Indian Forester, 142 (1) : 62-67, 2016
http://www.indianforester.co.in
ISSN No. 0019-4816 (Print)
ISSN No. 2321-094X (Online)

ERROR PROPAGATION IN FOREST BIOMASS ASSESSMENT

M. SIVARAM, S. SANDEEP[1] AND H. MATIEU[2]

Southern Regional Station,
National Dairy Research Institute, Bangalore (Karnataka)
E-mail: sivaram.ndri@gmail.com

ABSTRACT

Forest biomass is the basis for the estimation of carbon storage and emission due to forestry sector. Though the total forest biomass includes aboveground and belowground biomass, this paper deals with the issues related to aboveground biomass. The total aboveground biomass is estimated through a number of variables measured across various components of trees using non-destructive methods. The techniques employed range from simple measuring tape to regression models to satellite imageries. The total error in biomass estimates is the sum of errors in the variables propagated in a hierarchical fashion. The knowledge of prediction errors helps to know the quality of biomass and subsequently bio-energy and carbon estimates. In this paper, various sources of error in biomass estimation, error quantification and error propagation are discussed. The sources of error include tree measurements, sampling strategy, choice of an allometric model and satellite imageries. In South Asia, the standard errors of co-efficient of biomass equations and R^2 are often depicted as indicators for the quality of volume and biomass equations. The studies on error propagation in biomass estimates are scarce. Monte Carlo analysis, Pseudo-meta-analysis and Bayesian model averaging have been investigated to address the issues of error propagation in biomass estimation. Among these Bayesian model averaging appears to be a promising technique.

Key words: Biomass, Error propagation, Allometric equation, Monte Carlo analysis, Pseudo-meta-analysis and Bayesian model

Introduction

Gases that trap heat in the atmosphere are called greenhouse gases. Carbon dioxide (CO_2), methane (CH_4), nitrous oxide (N_2O) and fluorinated gases are the major greenhouse gases. Deforestation and forest degradation account for nearly 20% of global greenhouse gas emissions (Sukhdev *et al.*, 2015). However, forests form an important aspect of active carbon pool as they account for 60% of terrestrial carbon storage (Wilson and Daff, 2003). Hence, forest ecosystems act as both source and sink of carbon and thus play a crucial role in global carbon cycles. Forest biomass estimates are essential for the estimation of carbon sequestration and emission potential of forests.

The accurate measurement of volume or biomass of trees is a destructive process. However, researchers estimate volume or biomass following non-destructive methods. This will involve determination of actual volume or biomass of a sample set of trees and relating them to non-destructive measures like diameter at breast-height (dbh) and height of trees through regression analysis. Determination of volume of any specified part of the tree such as stem or branch is usually achieved by cutting the tree part into logs and making measurements on the logs. The volume of individual logs may be calculated by using one of the formulae like Smalian's formula, Huber's formula and Newton's formula depending on the measurements available (Jayaraman, 1999). Biomass is usually expressed in terms of dry weight of component parts of trees such as stem, branches and leaves. Biomass of individual trees are determined destructively by felling the trees and separating the component parts like main stem, branches, twigs and leaves The data collected from sample trees on their volume or biomass along with the dbh and height of sample trees are utilized to develop prediction equations through regression techniques. It is evident that the regional and global level biomass estimates that use remote sensing and statistical techniques ultimately depend upon the field measurements of individual trees.

In the tree allometric database developed for South Asia, out of 4456 equations only 1.5% of the equations describes carbon as output variable (Sandeep *et al.*, 2014). Of the remaining equations 48% are biomass equations and 40% are volume equations. Most of the biomass allometric equations in tropics are from

A review on various sources of error in forest biomass estimation and error propagation methods recommended the use of Bayesian model averaging in improving the biomass estimates.

[1]Kerala Forest Research Institute, Peechi
[2]Food and Agricultural Organization, Rome

ecological studies, which are limited to small tree diameter classes. Therefore, conversion of volume to biomass and then to carbon is inevitable to obtain national and regional level biomass and carbon estimates.

Biomass calibration

Approach by individual studies

The studies in limited areas have used tree biomass regression equations (Chave *et al.*, 2005; Thomas *et al.,* 2013). This involves estimating the biomass per average tree of each dbh class of the stand table, multiplying by the number of trees in the class, and summing across all classes (Brown, 1997). Light Detection And Ranging (LiDAR) data and remote sensing images are employed to calibrate allometric relations to estimate biomass at large spatial scales (Case and Hall, 2008; Colgan *et al.*, 2013).

National Level approach in South Asia

In South Asia, the national level biomass density is estimated from volume over bark per hectare by first estimating the biomass of the inventoried volume with the use of wood density estimates and then "expanding" this value to take into account the biomass of the other aboveground components (Brown, 1997). The inventory tree volume are based on tree volume allometric equations.

Volume and biomass estimation studies have several sources of error and estimating their extent in biomass estimates is cumbersome (Chave *et al.,* 2005; Pare *et al.*, 2013). However, the knowledge of prediction errors aids in formulating error budgets and in knowing the quality of biomass and carbon estimates which affect greenhouse gas inventory. In this paper, we discuss about various sources of error in the aboveground biomass estimation, error quantification and error propagation.

Sources of errors in Forest Biomass Estimates

Error is the difference between reality and representation of reality. In statistical jargons, it is the absolute difference between observed and estimated value. The statistical error is also called as uncertainty. Many times an estimate is obtained from a function of several variables. The error propagation is the effect of the ' errors on the uncertainty of a based on them (http://en.wikipedia.org /wiki/Propagation_of_uncertainty). Such error propagation issues arise in the inventory of forest.

Forest biomass includes the aboveground and belowground living mass. However, in this paper, we restrict to aboveground biomass which is a function of a number of variables measured across various components of trees (stem, leaves, branches etc.) using techniques ranging from simple measuring tape to regression models to satellite imageries. The total error in biomass estimates is the sum of errors in the variables.

Case and Hall (2008), in their review, have listed a potential biological, environmental, and data modelling factors affecting tree allometry and related estimation of above ground tree biomass and total volume. They include i) Biological factors: Tree species, tree size range, tree structure and height, wood density, stand composition and structure, slivicultural interventions and biotic disturbance; ii) Environmental factors: Eco-region, biome, range, extent, climate, elevation, site quality and resource availability and iii) Data modelling factors: field data collection, measurement protocols, sampling strategies and modelling process ie. Statistical models and issues related to calibration of biomass estimates from sample plot to regional and global level.

Data Measurements

Tree level measurements

Diameter and height

Tree volume and biomass estimation using non-destructive methods rely on diameter and height measurements. The dbh is very easy to measure, and common attribute used in developing allometric equation. In the database on tree allometric database for South Asia (Sandeep *et al.,* 2014), about 76% of the equations used dbh as input variable. In 12% of the equations gbh was used as input variable. Only 4% of the equations used height as input variable.

The inclusion of height in the allometric equation along with dbh improves the accuracy of the biomass estimate (Rizvi *et al.*, 2008). However, tree height is very difficult to measure especially in dense forests and thus large measurement errors could occur. Due to practical difficulties in measuring tree height, height prediction equation is often used in the place of original height

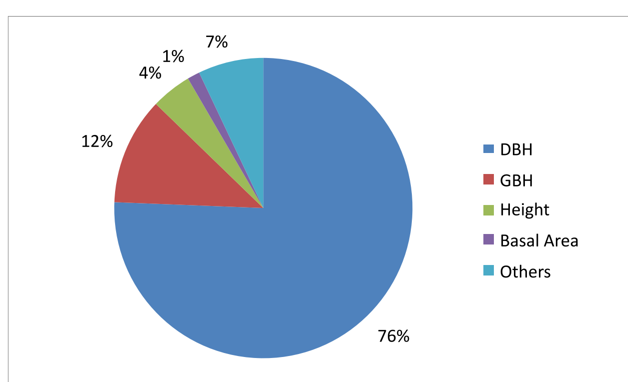

Fig. 1: Statistics on input variables used in allometric equations in South Asia

(Molto *et al.*, 2012). In the tree allometric database for South Asia, there are about 253 height predictive equations which is about 0.06% of the total number of equations.

Hunter *et al.* (2013) found that the precision of height measurements of individual trees led to a mean error of 16% in the estimate of individual tree biomass and 5 to 6% error in overall plot level biomass. Chave *et al.* (2004) estimated 48% and 78% of error at tree level for the trees greater than 10 cm diameter and less than 10 cm diameter respectively.

Wood specific gravity

Williamson and Wiemann (2010) have identified various sources of error in estimating wood specific gravity: i) extracting wood samples that are not representative of tree wood, ii) differentiating wood specific gravity from wood density, iii) drying wood samples at temperatures below 100°C and the resulting moisture content complications, and iv) improperly measuring wood volumes. Chave *et al.* (2004) found that wood specific gravity decreased significantly across different diameter classes.

Sampling strategy

Tree diversity in tropical forests is very rich. Thus, representativeness of a network of sufficient number of plots of appropriate size over a vast diverse forest tropical landscape has been a difficult issue to resolve with the limited resources. Chave *et al.* (2004) estimated within plot error of 5% to 16% for the plot size of 1 ha to 0.1 ha. The among-plot error uncertainty was 11%. Even in a limited land-scape, often trade-off is made between the accuracy and resource constraints.

Model Development

Choice of an allometric model

The choice of allometric model is the most important source of error in volume and biomass quantification. Chave *et al.* (2004) found the error on this account reaching up to 22% of the final estimate. Various tree allometric models reported so far belong to linear, polynomial or non-linear category (Rizvi and Khare, 2006; Rizvi *et al.*, 2008; Picard *et al.*, 2012). Though simple linear models are most frequently used for tree growth, they suffer from the problem of 'negative estimation' of tree size exclusively for the lower range of the explanatory variable (Ajith *et al.*, 1999). The sigmoid shape functions, belonging to the non-linear models serve the purpose of a more realistic and process based tool for allometric modelling. An insight into the prediction capabilities of these sigmoid functions revealed that although they solved the problem of 'negative estimation' associated with linear models, they suffered from the problem of 'constant estimation' of tree size exclusively for the higher range of the explanatory variable (Srivastava and Ajit, 2002).

An analysis of the tree volume and allometric database for South Asia reveals that the simple linear, quadratic and power equations relating dbh alone contributed to about 52%. The simple linear equation relating product of square of dbh and height was found by 7.3%. About 12 functional forms contributed to the development of as many as 71% of the total number of equations (Table 1).

Table 1: Statistics on the use of various functional forms in the allometric equations for South Asia

S.No.	Functional form of equation	%
1	Y = a+b*X	18.44
2	Y = a*X^2+b*X+c	17.09
3	ln(Y) = a+b*ln(X)	16.60
4	Y = a+b*X^2*Z	7.37
5	ln(Y) = a+b*ln(X)+c*ln(Z)	5.56
6	vY = a+b*X	2.48
7	ln(Y) = a+b*ln(X^2*Z)	2.21
8	Y = a*X^b	2.17
9	Y = a*X	1.75
10	Y = a+b*ln(X)	1.67
11	Y = a+b*X^2	1.60
12	Y = a+b*X+c*Z+d*W	1.06
13	Other forms	22.00

Y-output variable such as volume/biomass/carbon ; X-dbh(cm); Z-height (m)

Model selection criteria

To choose the best biomass regression model from among the several candidate models, goodness of fit statistics such as R^2, Furnival Index, mean square error, root mean square error (RMSE), mean absolute percentage error, asymptotic error, Wald confidence interval, chi-square tests and information criterion such as Akaike information criterion are used for model selection. Residual plots and autocorrelation plots have also been used to aid deciding the best model (Rizvi *et al.*, 2008). The authors have used more than one criteria. Often trade-off is made for selecting the final model (Rizvi *et al.,* 2008). In the allometric database for South Asia (Sandeep *et al.*, 2014), R^2 is available only for 33.8% of the equations. Other fit statistics such as RMSE is available for 13.12% of the equations.

Model testing and validation

Once the best model is chosen, it used to be tested on independent data which is set apart before model building. Once the testing is satisfactory the model is validated again with the data collected afresh from the field (Rizvi and Khare, 2006). However, in South Asia such testing and validation procedures are rarely done. Use of re-sampling procedures is seen in the literature to assess the performance of the chosen model. However, in South Asia, the standard errors of co-efficient of regression equations and R^2 are often taken as indicators for the quality of the biomass equations (Rizvi *et al*., 2008).

LiDAR and Satellite imageries

LiDAR data is used for obtaining the height information (Nguyet *et al.*, 2012). Satellite imageries are used for computing the crown projected area. Both the measurements are used for estimating the biomass at large spatial scales. However, errors occur in acquiring and processing of LiDAR and satellite imageries (Gurdak *et al.*, 2009; Nguyet 2012; Hunter *et al.*, 2013). To translate remote sensing observations to forest biomass estimates field data are relied (often called “ground-truth”) for calibration and validation. The algorithms and models used for calibration are also one of the sources of error.

Error quantification and Propagation

Boostraping

Monte carlo simulation

The forest biomass estimation involves measurements to be made on trees in the field to regional levels (spatial scale) using variety of techniques ranging from simple measuring tape to satellite imageries in a hierarchical fashion. Considering such complexities, Monte Carlo analysis is recommended to quantify the uncertainty of forest biomass estimates. Monte Carlo Simulation is based on the repetition of many individual model realizations with each realization using a randomly constructed range of estimates of the parameters (Larocquea *et al.*, 2008; Schade and Wiesenthal, 2011; Hunter *et al.*, 2013). The range may be based on expert knowledge or possible extent of errors associated with the estimates. The models are then aggregated into a probability distribution which shows the variation in the output.

Pseudo-meta-analysis approach

In this approach, allometric equations from the literature are combined and new set of equations are produced. The pseudo-data set are created using the following steps (i) Create an original biomass population that could produce a fit with an R^2 as seen in published results by inputting the original equation and then creating datasets with low to high levels of variance ii) Re-fit original equation with pseudo-data to calculate an R^2 of the fit and select the pseudo-dataset with the R^2 most closely matching the published R^2 and iii) From each 'population' sub-sample with replacement the same number (*n*) as used in the original work to bootstrap the error. The approach was tested against a large dataset of raw dbh and biomass numbers from the Canadian ENFOR data (Wayson *et al.*, 2013).

Bayesian model averaging (BMA)

Data analysts typically select a model from some class of models and then proceed as if the selected model had generated the data (Hoeting *et al.*, 1999). This approach ignores the uncertainty in model selection, leading to over con?dent inferences and decisions that are more risky than one thinks they are. Bayesian model averaging provides a coherent mechanism for accounting for this model uncertainty (Molto *et al.* 2012). While estimating biomass uncertainty arises as there are number of competing models to be compared and chosen. Further, BMA method can group models, *e.g.* by climate zone. The BMA method has the advantage of generating weighted estimates. The weighting can follow the stratification used for inventory or some prior knowledge and improve the estimates to include a greater range of variation in diameters of samples or species (Jara and Henry, 2013).

Conclusions

The quality of allometric equations varies due to several sources of error. The sources range from tree measurements to model selection to techniques used for spatial extrapolation. In South Asia, the tree volume and biomass allometric models used for the estimation of aboveground biomass estimation suffer from certain shortcomings. Although it is necessary to take into account the different sources of error in the calculation of the total error, this is rarely done. Monte Carlo analysis, Pseudo-meta-analysis and Bayesian model averaging have been explored to tackle the problem of error propagation in biomass estimation. Among these Bayesian model averaging seems to be a promising technique.

As regards South Asia, it is suggested to update the existing database of tree volume and biomass allometric equations developed by Kerala Forest Research Institute with the support of FAO and undertake meta-analysis (including quality assessment of existing equations) for further research. Since proper allometric work requires lot of effort and time, adequate allocation of financial and human resources and capacity building will improve the accuracy of the biomass estimates significantly.

Acknowledgement

The authors gratefully acknowledge the reviewer for valuable comments which have improved the paper.

ou toek=kk eY;kdu e =kfV io/u
,e- flokjke- ,l- lnhi ,o ,p- efr;
lkjk'k

okfudh lDVj d dkj.k dkcu Hk.Mkj.k vkj mRltu d vkdyu d fy, ou toek=kk vk/kj gA ;|fi dy ou toek=kk e HkE;ifjd vkj Hkfe d uhp dh toek=kk 'kkfey g] bl 'kk/i=k e HkE;ifjd toek=kk l lcf/r fo"k;k ij fopkj fd;k x;k gA xj&fouk'kd fof/;k dk mi;kx djd o{kk d fofHkUu ?kVdk e ekfir db ifjofr;k d tfj, dy HkE;ifjd toek=kk vkdfyr dh xbA rduhdk e lk/kj.k ekiu Vi l lekJ;.k ekMyk l lVykbV betjh rd i;Dr dh xb gA toek=kk vkdyuk e dy =kfV J.khc¼ rjhd l iof/r ifjofr;k e =kfV;k dk ;kx gA Hkfo";dFku =kfV;k dh tkudkjh toek=kk dh x.koRrk dk tkuu vkj ckn e to Åtk vkj dkcu vkdyuk e lgk;rk djrh gA bl 'kk/i=k e toek=kk vkdyu e =kfV d fofHkUu lkrk] =kfV ifjek.ku vkj =kfV ic/u ij fopkj fd;k x;k gA =kfV d lkrk e o{k eki] ifrp;u j.kuhfr] ,d lki{kferh; ekMy dh ilUn vkj lVykbV betjh 'kkfey gA nf{k.k ,f'k;k e toek=kk lehdj.kk d x.kkd dh ekud =kfV vkj R^2 dk ik;% vk;ru vkj toek=kk lehdj.kk dh x.koRrk d fy, fun'kdk d :i e fpf=kr fd;k tkrk gA toek=kk vkdyu e =kfV io/u d fo"k;k d lek/ku d fy, मोन्टो कार्लो fo'y"k.k] स्यूडोन्मेटा fo'y"k.k vkj बेसियन मॉडल ,ojftx ,d vk'kktud rduhd fn[kkb iMrh gA

References

Ajith, Srivastava P.N., Gupta V.K. and Solanki K.R. (1999). Tree growth modeling: Problem of negative and constant estimation size. *Agro Forestry News Letter*, 11(1): 4.

Brown S. (1997). *Estimating biomass and biomass change of Tropical Forests: a Primer.* FAO Forestry Paper – *134.* FAO, Rome (also available at http:// www.fao.org/ docrep/ w4095e/w4095e00.HTM).

Case B.S and Hall R.J. (2008). Assessing prediction errors of generalized tree biomass and volume equations for the boreal forest region of west-central Canada, *Can. J. For. Res*., 38: 878-889.

Chave J., Condit R., Aguilar S., Hernandez A., Lao S. and Perez R. (2004). Error propagation and scaling for tropical forest biomass estimates, *The Royal Society*, 359: 409-420.

Chave J., Andalom C., Brown S.,Cairns M.A., Chambers J.Q., Eamus D., Folster H., Fromard F., Higuchi N., Kira T., Lescure J.P., Nelson B.W., Ogawa H., Puig H., Rie´ra B., and Yamakura, T. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests, *Oecologia*, 145: 87–99.

Colgan M.S., Asner G.P. and Swemmer T. (2013). Harvesting tree biomass at the stand level to assess the accuracy of field and airborne biomass estimation in savannas, *Ecological Applications*, 23(5):1170–1184.

Gurdak J.J., Qi S.L., and Geisler M.L., (2009). *Estimating prediction uncertainty from geographical information system raster processing—A user's manual for the Raster Error Propagation Tool (REPTool): U.S. Geological Survey Techniques and Methods,* 11–C3, 71 p.

Hoeting, A., Madigan. D., Raftery, A.E. and Volinsky, C.T. (1999). Bayesian Model Averaging: A Tutorial, *Statistical Science,* 14 (4): 382–417.

Hunter M.O., Keller M., Victoria D. and Morton D.C. (2013).Tree height and tropical forest biomass estimation, *Biogeosciences*, 10: 8385–8399.

Jara M.C. and Henry M. (2013). *Proceedings of the regional technical workshop on Tree Volume and Biomass Allometric Equations in South and Central America*, 21-24 May 2013, UN-REDD MRV Report 12, Turrialba, Costa Rica.

Jayaraman K. (1999). *A Statistical Manual for Forestry Research. Regional Office for Asia and Pacific*, FAO, Bangkok. ftp://ftp.fao.org/docrep/ fao/003/X6831E/X6831E00.pdf .

Larocquea G.R., Bhattib J.S., Boutina R. and Chertovc O. (2008). Uncertainty analysis in carbon cycle models of forest ecosystems: Research needs and development of a theoretical framework to estimate error propagation, *Ecological Modelling*, 219: 400–412.

Molto Q., Rossi V. and Blanc L. (2012). Error propagation in biomass estimation in tropical forests, Methods in Ecology and Evolution, *British Ecological Society*: 1-9.

Nguyet D.A. (2012). *Error propagation in carbon estimation using the combination of Airborne LiDARr data and very high resolution Geo-eye satellite imagery in Ludhikhola Watershed*, Gorkha, Nepal.

Pare D., Bernier P., Lafleur B., Titus B.D., Thiffault E., Maynard D.G., and Guo X. (2013). Estimating stand-scale biomass, nutrient contents, and associated uncertainties for tree species of Canadian forests, *Can. J. For. Res*., 43: 599–608.

Picard N., Saint-André L., and Henry M. (2012). *Manual for building tree volume and biomass allometric equations: from field measurement to prediction*. Food and Agricultural Organization of the United Nations, Rome, and *Centre de Coopération Internationale en Recherche Agronomique pour le Développement*, Montpellier, 215 pp.

Rizvi R.H. and Khare D. (2006). Prediction models for timber weight of Poulus deltoids planted on farmelands in Haryana. *Indian J. of Agroforestry*, 8(1): 77-85.

Rizvi R.H., Diwakar K. and Dhillon R.S. (2008). Statistical models for aboveground biomass of *Populus deltoides* planted in agroforestry in Haryana. *Tropical Ecology*, 49(1): 1-8.

Sandeep S., Sivaram M., Henry M., and Birigazzi L. (2014). *Inventory of volume and biomass tree allometric equations for South Asia.* KFRI, Peechi, India, Food & Agriculture Organization of the United Nations, Rome, Italy. (database is also available at http://www.globallometree.org/database/)

Schade B., and Wiesenthal T. (2011). Biofuels: A model based assessment under uncertainty applying the Monte Carlo method. *Journal of Policy Modeling*, 33(1): 92-126.

Srivastava P.N. and Ajit (2002). Illogical estimation intricacies associated with tree growth modelling and their alternative solutions. In: *Abstracts of fifth annual conference of the International Academy of Physical Sciences on Interdisciplinary trends in Physical Sciences" and symposia on "The Role of Information Technology and its Implications"* held at Bundelkhand University, Jhansi during 7- 9 April, 2002. Pp 31-32.

Sukhdev P., Prabhua R., Kumara P., Bassic A., Patwa-Shaha W., Entersa T , Labbatea G. and Julie G. (2015). *REDD+ and a Green Economy: Opportunities for a mutually supportive relationship*. UN-REDD Programme Policy Brief 01.

Thomas T.P., Rugmini P. and Balagopalan M. (2013). *Carbon storage potential of different age teak plantations in Kerala.* KFRI Research Report No. 458. Kerala Forest Research Institute, Peechi.

http://en.wikipedia.org/wiki/Propagation_of_uncertainty (accessed on 01 August 2014).

Wayson C., Johnson K., Cole J., and Ung C.H. (2013). Estimating uncertainty of allometric biomass equations using a pseudo-meta-analysis approach In: *Proceedings of the regional technical workshop on Tree Volume and Biomass Allometric Equations in South and Central America (*M.C. Jara, and M.Henry, Eds.), 21-24 May 2013, UN-REDD MRV Report 12, Turrialba, Costa Rica.

Williamson G.B. and Wiemann M.C. (2010). Measuring wood specific gravity...Correctly. *Am J Bot.*, 97(3):519-24.

Wilson B.R.S.A. and Daff J.T. (2003). *Australia's State of the Forests report*. Department of Agriculture, Fisheries and Forestry, Govt. of Australia.