XIV WORLD FORESTRY CONGRESS, Durban, South Africa, 7-11 September 2015

ALLOMETRIC EQUATIONS AT NATIONAL SCALE FOR ESTIMATING TREE AND FOREST BIOMASS IN VIET NAM

Bao Huy[1]

[1]Assoc.Prof.Dr., Tay Nguyen University, Vietnam; Email: baohuy.frem@gmail.com

Abstract

Natural forest types of Viet Nam thus no tools have been developed to estimate its carbon stock at national level. In the context of climate change mitigation through forestry sector and REDD+ mechanism, developing allometric equations with a dataset of tree biomass at national level could provide improved carbon stock estimates, with a known accuracy. This study developed a set of models to estimate above ground biomass (AGB) in evergreen broadleaf forests (EBLF). The dataset for EBLF included 860 trees located in five eco-regions of Viet Nam. DBH (diameter at breast height), H (tree height), WD (wood density) were used as input variables. The modelling was performed by applying non-linear mixed effect models and power models on residuals, with or without random effects of eco-regions or environment variable to models' parameters. Indicators to select the best models were the Akaike information criterion (AIC), the sum of squared error (SSE), and R^2 as well as visual interpretation of each model. An independent dataset with 1,303 trees was used for validation, with bias (percent error of total trees, S%), efficiency factor (EF) and mean absolute percent error (MAPE%) were the main indicators for validation of selected. The results indicated that random effects of eco-regions improved estimates; the best model of AGB = a*[(DBH/100)^2*H*WD]^b had the highest accuracy with a S% bias under 3%, MAPE% < 18% and EF of more than 0.95. By using the best model of this study, MAPE% was reduced significantly by 14 percent compared to models from IPCC (2003) and Brown (1997) and 10-18 percent to Chave (2005, 2014).

Keywords: Allometric equation, evergreen broadleaved forest, tree biomass, Viet Nam

Introduction, scope and main objectives

In the context of global climate change, forest management to mitigate climate change through CO_2 absorption by forest ecosystems deserves urgent attention from governments. To help support this need, the UN-REDD Programme has been taking action in developing countries and in Viet Nam since 2009. The IPCC, a scientific body set up under the auspices of the UN, in 1996, 2003 and 2006 also provided guidelines to measure and monitor forest carbon. However, there is a significant need globally and in Viet Nam to develop models for biomass and carbon estimations for national measuring, reporting and verifying (MRV) systems and produce accurate emission factors and reliable training datasets of biomass, carbon per hectare (ha) for activity data.

For rainforests over the tropics, authors have provided biomass models such as IPCC (2003), Brown *et al.* (1997), Chave *et al.* (2005, 2014) and Basuki *et al.* (2009). But, these developed models have no data on forest types, ecological zones and have not been evaluated for reliability in Viet Nam. Thus, analysis of the national scale dataset was required.

This study has further analysed this data to achieve the following objectives: i) Develop models to estimate Above Ground Biomass (AGB) in Evergreen Broadleaf Forest (EBLF); ii) Consider eco-regions and environment factors to set up models with adaptive parameters to increase the reliability of biomass estimates in different forest ecological conditions in Viet Nam; iii) Validate the reliability and accuracy of selected models in this study and compare with pan-tropical models to provide a proposal to apply these models in the UN-REDD Programme in Viet Nam.

Methodology

Dataset description:

The data was collected by Viet Nam Forestry University (VFU), Vietnamese Academy of Forest Sciences (VAFS), Forest Inventory and Planning Institute (FIPI) and Tay Nguyen University (TNU) with the support of UN-REDD Phase I Programme. Within each of the 1 ha sample plots, the number of trees sampled was determined based on the ratio of trees within each diameter class. In each sample plot, the fresh biomass of stems, branches and leaves from 55 trees were collected and prepared for biomass and WD calculations. To find the fresh to dry ratio of each tree to calculate the biomass, samples were taken from stem, branches and old and new leaves. For WD, samples were taken from every one-fourth or fifth of stem length. For EBLF, total of 869 tree samples were collected in five eco-regions (Central Highlands, North Central Coastal, Northeast, South Central Coastal and Southeast).

Model development:

The development of AGB (included biomass of stems, branches and leaves) models with a single or group of input variables such as DBH (diameter at breast height, cm), H (tree height in m) and WD (wood specific gravity in $g.cm^{-3}$).

The list of models to test for each group of input variables was comprised of:

- DBH:
  - AGB = a*DBH^b Eq. 1
  - AGB = a + b*DBH + c*DBH^2 Eq. 2
  - AGB = a + b*DBH + c*DBH^2 + d*DBH^3 Eq. 3
- DBH + H:
  - AGB = a*DBH2H Eq. 4
  - AGB = a*DBH2H^b Eq. 5
  - AGB = a*DBH^b*H^c Eq. 6
- DBH + WD:
  - AGB = a*DBH^b*WD Eq. 7
  - AGB = a*DBH^b*WD^c Eq. 8
- DBH + H + WD:
  - AGB = a*DBH2HWD^b Eq. 9
  - AGB = a*DBH^b*H^c*WD^d Eq. 10

Group of input variables was calculated as follows:

$$\mathrm{DBH2H\ (m3)} = \left(\frac{DBH}{100}\right)^2 \times \mathrm{H} \quad \text{Eq. 11}$$

$$\mathrm{DBH2HWD\ (kg)} = \mathrm{DBH2H} \times \mathrm{WD} \times 1000 \quad \text{Eq. 12}$$

The modelling was performed by applying non-linear mixed effect models (nlme) in R software with power models on residuals.

Random effect tested:

The effect of the following variables was tested:

Ecological zone: Vietnamese eco-region (Ministry of Agriculture and Rural Development (MARD)

Environmental variable: i) Basal area classes: Four classes - Poor: BA ≤ 10 m^2/ha, Medium: 10 < BA ≤ 20 m^2/ha, Rich: 20 < BA ≤ 30 m^2/ha and Very Rich: BA > 30 m^2/ha; ii) WD classes: Four classes - WD1 = 0-0.4, WD2 = 0.4-0.6, WD3 = 0.6-0.8 and WD4 >0.8, in g/cm^3; iii) Soil type (from an old soil map of Indochina): Three types of soil - Cristalline shists, Igneous rocks and Sedimentary rocks; iv) Soil type from HWSD (Harmonised World Soil Database): Two types - Acrisols and Ferralsols; v) The length of dry season (number of months with less than 60 mm rain): One, two, three or five months; vi) Rain classes: Five classes - Rain1 <1,400, Rain2 = 1,400-1,600, Rain3 = 1,600-1,800, Rain4 = 1,800-2,000, and Rain5 >2,000 mm/year.

Indicators used for model selection:

For each group of input variables, one was selected and considered the best model. The criteria used to select the best models were (Picard, Saint Andre et al. (2012)):

- Visible issues in the three graphs**.** predicted values and observations against input variables, predicted values against observations, and residuals (or weighted residuals) against predicted values.
- AIC: Akaike information criterion**.** The model with the smaller AIC value is preferred:

$$\mathrm{AIC} = -2\ln(\mathrm{L}) + 2\mathrm{p} \qquad \text{Eq. 13}$$

where, L is the likelihood of the fitted model and p is the total number of parameters in the model.

- SSE: Sum of squared errors. The model with the smaller SSE value is preferred:

$$\mathrm{SSE} = \sum_{\mathrm{i}=1}^{\mathrm{n}}(\mathrm{Yi} - \mathrm{Yipre})^2 \qquad \text{Eq. 14}$$

where, Yipre: the predicted biomass, Yi: the observed biomass, n = number of observations.

- R^2: Coefficient of determination of the regression**.** Generally, the highest R^2 value with statistical significance level exhibits the optimal model. In some cases, despite the R^2 value being high, the model is not optimal. Therefore, this criterion is considered after the other ones presented above.

Model validation and comparison to pan-tropical models:

An independent dataset for validation was set up. The independent data was synthesized by combination of volume dataset (from FIPI) and WD database per species (from VAFS, VFU and TNU). Some 1,303 independent trees were collected, mainly located in two eco-regions (North Central Coast and Northeast of Viet Nam).

Based on this independent data, all the selected models of this study were validated and compared with pan-tropical models of Brown (1997), IPCC (2003) and Chave et al. (2005, 2014).

The validation indicators used Bias (S%), efficiency factor (EF) and mean absolute percent error (MAPE%):

- S%**:** Bias. It stands for the percentage of average error for a group of trees. Smaller values are preferred (Chave et al., 2014):

$$\mathrm{S\%} = 100 * \frac{\sum_1^n(\mathrm{Yipre} - \mathrm{Yi})}{\sum_1^n Yi} \qquad \text{Eq. 15}$$

- Efficiency Factor (EF)**:** This is a dimensionless statistic which relates model predictions to observed data (Loague and Green (1991) referred by Mayer et al., 1993). A good fit model has an EF close to one:

$$EF = 1 - \frac{\sum_1^n (Yi - Yipre)^2}{\sum_1^n (Yi - \overline{Y})^2} \quad \text{Eq. 16}$$

- MAPE%**:** Mean absolute percent error (or Average deviation percent), gives the average absolute error for a single tree prediction. Smaller MAE% value is preferred (Mayer et al., 1993):

$$MAPE\% = \frac{100}{n} \sum_{i=1}^{n} \frac{|Yipre - Yi|}{Yi} \quad \text{Eq. 17}$$

where, Yipre: the predicted biomass, Yi: the observed biomass, $\overline{Y}$ = the average of observations and n = number of trees for validation.

Results

i) Aboveground biomass (AGB) in EBLF

Model AGB = f(DBH) - Comparison of model forms and selection of the best equations:

By comparing the three model forms, the optimal result was generated through the model form: AGB = a*DBH^b in table 1

When testing random effects on the power model, the results also indicated that eco-regions (MARD), and four environment parameters including WD class, Soil type, Soil type HWSD, Dry season length affected AGB = a*DBH^b model. Compared to the model without random effect, the models with random effect reduced SSE and AIC and increased R^2. This shows a lack of H or WD variables, random effect models on ecological zones and environment factors important to improve the reliability of biomass estimations.

Table 1: Comparison of different AGB = f(DBH) models with and without random effect

Id	Model form	Random effect	Residual function	AIC	SSE	Adj. R^2
1	AGB = a + b*DBH + c*DBH^2 + d*DBH^3	No	1/DBH^k	10 214	149 857 129	0.848
2(*)	**AGB = a*DBH^b**	**MARD**	**1/DBH^k**	**9 957**	**144 519 193**	**0.854**

(*) Selected Model

The selected model with only DBH as input variable has a random effect of MARD eco-regions of MARD (in table 2).

Table 2: The selected model AGB = a*DBH^b with random effect of eco-region

Id	Random effect class	N trees	Equation
1	All trees	860	AGB = 0.139436*DBH^2.415395
2	Central Highlands	114	AGB = 0.198658*DBH^2.415393
3	North Central Coastal	331	AGB = 0.121155*DBH^2.415395
4	Northeast	215	AGB = 0.124830*DBH^2.415395
5	South Central Coastal	110	AGB = 0.132507*DBH^2.415395
6	Southeast	110	AGB = 0.120032*DBH^2.415395

Model AGB = f(DBH, H) - Comparison of model forms and selection of the best equations:

On comparing the three model forms, the optimal result was generated through the following model form: AGB = a*DBH2H^b with the lowest AIC and SSE values in table 3

Table 3: Comparison of different AGB = f(DBH, H) models with and without random effect

Id	Model form	Random effect	Residual function	AIC	SSE	Adj. R^2
1	AGB = a*DBH2H^b	No	1/DBH2H^k	9 972	135 723 726	0.863
2(*)	**AGB = a*DBH2H^b**	**MARD**	**1/DBH2H^k**	**9 816**	**121 080 777**	**0.878**
3	AGB = a*DBH2H^b	WD classes	1/DBH2H^k	9 416	80 205 987	0.919

(*) Selected Model

The overall best model here was again the power model with random effect of WD classes. It was not selected as WD classes are not developed for most tree species, but it emphasizes the importance of Wood density for biomass models. Therefore, the model with random effect on MARD classes was chosen (table 4).

Table 4: The selected model AGB = a*DBH2H^b with random effect on eco-region

Id	Random effect class	N trees	Equation
1	All trees	860	AGB = 277.27292*DBH2H^0.94705
2	Central Highlands	114	AGB = 363.43768*DBH2H^0.94705
3	North Central Coastal	331	AGB = 254.49543*DBH2H^0.94705
4	Northeast	215	AGB = 255.33956*DBH2H^0.94705
5	South Central Coastal	110	AGB = 277.88007*DBH2H^0.94705
6	Southeast	110	AGB = 235.21185*DBH2H^0.94705

Model AGB = f(DBH, WD) - Comparison of model forms and selection of the best equations:

Among two model forms were tested, AGB = a*DBH^b*WD had the lowest AIC and SSE and was therefore selected (table 5).

Table 5: Comparison of different AGB = f(DBH, WD) models with and without random effect

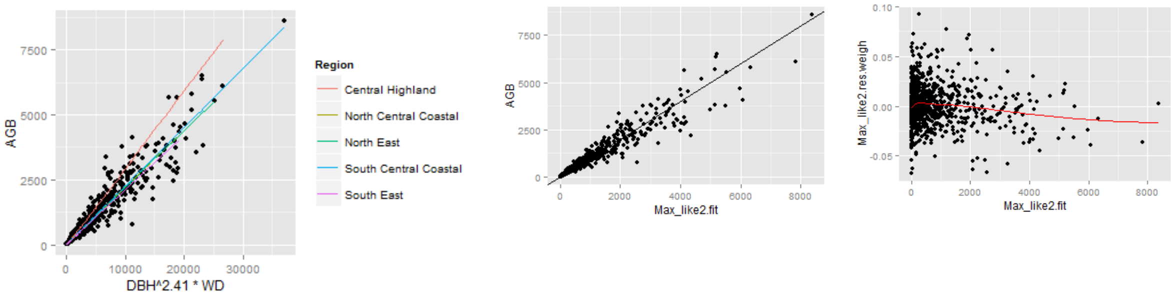
Id	Model form	Random effect	Residual function	AIC	SSE	Adj. R2
1	AGB = a*DBH^b*WD	No	1/DBH^k	9 536	87 118 189	0.912
2(*)	**AGB = a*DBH^b*WD**	**MARD**	**1/DBH^k**	**9 404**	**80 739 721**	**0.918**

(*) Selected Model

With random effect tested, the results indicated that eco-region (MARD) affected parameters of the AGB = a*DBH^b*WD model (table 6 and fig. 1). Compared to the model without random effect, models with random effect reduced SSE and AIC. When WD contributed to input variables, the eco-zone factor still affected AGB, because different conditions caused H variable changes within the DBH class.

Table 6: The selected model AGB = a*DBH^b*WD with random effect on eco-region

Id	Random effect class	N trees	Equation
1	All trees	860	AGB = 0.23342*DBH^2.40963*WD
2	Central Highlands	114	AGB = 0.23342*DBH^2.46615*WD
3	North Central Coastal	331	AGB = 0.23342*DBH^2.39720*WD
4	Northeast	215	AGB = 0.23342*DBH^2.39623*WD
5	South Central Coastal	110	AGB = 0.23342*DBH^2.40257*WD
6	Southeast	110	AGB = 0.23342*DBH^2.38600*WD



Model and Data Data = f(Fitted) Res.weighted = f(Predict)

Fig. 1: The plots of fitted values and observed, observed values against predicted, weighted residuals against predicted values of selected model AGB = 0.2334195*DBH^2.4096317*WD with random effect on region

Model AGB = f(DBH, H, WD) - Comparison of model forms and selection of the best equations:

The tree biomass content may be different, even for trees with the same volume or the same DBH and H. This is due to the tree wood density, depending greatly on tree species. While it is difficult to develop models for each species in tropical forests, the variable WD is considered a representative factor, reflecting dry biomass stored in different species. Relationships between AGB with a combination of the three variables DBH, H and WD were examined through alternative models to obtain the optimum one as shown in table 7. Compared to the two forms, the optimum result for AGB selection was generated through the models with one variable combining: AGB = a*DBH2HWD^b

Table 7: Comparison of different AGB = f(DBH, H, WD) models with and without random effect

Id	Model form	Random effect	Residual function	N trees	AIC	SSE	Adj. R^2
1(*)	**AGB = a*DBH2HWD^b**	**No**	**1/ DBH2HWD^k**	**860**	**9 200**	**65 965 670**	**0.933**
2	AGB = a*DBH^b*H^c*WD^d	No	1/ DBH^k	860	9 304	67 510 566	0.932

(*) Selected Model

The selection form AGB = a*DBH2HWD^b was tested with random effect on eco-zones and environment parameters and the results indicated no factors to affect the AGB model. This showed that three variables together were able to account for AGB, reflecting tree size and biological characteristics of species in different site conditions. Moreover, this model had the lowest AIC of all the models tested with different combinations of input variables and random effects, meaning that it is the overall best model developed.

The best model developed to predict tree AGB in evergreen broadleaved forest is:

AGB = 0.66609 *DBH2HWD^0.94304 Eq. 18

ii) Validation of selected models

The dataset of independent sampled trees in EBLFs of two eco-regions (520 trees in North Central Coast (NCC) and 783 trees in Northeast (NE)) was used to validate selected model. The results are presented in table 8.

As a result, selected model validated had a good fit with S% bias < 3% and MAPE% < 18%, EF was over 0.95.

Table 8: Validation of selected model

Equations	Eco-region	N trees	S% Bias	EF	MAPE%
AGB = 0.66609*DBH2HWD^0.94304	NCC	520	2.2	0.97	17.6
	NE	783	2.6	0.95	18.0

Discussion

Data from independent trees in EBLFs were employed to validate and compare the selected models to pan-tropical models. AGB for each independent tree was estimated following selected and pan-tropical models, plotted in Fig. 2 and then S% bias, EF and MAE% were calculated as indicators to compare the models (table 9). As a resul, by use of the best model of the current study, MAPE% was reduced by 14 percent compared to models from IPCC (2003) and Brown (1997) and 10-18 percent from Chave (2005, 2014).

These comparison results support the use of national allometric equations over generic global and regional tropical forest equations.

Table 9: Comparison of the selected models to pan-tropical models

Model	Author	Equation	N trees	S% Bias	EF	MAPE %
AGB = f(DBH)	Brown (1997)	AGB = exp(-2.134 + 2.530*ln(DBH))	1 303	24.6	0.73	49.7
	IPCC (2003)	AGB=exp(-2.289 + 2.649*ln(DBH) - 0.021*(ln(DBH))^2)	1 303	22.3	0.77	49.9
	This study (Huy I)	AGB = 0.1394363*DBH^2.4153948	1 303	-7.1	0.87	35.6
AGB = f(DBH, H, WD)	Chave II (2005)	AGB = 0.0509*WD*DBH^2*H	1 303	28.3	0.82	27.5
	Chave III (2014)	AGB = 0.0673*(WD*DBH^2*H)^0.976	1 303	30.2	0.82	35.3
	This study (Huy III)	**AGB = 0.6660939 *DBH2HWD^0.9430468**	**1 303**	**2.4**	**0.96**	**17.8**

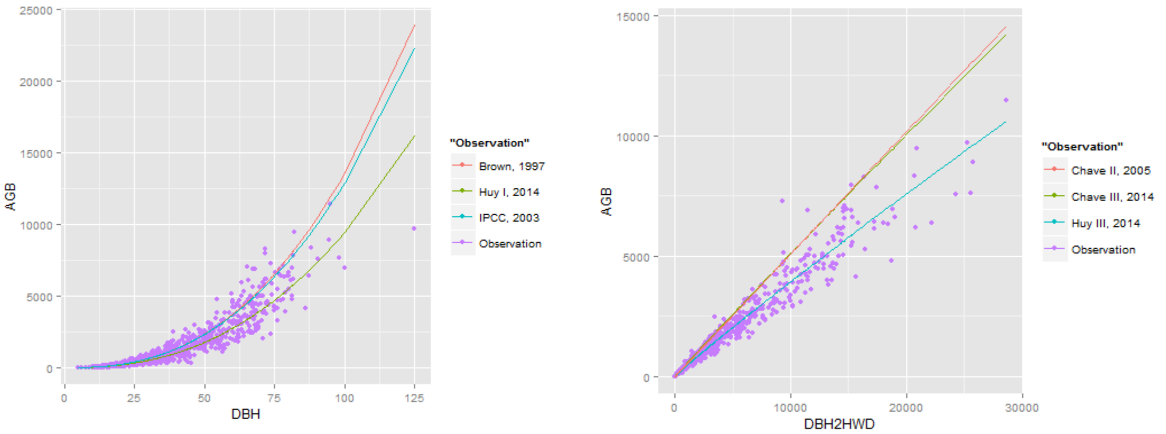


Figure 2: Graphs of the independent data and predictions by this study and different pan-tropical models.

Conclusions

From the results of the dataset analysis of evergreen broadleaved forest, national scale biomass models were developed. The main conclusions are:

- The best option is biomass models with three variables DBH, H and WD using the equation AGB = a*DBH2HWD^b. WD is more important than tree height for biomass models.
- Biomass models with random effect of eco-regions or environmental parameters improved the reliability of estimating biomass at a national scale.
- The selected model were validated had a good fit with S% bias < 3% and MAPE% < 18%, EF was over 0.95.
- AGB models at a national scale showed greater reliability than existing pan-tropical models through validations by S% bias, MAPE% and EF indicators. By use of the best model of the current study, MAPE% was reduced by 14 percent compared to models from IPCC (2003) and Brown (1997) and 10-18 percent from Chave (2005, 2014).

Acknowledgements

The author would like to acknowledge technical support during data analysis and model development with the R software provided by Gael Sola, for report edited by Gael Sola, Akiko Inoguchi, Matieu Henry (FAO), and data contributed by Viet Nam Forestry University (VFU), Vietnamese Academy of Forest Sciences (VAFS), Forest Inventory and Planning Institute (FIPI) and Tay Nguyen (TNU) with funding from the UN-REDD Programme Viet Nam.

References

1. Basuki TM, Van Lake PE, Skidmore AK, Hussin YA 2009. Allometric equations for estimating the above-ground biomass in the tropical lowland Dipterocarp forests. *For. Ecol. and Manag.* 257(2009): 1684-1694.
2. Brown S. 1997. *Estimating biomass and biomass change of tropical forests: A Primer*. FAO Forestry paper – 134. ISBN 92-5-103955-0. Available on web site: http://www.fao.org/docrep/W4095E/w4095e00.htm#Contents.
3. Chave J, Andalo C, Brown S, Cairns MA, Chambers JQ, Eamus D, Folster H, Fromard F, Higuchi N, Kira T, Lescure JP, Nelson BW, Ogawa H, Puig H, Rier B, Yamakura T. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* 145 (2005): 87-99. DOI 10.1007/s00442-005-0100-x.
4. Chave J, et al. 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. *Global change biology* (2014). Doi: 10.1111/gcb.12629.
5. IPCC 2003, 2006. *IPCC Guidelines for National Greenhouse Gas Inventories*. Prepared by the National Greenhouse Gas Inventories Programme, Eggleston H.S., Buendia L., Miwa K., Ngara T., Tanabe K., (eds). Published: IGES, Japan.
6. Mayer DG, Butler DG. 1993. Statistical validation. *Ecological Modelling*, 68(1993): 21-32.
7. Picard N, Saint Andre L, Henry M. 2012. *Manual for building tree allometric equations: from the field to the prediction*. Food and Agriculture Organization of the United Nations, Centre de Coopération Internationale en Recherche Agronomique.
8. UNFCCC 1992. *United Nations Framework Convention on Climate Change*. United Nations.
9. UN-REDD 2011. *Measurement, Reporting & Verification (MRV) Framework Document*. UN-REDD Viet Nam Programme.