squared error obtained using volume equations. Biomass equations for a given species often differ over that species' range because of variation in tree component ratios and taper associated with changing site and stand conditions (Van Lear et al. 1986). Therefore, formulation of consistent taper equations and biomass equations using the same data should improve forest biomass and carbon estimates.

Given the costs associated with direct biomass measurements, it is also useful to capitalize on existing taper equations and wood density information to improve the efficiency of biomass sampling efforts. In particular, IS and control variate methods have been developed to provide accurate estimates of stem and whole-tree mass (e.g. see Van Deusen & Baldwin 1993; Van Deusen & Roesch 2011). These methods typically utilize regional tree taper equations at the selection and/or estimation phases of sampling to improve precision while disc- or core-based measurements ensure unbiasedness for the individual tree. The extensive literature and some of the past data collected on wood density could be similarly employed to improve the efficiency of data collection at the tree level, as well as to help identify trends in tree biomass associations at broad levels. For example, the US western wood density survey program collected data from over 30,000 trees of 15 commercial species across more than 4000 sites in the western USA (Maeglin & Wahlgren 1972). Given the important effects of stem wood density on total tree biomass, it is likely that considerable information on species, site-level, and geographic variations in biomass could be recovered from these legacy data-sets.

Finally, future biomass and carbon modeling efforts should draw on recent advances in biomechanical models of tree form. Allometric scaling theory derived from first principles could be used to guide the development of model forms for the allocation of tree biomass in order to improve robustness over what could be achieved from empirical data alone. For example, the quarter-power scaling theory of West et al. (1999) implies specific constraints on the distribution of tree biomass, while the work of Mäkelä and Valentine (2006) establishes specific tree-level covariates besides DBH that are needed for component biomass estimation.

Combining data sources

Consistent methods for data collection and analysis are crucial in developing standard biomass estimation procedures. While forest biomass estimation has been done using stand-based inventory data and remotely sensed data, little research has been dedicated toward integrating these two approaches. In assessing the accuracy of regional LiDAR-based biomass estimation using a simulation approach, Ene et al. (2012) found approximately 1.8 times larger estimated standard errors for the airborne laser scanning estimates compared to the ground-based inventories; however, when LiDAR and ground data are combined the accuracy, measured in terms of root mean squared error, was improved by 59%. Traditional field measurement methods are considered to be more accurate but this information may change very quickly and become outdated because of the dynamic nature of forest environments. More frequent satellite remote sensing data may provide a supplement or a substitute (Main-Knorn et al. 2011; Goerndt et al. 2013). More generally, the integration of multiple data sources and advanced technology will become critical for estimating forest biomass and carbon accurately over time (Temesgen et al. 2007). Moreover, because of the independence of the data-sets, performance of the models could be cross-verified by using stand-based inventory data as validation data-sets for models developed from remotely sensed data.

Zapata-Cuartas et al. (2011) proposed a Bayesian approach for estimating aboveground tree biomass which outperformed the classical statistical approach of least square regression. They found similar significant values in the estimation of parameters using a sample size of six trees compared to 40–60 trees in the classical approach.

Discussion

The global issue of climate change, trade in carbon credits, and the interest in reducing fossil fuel carbon dioxide emission by using forest biomass for energy production has increased the importance of forest biomass quantification in recent years. Different national and international reports have presented the amount of carbon sequestered by forests, e.g. IPCC (2007) reports that forests contain about 80% of aboveground and 40% of belowground carbon stocks. Additionally, it has been reported that the amount of carbon stored in dry wood is approximately 50% by weight. All these numbers are based on allometric equations developed to relate forest biomass to easily measurable attributes of forest stands or trees. Thus, an accurate estimate of carbon stocks requires improved and consistent methods for forest biomass quantification.

The most common biomass modeling approach selects some trees for destructive sampling and weighs their components. Then regression models are fitted to describe the relationship of biomass to species and size variables. However, there are questions that need to be answered before we apply these models at different spatial scales. The questions include would the species composition and stand density at new sites be similar to those sites where sample trees were selected.

In this regard, caution should be taken in developing and evaluating methods for estimating aboveground biomass and its components. Application of more sophisticated fitting methods such as seemingly unrelated regression and nonlinear three-stage least squares and their wide use are necessary. Additionally, assessing the effect of sub-sampling design and sample size in estimating component biomass, the uncertainty involved in the estimation, scope of inference of the models, and accuracy of biomass conversion and expansion factors are important subjects for biomass studies and methodological research.

We foresee future work in several directions. Examination of Bayesian methods to preserve the covariance among biomass components is warranted. In addition to considering the below ground biomass, there is a need to develop comprehensive biomass estimation methods that account for differences in site and stand density and improve forest biomass modeling and validation at a range of spatial scales.

Because biomass estimation is a complex process, when possible, we should make use of already available resources such as wood density and forest inventory databases. Combining different data-sets for model development and using independent data-sets for model verification will offer opportunities to improve biomass estimation. Focus should also be made on belowground biomass estimation to accurately estimate the full forest contribution to carbon sequestration. The increasing demand for credible biomass and carbon estimates at widely varying spatial scales requires circumspect evaluations of the sampling and modeling practices used for biomass equation development. The accuracy of forest biomass estimates can be improved, including estimates of the aboveground live tree components as well as the dead and nontree fractions.

Acknowledgments

We gratefully acknowledge the support provided by the Forest Inventory and Analysis program, Pacific Northwest Research Station, United States Forest Service (USFS). The authors also acknowledge the support of the USFS national office and its regional program directors in instituting a review of FIA biomass sampling and modeling protocols with the collaboration of academic and private-sector cooperators.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Baldwin VC. 1987. Green and dry-weight equations for aboveground components of planted loblolly pine trees in the west gulf region. South J Appl For. 11:212–218.

Barclay H, Pang P, Pollard D. 1986. Aboveground biomass distribution within trees and stands in thinned and fertilized Douglas-fir. Can J For Res. 16:438–442.

Bergman R, Cai Z, Carll CG, Clausen CA, Dietenberger MA, Falk RH, Frihart CR, Glass SV, Hunt CG, Ibach RE, et al. 2010. Wood handbook, wood as an engineering material. General Technical Report FPL-GTR-190. Madison (WI): U.S. Department of Agriculture, Forest Service, Forest Products Laboratory; 508 p.

Bockheim J, Lee S. 1984. Biomass and net primary production equations for thinned red pine plantations in central Wisconsin. University of Wisconsin Coll. of Agric. For. Res. Notes 256. Wisconsin.

Bouffier LA, Gartner BL, Domec JC. 2003. Wood density and hydraulic properties of ponderosa pine from the Willamette Valley vs. the Cascade Mountains. Wood Fiber Sci. 35:217–233.

Callaway RM, DeLucia EH, Schlesinger WH. 1994. Biomass allocation of montane and desert ponderosa pine: an analogy for response to climate change. Ecology. 75:1474–1481.

Canavan SJ, Hann DW. 2004. The two-stage method for measurement error characterization. For Sci. 50:743–756.

Chave J, Condit R, Aguilar S, Hernandez A, Lao S, Perez R. 2004. Error propagation and scaling for tropical forest biomass estimates. Philos Trans Royal Soc B. 359: 409–420.

Cochran P, Jennings J, Youngberg C. 1984. Biomass estimators for thinned second-growth Ponderosa pine trees. Res. Note PNW-415. Portland (OR): US Department of Agriculture, Forest Service, Pacific Northwest Forest and Range Experiment Station.

Domke GM, Woodall CW, Smith JE. 2011. Accounting for density reduction and structural loss in standing dead trees: implications for forest biomass and carbon stock estimates in the United States. Carbon Balance Manag. 6:1–14.

Dong L, Zhang L, Li F. 2014. A three-step proportional weighting system of nonlinear biomass equations. Advance online publication. For Sci. 60.

Ene LT, Næsset E, Gobakken T, Gregoire TG, Ståhl G, Nelson R. 2012. Assessing the accuracy of regional LiDAR-based biomass estimation using a simulation approach. Remote Sens Environ. 123:579–592.

Fang JY, Wang ZM. 2001. Forest biomass estimation at regional and global levels, with special reference to China's forest biomass. Ecol Res. 16:587–592.

Goerndt ME, Monleon VJ, Temesgen H. 2013. Small area estimation of county-level forest attributes using ground data and remote sensed auxiliary information. Forest Sci. 59:536–548.

Gonzalez JS. 1990. Wood density of Canadian tree species. Edmonton: Information Report – Northwest Region, Forestry Canada–NOR-X-315.

Gower S, Haynes B, Fassnacht K, Running S, Hunt EJ. 1993. Influence of fertilization on the allometric relations for two pines in contrasting environments. Can J For Res. 23:1704–1711.

Grier C, Lee K, Archibald R. 1984. Effect of urea fertilization on allometric relations in young Douglas-fir trees. Can J For Res. 14:900–904.

Grigal DL, Kernick L. 1984. Generality of black spruce biomass estimation equations. Can J For Res. 14:468–470.

Hakkila P. 1989. Utilization of residual forest biomass. Heidelberg: Springer-Verlag Berlin; 568 p.

Heath LS, Hansen MH, Smith JE, Smith WB, Miles PD. 2008. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. In: McWilliams W, Moisen G, Czaplewski R, comps. Forest Inventory and Analysis (FIA) symposium. Proc. RMRS-P-56 CD. Fort Collins (CO): U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. [CD–ROM].

Hollinger DY. 2008. Defining a landscape-scale monitoring tier for the North American carbon program. In: Hoover CM, editor. Field measurements for forest carbon monitoring. Durham: © Springer Science + Business Media B.V. 2008; p. 3–16.

IPCC. 2007. Climate change 2007: synthesis report. In: Core Writing Team, Pachauri RK, Reisinger A, editors. Contribution of working groups I, II and III to the fourth assessment report of the intergovernmental panel on climate change. Geneva (Switzerland): IPCC; 104 p.

Jenkins J, Chojnacky D, Heath L, Birdsey R. 2003. National-scale biomass estimators for United States tree species. For Sci. 49:12–35.

Lines ER, Zavala MA, Purves DW, Coomes DA. 2012. Predictable changes in aboveground allometry of trees along gradients of temperature, aridity, and competition. Global Ecol Biogeogr. 21:1017–1028.

Lu D. 2006. The potential and challenge of remote sensing-based biomass estimation. Int J Remote Sens. 27:1297–1328.

Maeglin RR, Wahlgren HE. 1972. Western wood density survey, Report No. 2. Madison (WI): USDA Forest Service Research Paper FPL-183.

Main-Knorn M, Moisen GG, Healey SP, Keeton WS, Freeman EA, Hostert P. 2011. Evaluating the remote sensing and inventory-based estimation of biomass in the western Carpathians. Remote Sens. 3:1427–1446.

Mäkelä A, Valentine HT. 2006. Crown ratio influences allometric scaling trees. Ecology. 87:2967–2972.

Nave LE, Vance ED, Swanston CW, Curtis PS. 2010. Harvest impacts on soil carbon storage in temperate forests. Forest Ecol Manag. 259:857–866.

Oderwald RG, Rayamajhi JN. 1991. Biomass inventory with tree taper equations. Bioresour Technol. 36:235–239.

Ozcelik R, Eraslan T. 2011. Two-stage sampling to estimate individual tree biomass. Turk J Agric For. 36:389–398.

Parresol BR. 2001. Additivity of nonlinear biomass equations. Can J For Res. 31:865–878.

Pollard D. 1972. Above-ground dry matter production in three stands of trembling aspen. Can J For Res. 2:27–33.

Ritchie MW, Zhang J, Hamilton TA. 2013. Aboveground tree biomass for *Pinus ponderosa* in Northeastern California. Forests. 4:179–196.

Sabatia CO, Lynch TB, Will RE. 2008. Tree biomass equations for naturally regenerated shortleaf pine. South J Appl For. 32:163–167.

Standish JT, Manning GH, Demaerschalk JP. 1985. Development of biomass equations for British Columbia tree species. Information Report. BC-X-264. Victoria: Canadian Forest Service, Pacific Forest Resource Center; 47 p.

Temesgen H, Goerndt ME, Johnson GP, Adams DM, Monserud RA. 2007. Forest measurement and biometrics in forest management: status and future needs of the Pacific Northwest USA. J For. 105:233–238.

Ter-Mikaelian MT, Korzukhin MD. 1997. Biomass equations for sixty-five North American tree species. For Ecol Manag. 97:1–24.

Tobin B, Nieuwenhuis M. 2007. Biomass expansion factors for Sitka spruce (*Picea sitchensis* (Bong.) Carr.) in Ireland. Eur J For Res. 126:189–196.

Tuskan G, Rensema T. 1992. Clonal differences in biomass characteristics, coppice ability, and biomass prediction equations among four Populus clones grown in eastern North Dakota. Can J For Res. 22:348–354.

Valentine HT, Tritton LM, Furnival GM. 1984. Subsampling trees for biomass, volume, or mineral content. For Sci. 30:673–681.

Van Deusen P, Baldwin VC, Jr. 1993. Sampling and predicting tree dry weight. Can J For Res. 23:1826–1829.

Van Deusen P, Roesch FA. 2011. Sampling a tree for total volume, biomass, and carbon. J For. 109:131–135.

Van Lear DH, Taras MA, Waide JB, Augspurger MK. 1986. Comparison of biomass equations for planted vs. natural loblolly pine stands of sawtimber size. For Ecol Manage. 14:205–210.

Ver Hoef J, Temesgen H. 2013. A comparison of the spatial linear model to nearest neighbor (k-NN) methods for forestry applications. PLoS ONE. 1:1–11.

West GB, Brown JH, Enquist BJ. 1999. A general model for the structure and allometry of plant vascular systems. Nature. 400:664–667.

Williams RA. 1989. Use of randomized branch and importance sampling to estimate Loblolly pine biomass. South J Appl For. 13:181–184.

Zapata-Cuartas M, Sierra CA, Alleman L. 2011. Probability distribution of allometric coefficients and Bayesian estimation of aboveground tree biomass. For Ecol Manag. 227:173–179.

Zhou X, Hemstrom MA. 2009. Estimating aboveground tree biomass on forest land in the Pacific Northwest: a comparison of approaches. Res. Pap. PNW-RP-584. Portland (OR): U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station; 18 p.