ORIGINAL PAPER

Modelling height-diameter relationships in complex tropical rain forest ecosystems using deep learning algorithm

Friday Nwabueze Ogana[1] · Ilker Ercanli[2]

Abstract   Modelling tree height-diameter relationships in complex tropical rain forest ecosystems remains a challenge because of characteristics of multi-species, multi-layers, and indeterminate age composition. Effective modelling of such complex systems required innovative techniques to improve prediction of tree heights for use for aboveground biomass estimations. Therefore, in this study, deep learning algorithm (DLA) models based on artificial intelligence were trained for predicting tree heights in a tropical rain forest of Nigeria. The data consisted of 1736 individual trees representing 116 species, and measured from 52 0.25 ha sample plots. A K-means clustering was used to classify the species into three groups based on height-diameter ratios. The DLA models were trained for each species-group in which diameter at beast height, quadratic mean diameter and number of trees per ha were used as input variables. Predictions by the DLA models were compared with those developed by nonlinear least squares (NLS) and nonlinear mixed-effects (NLME) using different evaluation statistics and equivalence test. In addition, the predicted heights by the models were used to estimate aboveground biomass. The results showed that the DLA models with 100 neurons in 6 hidden layers, 100 neurons in 9 hidden layers and 100 neurons in 7 hidden layers for groups 1, 2, and 3, respectively, outperformed the NLS and NLME models. The root mean square error for the DLA models ranged from 1.939 to 3.887 m. The results also showed that using height predicted by the DLA models for aboveground biomass estimation brought about more than 30% reduction in error relative to NLS and NLME. Consequently, minimal errors were created in aboveground biomass estimation compared to those of the classical methods.

Introduction

Tree height (h) and diameter (d) are important variables that are frequently measured in forest inventories for the determination of volume, biomass, and basal area (Gomez-Garcia et al. 2014; West 2015), and used for forest stand structure analysis (Ogana and Gorgoso-Varela 2020). They provide information on the competitive status of a tree within a stand (West 2015) and their ratio is used as a stability index, i.e., tree slenderness coefficient (Sharma and Parton 2007; Zhang et al. 2020). Equally important, height and diameter measurements are used for assessing site productivity (West 2015). In fact, height-diameter allometry is regarded as the fundamental component of forest growth and yield models (Gomez-Garcia et al. 2014; Bravo et al. 2019).

The ease by which diameter and height are measured vary, with the former being easier to measure and at low cost (Ferraz-Filho et al. 2018). On the other hand, measurement

✉ Friday Nwabueze Ogana
  fn.ogana@ui.edu.ng

[1]   Department of Social and Environmental Forestry, Faculty of Renewable Natural Resources, University of Ibadan, IbadanOyo 200284, Nigeria

[2]   Department of Forest Engineering, Faculty of Forestry, Çankiri Karatekin University, Çankiri 18200, Turkey

of tree height is costly, often difficult and time- consuming (Özçelik et al. 2018; Ciceu et al. 2020; Magnussen et al. 2020), especially in complex forest ecosystems with closed-canopies (Larjavaara and Muller-Landau 2013), and as such, foresters find it more acceptable to estimate this variable (Temesgen et al. 2014). To do this, a few heights are measured and an appropriate height-diameter (h–d) function is then used to estimate other tree heights for which diameters have been measured (Kalbi et al. 2018). Modelling tree height-diameter relationships in even-aged, single-layer and monospecific or conspecific stands is straight-forward and less variable compared with complex tropical mixed forest ecosystems, characterised by multi-species, multi-layers, and indeterminate age composition (Temesgen et al. 2014).

A good example of a complex forest ecosystem is the tropical rain forest biome, regarded as one of the world's major vegetation types and the most diverse terrestrial ecosystem (Turner 2001). It serves as habitat for more fauna and flora species compared to other biomes (Turner 2001). Studies have shown that in Nigerian rainforests there are more than 4600 identified plant species (Sarumi et al. 1996) and a majority are locally endemic (Richards 1996). Turner (2001) also suggested that some tropical rain forests may have over 100 tree species with ≥ 10 cm diameter at breast height (1.3 m aboveground) on one hectare. Thus, the complex of species and structural composition within a small area makes it difficult to develop models for estimating some dendrometric variables e.g., tree height (Akindele and LeMay 2006; Bravo et al. 2019).

However, attempts have been made to develop height-diameter (h–d) models for tropical forest ecosystems using different approaches. For example, Fang and Bailey (1998) developed h–d models for all species combined in a tropical forest in Hainan, China. Feldpausch et al. (2011) developed regional h–d allometry models for tropical forest ecosystems using the ordinary least square technique. A similar approach was used by Ogana (2019) to fit h–d models in tropical mixed forests in Nigeria. However, procedures that do not take into consideration species-specific variability may not give precise predictions of height (Temesgen et al. 2014). Another alternative that has been frequently used involves the identification of major tree species, arrange the species into groups if there are many major species and use ordinary least squares (OLS) or mixed-effect modelling technique to develop models for the groups. Temesgen et al. (2014) used this methodology to develop h–d relationships for major tree species in tropical forests in Northeast China. Kearsley et al. (2017) also used a similar procedure for tropical forests in the Congo basin. This approach seems appropriate and logical, however, when aboveground biomass estimates of a tropical mixed forest is the objective, the issue of major species selection may be irrelevant. Since tropical biomass equations like those developed by Chave et al.

(2014) and Fayolle et al. (2018) require tree height as one of the input variables, it is therefore important to develop h–d models that would account for the complex nature of tropical forest ecosystems. In Nigeria, Chenge (2021) classified all the sampled species in Omo biosphere into groups and fitted both ordinary least squares (OLS) and non-linear mixed-effect (NLME) models to the group data.

A more recent approach that could be used to address the problem with modelling h–d relationships in a complex forest ecosystem is artificial neural networks (ANNs). ANNs are a subunit of artificial intelligence (AI) whose functionality mimics that of the human brain (Strobl and Forte 2007). ANNs have been consistently used in forestry with significant success for modelling tree height (Özçelik et al. 2013; Vieira et al. 2018; Bayat et al. 2020; Ercanli 2020a; Hamidi et al. 2021), tree taper (Nunes and Görgene 2016), site productivity (Aertsen et al. 2010), tree biomass and volume (Miguel et al. 2016; Özçelik et al. 2017), basal area increment (Ashraf et al. 2013) and mortality and regeneration (Hamidi et al. 2021). These researchers reported reasonable predictions of tree dendrometric variables with ANNs compared with ordinary least square and mixed-effect models. However, most of the studies have been limited to conspecific stands or stands with a few tree species. In addition to ANNs, the deep learning algorithm (DLA) is another form of AI that has been recently introduced. DLA models are multi-layered ANNs with at least three hidden layers and hundreds to thousands of neurons (Ercanli 2020a). They represent a more complex structure similar to the human brain than those of ANNs. Recent studies by Ercanli (2020a, b) showed that the DLA had better prediction of tree height in an even-aged pine stand compared to ANNs, mixed-effect and ordinary least square models.

Application of the DLA models in complex tropical forests of Africa, including Nigeria, has not apparently been documented. Yet accurate prediction of dendrometric variables such as total tree height is necessary for quantifying the aboveground biomass (AGB) of the region. When tree heights are accurately estimated for complex tropical forests, minimal errors will be introduced into the estimation of AGB. Therefore, the objectives of this study were to: (1) develop DLA models for a tropical rain forest of Nigeria; (2) compare the predictions from DLA with h–d models developed with classical methods; and, (3) evaluate the models based on aboveground biomass estimations.

Materials and methods

Data

The data used for this study were collected in Cross River State of Nigeria during a REDD+ research project funded

by the African Forest Forum (AFF) in collaboration with the Swiss Agency for Development and Cooperation (SDC). Additional inventory data from research in the Ekuri Forest Reserve in the same state were also included. The data comprise diameter and total height of 1,736 individual trees representing 116 species measured from 52 0.25 ha sample plots. The number of individual trees (n) per species ranged from 1 to 378. Of this number, only 12 species ≥ 30. Because of the multiple tree species composition, it was not possible to develop species-specific height functions. Therefore, a cluster analysis was carried out.

A K-means clustering (MacQueen 1967) was used to classify the species into groups based on height-diameter ratios; this ensures high intra-class and low inter-class similarities. The Hartigan-Wong algorithm (Hartigan and Wong 1979 cited in Kassambara 2017) was used. The algorithm minimizes the total intra-cluster variation, defined as the sum of squared Euclidean distances between the height-diameter ratio of the species and corresponding mean.

$$TWSS = \sum_{k=1}^{k} W\left(C_k\right) = \sum_{k=1}^{k} \sum_{x_i \in C_k} \left(x_i - \mu_k\right)^2 \tag{1}$$

where $TWSS$ is the total within sum of squared, W represents within, C_k is the individual cluster (group), x_i represents height-diameter ratio of a species belonging to the cluster C_k, μ_k is the mean value of the height-diameter ratio assigned to the cluster C_k. The cluster (Maechler et al. 2019) and factoextra (Kassambara and Mundt 2020) packages both implemented in R (R Core Team 2020) were used in the analysis. The 116 tree species were classified into three groups: group 1 had 68 species, group 2 and 3 had 25 and 23 species, respectively, (see group 1 had 68 species, group 2 and 3 had 25 and 23 species, respectively, (see Appendix Tables S1, S2, and S3).).

Descriptive statistics of the tree variables: diameter (d in cm), total tree height (h in m) and height-diameter ratio (h–d r); computed stand variables: quadratic mean diameter (Dg, cm), basal area per ha (G, m² ha⁻¹), basal area per ha of larger trees (BAL, m² ha⁻¹) and number of trees per ha (N, trees ha⁻¹); computed diversity indices: dominance, evenness, Simpson and Shannon indices of the data by species-group are shown in Table 1. The species-group data were randomly split into training (85%) and validation (15%) sets. Diameter histograms (pooled data) and scatter plots by species-group are presented in Fig. 1a and b, respectively.

Modelling the height-diameter (h–d) relationships

Two sets of h–d models were developed for each species-group (SG) data from tropical rain forest ecosystems of Nigeria: those based on classical methods, i.e., nonlinear least squares (NLS) and nonlinear mixed-effects (NLME),

Table 1 Descriptive statistics of the species group data from tropical rain forests of Nigeria

		d (cm)	h (m)	hdr	Dg (cm)	G (m² ha⁻¹)	BAL (m² ha⁻¹)	N (trees ha⁻¹)	Dominance	Evenness	Shannon	Simpson
Group: 1 (SG1)	Species: 68; n trees: 1,386											
	Min	10	5.7	0.2	18.8	3.7	0	44	0.045	0.224	0.136	0.059
	Max	113.4	51.5	2.1	64.9	40.9	40.9	440	0.941	0.982	3.331	0.955
	Mean	29.8	22.8	0.9	35.5	19.8	15	216.4	0.227	0.619	2.164	0.773
	SD	16.1	8.1	0.3	6.8	11.2	10	125.5	0.186	0.248	0.661	0.186
Group: 2 (SG2)	Species: 25; n trees: 139											
	Min	11.6	9.4	0.2	18.8	3.7	0	48	0.045	0.224	0.136	0.059
	Max	130	45	1.5	64.9	40.9	39	440	0.941	0.982	3.331	0.955
	Mean	54.4	27.9	0.5	37.5	16.2	7.2	165.3	0.175	0.721	2.316	0.825
	SD	23.5	8.3	0.2	7.6	10.1	7.6	115.7	0.154	0.237	0.575	0.154
Group: 3 (SG3)	Species: 23; n trees: 211											
	Min	10.1	6.4	0.2	27.2	4.9	0	48	0.045	0.224	1.271	0.473
	Max	125	43	2.2	55	40.9	39.9	440	0.527	0.982	3.331	0.955
	Mean	24.1	22.9	1.1	32.5	19.8	16.5	250.3	0.147	0.646	2.541	0.853
	SD	13.4	6.9	0.3	4.7	8.8	8.5	111.9	0.109	0.215	0.526	0.109

d Diameter, h Height, hdr Height-diameter ratio, Dg Quadratic mean diameter, G Basal area per ha, BAL Basal area in larger trees, N Number of trees per ha

Fig. 1 a Diameter histograms
(pooled data) and **b** scatterplots
by species group (SG)
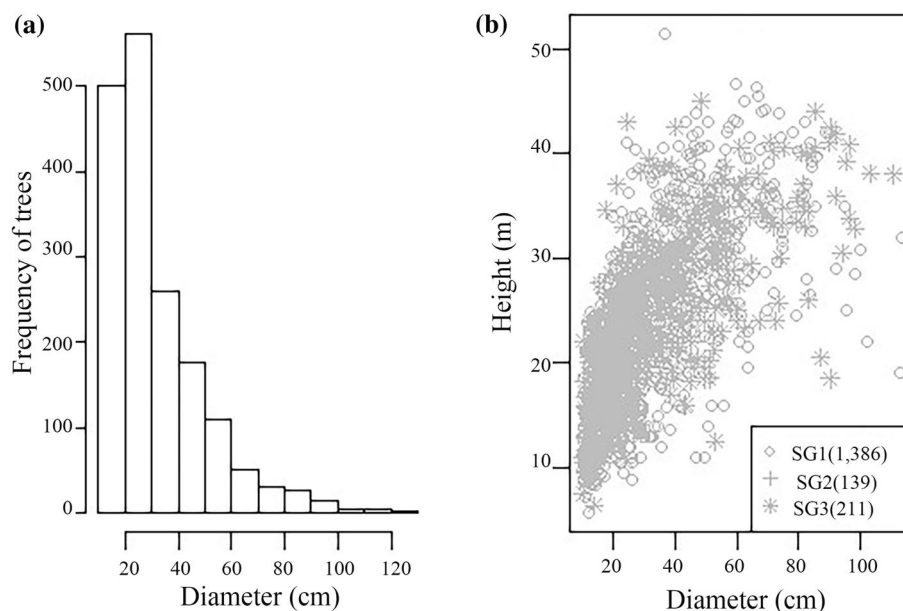
and those based on artificial intelligence (AI), i.e., the deep
learning algorithm (DLA).

Models based on classical methods: NLS and NLME

Several nonlinear single predictor height-diameter func-
tions have been used to describe tree height and diameter
relationships in both even-aged and uneven-aged stands
(Mehtätalo et al. 2015; Corral-Rivas et al. 2019; Bronisz
and Mehtätalo 2020; Ciceu et al. 2020; Ercanli 2020a;
Ogana et al. 2020; Xie et al. 2020), and in complex natu-
ral forests (Feldpausch et al. 2011; Temesgen et al. 2014;
Kearsley et al. 2017; Ogana 2019; Chenge 2021). To select
the base model for the complex tropical forests, 18 single
predictor h–d models were initially evaluated. The models
include: Curtis (1967), Meyer (1940), Chapman-Richards
(Richards 1959), Michailoff (1943), Michaelis-Menten
(Michaelis and Menten 1913), Korf (Lundqvist 1957), Näs-
lund (1937), Power (Stoffels and van Soest 1953), modified
power (Ogana and Gorgoso-Varela 2020), Prodan (Strand
1959), Gompertz (1825), Logistic (Pearl and Reed 1920),
Ratkowsky (1990), Schenute (1981), Wykoff (Wykoff et al.
1982), modified Hossfeld IV (Ogana et al. 2020), Weibull
(Yang et al. 1978), and Burkhart (Burkhart and Strub 1974).
Nonlinear least square (NLS) was used to fit the models in R
(R Core Team 2020) and were evaluated and ranked based
on five indices. Preliminary results showed that Meyer had
the minimum rank sum (see Appendix Table S4). Thus, the
model was selected and expanded.

The Meyer model (Eq. 2) was expanded with the inclusion
of stand variables and biodiversity indices. Stand variables
(*Dg*, *G*, *BAL* and *N*) and biodiversity indices (dominance,
evenness, Shannon and Simpson) in Table 1 were all evaluated

first. However, only the inclusion of the quadratic mean diam-
eter (*Dg*) and number of trees per ha (*N*) in a linear combina-
tion as replacement for the asymptotic parameter b_0 improved
the models significantly. The generalised model is expressed
as Eq. (3):

$$E(h) = 1.3 + b_0\left(1 - e^{-b_1 d}\right) \tag{2}$$

$$E(h) = 1.3 + \left(a_0 + a_1 Dg + a_2 N\right)\left(1 - e^{-b_1 d}\right) \tag{3}$$

where $E(h)$ and d represent expected total tree height (m)
and diameter at breast height (cm), respectively; *Dg* is quad-
ratic mean diameter (cm), *N* is number of trees per ha (trees
ha^{-1}), a_0, a_1, a_2, b_1 are model parameters. Equations (2) and
(3) were both fitted with NLS and NLME to the individual
species-group data. The NLS has only fixed-effects param-
eters which explain the trend in tree height common to the
overall stand (Ercanli 2020a). Contrary to the NLS, NLME
has both fixed and random effects parameters. The fixed
effects parameters play a similar role as those of ONLS;
the random effects parameters explain the variation in h–d
relationships across the plots.

The NLME model is represented in the general equation
(Pinheiro and Bates 2013) as:

$$h_{ij} = f\left(V_{ij}; \phi_i\right) + \varepsilon_{ij}; for \, which \, i = 1, ..., m \, and \, j = 1, ..., n_i \tag{4}$$

$$\phi_i = A_i \lambda + B_i b_i \tag{5}$$

where *m* represents the number of grouping factors (one
grouping factor was used in this study [plot]); n_i represents
the number of observation in the *i*th plot; h_{ij} is the height of
tree *j* on plot *i*, V_{ij} is a covariate vector; *f* represents the

nonlinear models [Eqs. (2) and (3)]; ϕ_i is the vector $r \times 1$; r is the model parameters; λ is a vector of the fixed parameters: $p \times 1$ (p the number of fixed parameters), b_i is a vector of the random parameter: $q \times 1$ (q equal number of random parameters) (Corral-Rivas et al. 2019), A_i is equal $r \times p$ and B_i is $r \times q$, respectively, and are the dimensional matrix for the fixed and random effects, for plot i (Corral-Rivas et al. 2019). The plot effects is presumed to have a common multivariate normal distribution with zero mean and variance–covariance matrix var(b_i) given as D for all values of i (Mehtätalo et al. 2015). The ε_{ij} represents random error with zero mean and constant variance var $(\varepsilon_{ij}) = \sigma^2$. A power type variance function was used to account for heteroscedasticity in the residuals: $\sigma^2 d_{ij}^{2\delta}$, where δ, is the power parameter to be estimated. The maximum likelihood through the 'nlme' function in R (R Core Team 2020) was used to estimate the parameters of the NLME models.

Deep learning algorithm (DLA)

The deep learning algorithm (DLA) is a multi-layer artificial neural networks (ANNs) with at least three hidden layers and hundreds to thousands of neurons, and gives a better representation of complex systems such as tropical forest ecosystems (Ercanli 2020a). The DLA requires sophisticated graphical processing units; thus, this study utilised the h2o.deeplearning function of the h2o package (LeDell et al. 2020) implemented in R (R Core Team 2020) to train the models. The h2o.deeplearning function has multi-layer feedback neural networks that provide well-supervised training procedures to predict output variable from input variable(s). In training the DLA models, diameter at beast height (d, cm), quadratic mean diameter (Dg, cm) and number of trees (N, trees per ha) were used as input variables, while tree height (h, m) was the output variable. The input variables were the independent variables used for the classical methods (NLS and NLME). The DLA was trained for each species-group.

Several factors influenced the convergence of DLA, e.g., number of hidden layers, number of neurons in the hidden layers, the activation function, distribution type, epochs, epsilon and rho. The adaptive learning rate algorithm called ADADELTA (Zeiler 2012 cited in Ercanli 2020b) was used to ensure fast convergence of the DLA. The ADADELTA has both momentum training and learning rate annealing. The rho parameter explains the rate of ADADELTA, while epsilon describes the strength of the learning rate during the training. Default values of 0.999 and 1×10^{-8} for rho and epsilon, respectively, were used to train the DLA models. A default value of 1000 was also used for the epochs. A similar value was used in Ercanli (2020a, 2020b). The Gaussian distribution

was selected amidst other distributions, (e.g., Bernoulli, Huber, Poisson, Multinomial, and Laplace) in the h2o.deeplearning function as the training distribution because it is a continuous distribution. The number of hidden layers initially evaluated in this study ranged from 3 to 10 and did not consider hidden layers > 10 because too complex a network makes it difficult to achieve convergence. For each hidden layer, 10 to 100 neurons with an increment of 10 per step were used. Of the three activation functions of the h2o.deeplearning function, the rectifier function was more suitable for the data set. The activation functions describe the nonlinear trends in the tropical data set (Ercanli 2020b). The best DLA models were selected for each species-group.

Model evaluation and equivalence test

The quality of model predictions was evaluated based on the comparisons of the root mean square error (RMSE), mean relative error (MRE), mean absolute percentage error (MAPE), critical error (E_{crit}) and Bayesian Information criteria (BIC). The smaller the RMSE, MRE, MAPE, E_{crit} and BIC statistics, the better models.

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)^2}{n - p}} \tag{6}$$

$$MRE = \left. \sum_{i=1}^{n} \frac{\hat{h}_i - h_i}{h_i} \middle/ n \right. \tag{7}$$

$$MAPE = \left(\left. \sum_{i=1}^{n} \frac{|h_i - \hat{h}_i|}{h_i} \middle/ n \right. \right) \times 100 \tag{8}$$

$$E_{crit} = \frac{\sqrt{\tau^2 \sum_{i=1}^{n} \left(h_i - \hat{h}_i\right)^2 \middle/ \chi_{crit}^2}}{\overline{H}_i} \tag{9}$$

$$BIC = n\ln\left(\frac{RSS}{n}\right) + p \ln n \tag{10}$$

where RSS is residual sum of squares; n is the number of observations; p the number of parameters; $\overline{h}_i$ is average tree height; h_i is observed tree height; $\hat{h}_i$ is the predicted height by the model; τ is the standard normal deviate (≈ 1.96 at probability level of $\alpha = 0.05$) and χ_{crit}^2 was obtained for $\alpha = 0.05$. In addition, relative rank (Poudel and Cao 2013) was used to determine the relative location of each model based on the evaluation statistics. It is expressed as:

$$R_i = 1 + \frac{(m-1)\left(S_i - S_{min}\right)}{S_{max} - S_{min}} \tag{11}$$

where R_i is relative rank of model i ($i = 1, 2, \ldots, m$); m is the number of models evaluated, S_i the evaluation statistic value of model i; S_{max} and S_{min} are the maximum and minimum values, respectively, of S_i. Relative rank is a real number with 1 as the best. For each model, the relative ranks were summed across the five statistics (RMSE, MRE, MAPE, E_{crit} and BIC). Thus, the relative rank sum was used to identify the best model for estimating tree height in complex tropical rain forest ecosystems.

The equivalence test of Robinson et al. (2005) was used to further assess height prediction by classical methods (NLS and NLME) and by DLA using the validation dataset (15% of the data). In this test, the size of the region of dissimilarity between the observed tree heights and predicted heights is an important factor for deciding on the acceptability of the model/method. The test begins with the null hypothesis (H_o) of significant difference between the observed and predicted values. Thus, a rejection of the H_o implies acceptance of the prediction of tree heights by the model.

The equivalence test was performed by regressing the relationships between the observed (X) and predicted (Y, predictions by NLS, NLME and DLA) heights and also by regressing the regression parameters with the intercepts (b_0) and slope (b_1) for this relation (Ercanli 2020b). Confidence intervals (CIs) for b_0 and b_1 were calculated using a two one-sided test (TOST) (Robinson et al. 2005). TOST tests the equality of slopes (b_1) to $1 \pm 10\%$ and the equality of intercepts (b_0) to $\bar{y} \pm 10\%$ (Ercanli 2020b). We used the nonparametric bootstrap technique described by Robinson et al. (2005) to obtain the predictions of the CIs for the parameters. The number of bootstrap replicates was set at 1000 as recommended and recently used by Ercanli (2020b). The equivalence test procedures for observed (X) and predicted (Y, predictions by NLS, NLME and DLA) heights were carried out using the "equivalence" package (Robinson 2016) implemented in R (R Core Team 2020).

Aboveground biomass estimations

A useful application of h–d models is the estimation of aboveground biomass (AGB). Different studies have shown that allometric models for estimating AGB perform better when information on tree height is incorporated (Chave et al. 2014; Popkin 2015; Kearsley et al. 2017; Fayolle et al. 2018). Thus, both observed and predicted tree heights by DLA and classical methods were used to estimate the AGB of the forests. The generalised pantropical AGB model (equation [12]) (Chave et al. 2014) was used.

$$AGB_{est} = 0.0673 \times \left(\rho d^2 h\right)^{0.976} \tag{12}$$

where AGB_{est} represents estimated aboveground biomass (kg); d is diameter at breast height (cm); h is tree height (m) and ρ is wood density (g cm^{-3}). Wood density for each species was extracted from the global wood density database (Chave et al. 2009; Zanne et al. 2009). For unidentified species, an average of 0.5 g cm^{-3} was used. A similar average was used by Ogana and Ogana (2019) in the same region. Reyes et al. (1992) also used 0.5 g cm^{-3} for wood density of tropical African species. The global wood density database and the AGB model (equation [12]) have been implemented in the BIOMASS package (Rejou-Mechain et al. 2017). They were obtained with "*wdData*" and "*computeAGB*" functions of the BIOMASS package in R. However, the AGB is in megagrams (Mg)—the conventional unit of AGB (Chave et al. 2014).

The observed AGB was calculated by substituting the density, and the measured diameters and heights into Eq. (12). The predicted AGB was obtained from the density, measured diameters and predicted heights by the classical methods (NLS and NLME) and DLA. Root mean square error (RMSE), critical error (Ecrit) and mean relative error (MRE) were used to assess the adequacy of the models for estimating AGB. A plot of relative error (i.e., predicted AGB minus observed AGB, divided by the observed AGB, in %) was also used to illustrate the bias in predicted AGB.

Results

Height-diameter (h–d) models

The estimated parameters of Eqs. (2) and (3) fitted with NLS for the species groups (i.e., SG1, SG2 and SG3) are presented in Tables 2, 3 and 4. Also in the tables are the parameter estimations and variance components of the fitted nonlinear mixed effect (NLME) models expressed as Eqs. (13) and (14), and the best of the DLA models. In SG1 data, the parameters of the models by NLS and NLME had low standard errors and were significantly different from zero ($p < 0.05$), except for Eq. 14. 14 where a_0 was not significant (Table 2). Similarly, in SG2 data, parameters a_1 and a_2 were not significant in Eq. (13) and (14) (Table 3). However, all parameters in the models were significant for the SG3 data set.

$$h = 1.3 + \left(b_0 + u_j\right)\left(1 - e^{-b_1 d}\right) \tag{13}$$

$$h = 1.3 + \left(\left[a_0 + u_j\right] + a_1 dg + \left[a_2 + u_j\right]N\right)\left(1 - e^{-b_1 d}\right) \tag{14}$$

Table 2 Species group 1: Information on parameters of models, root mean square error (RMSE), mean relative error (MRE), mean absolute percentage error (MAPE) critical error (Ecrit), Bayesian information criterion (BIC) and relative rank sum ($\sum R$)

Method	Eqn	Parameter	Estimate	SE	RMSE	MRE	MAPE	Ecrit	BIC	$\sum R$
NLS	2	b_0	36.3264	0.8839	5.4852	0.065	21.6913	0.4883	7349	
		b_1	0.0339	0.0015	(5.00)	(5.00)	(5.00)	(5.00)	(5.00)	25.00
	3	a_0	38.6096	2.269	5.2525	0.0607	20.3436	0.4676	7262	
		a_1	−0.1287	0.0436	(4.42)	(3.81)	(4.24)	(4.42)	(4.92)	21.81
		a_2	0.0161	0.0023						
		b_1	0.0322	0.0015						
NLME	13	b_0	34.8992	1.0467	4.8835	0.0539	19.1895	0.4348	7076	
		b_1	0.0342	0.0015	(3.49)	(1.92)	(3.60)	(3.50)	(4.74)	17.25
		σ_u^2	15.458	0.1147						
		σ^2	1.2113	0.0321						
		δ	0.4823							
	14	a_0	36.3931	3.4334	4.8872	0.0541	19.1152	0.4351	7092	
		a_1	−0.099	0.0759	(3.50)	(1.97)	(3.56)	(3.50)	(4.75)	17.28
		a_2	0.0175	0.0034						
		b_1	0.0337	0.0015						
		σ_u^2	21.8012	0.1363						
		σ_v^2	0.0001	0.0003						
		σ^2	1.4643	0.0353						
		σ_{uv}	−0.9998							
		δ	0.4517							
DLA	100 neurons in 6 hidden layers				3.8873	0.0506	14.5569	0.3461	3209	
					(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	5.00

SE Standard error, values in parenthesis are relative ranks

The results from the evaluation statistics (RMSE, MRE, MAPE, Ecrit and BIC) showed that the DLA models outperformed other models fitted by NLS and NLME for the three species-group (Tables 2, 3 4). The DLA models had the smallest statistics and lowest relative ranks (i.e., 1.00) across the five indices for the species groups. The optimal number of hidden layers and neurons for the DLA models were: 100 neurons in six hidden layers for SG1, 100 neurons in nine hidden layers for SG2, and 100 neurons in seven hidden layers for SG3. In these DLA models, the input variables were diameter, quadratic mean diameter and number of trees per ha. Thus, based on the relative rank sum, the order of ranking is: DLA models > NLME models > NLS models.

The graphical relationships between the observed (x-axis) and predicted (y-axis) tree heights by the best three models compared with the 1:1-line for each species-group is shown in Fig. 2. As seen in the graph, the DLA models 100 neurons in six hidden layers for SG1, 100 neurons in nine hidden layers for SG2 and 100 neurons in seven hidden layers for SG3 produced a more organised cluster of measured and predicted values along the main diagonal (i.e., 1:1-line) compared with those of NLS and NLME. Furthermore, the graph of residual against predicted tree heights by the models did not show any meaningful heteroscedasticity across the three species groups (Fig. 3).

The results from the equivalence test using the validation data showed that, for all models developed by NLS, NMLE and DLA, the null hypothesis (H_0) of dissimilarity for intercept (b_0) parameters was rejected, for which the bootstrap intercept (b_0) lies inside the equivalent regions ($\bar{y} \pm 10\%$) (Table 5). In the case of the null hypothesis for dissimilarity for slope parameters (b_1), only the DLA models 100 neurons in six hidden layers for SG1, 100 neurons in nine hidden layers for SG2 and 100 neurons in seven hidden layers for SG3 were rejected, in which the bootstrap slope (b_1) lies within the equivalent regions $1 \pm 10\%$. The predicted bootstrap (b_1) limit by the NLS and NLME models were not rejected for the three species groups. Since a rejection of the H_o implies acceptance of the prediction of tree heights, the DLA models were selected for the tropical rain forest ecosystems.

Aboveground biomass estimation

Aboveground biomass (Mg) estimations using tree height predicted by NLS, NLME and DLA models were assessed

Table 3 Species-group 2: Information on parameters of models, root mean square error (RMSE), mean relative error (MRE), mean absolute percentage error (MAPE) critical error (Ecrit), Bayesian information criterion (BIC) and relative rank sum ($\sum$R)

Method	Eqn	Parameter	Estimate	SE	*p*-value	RMSE	MRE	MAPE	Ecrit	BIC	$\sum$R
NLS	2	b_0	39.9357	2.6667	0.0000	5.2484	0.0292	16.8958	0.2642	716.0	
		b_1	0.0221	0.0031	0.0000	(5.00)	(5.00)	(5.00)	(5.00)	(4.95)	24.95
	3	a_0	48.2668	7.5297	0.0000	5.2012	0.0291	16.7201	0.2619	723.0	
		a_1	−0.1796	0.1378	0.1950	(4.94)	(4.98)	(4.94)	(4.94)	(5.00)	24.80
		a_2	−0.0034	0.0078	0.6620						
		b_1	0.021	0.0031	0.0000						
NLME	13	b_0	36.3742	1.9908	0.0000	4.3306	0.0282	14.2193	0.218	672.0	
		b_1	0.0275	0.0034	0.0000	(3.89)	(4.81)	(4.12)	(3.89)	(4.63)	21.34
		σ_u^2	10.8287	0.3082							
		σ^2	3.1555	0.1664							
		δ	0.2931								
	14	a_0	42.3565	7.0432	0.0000	4.1811	0.0276	13.6756	0.2105	673.0	
		a_1	−0.1454	0.1457	0.3220	(3.71)	(4.70)	(3.94)	(3.71)	(4.64)	20.70
		a_2	−0.0009	0.0077	0.9050						
		b_1	0.0269	0.0035	0.0000						
		σ_u^2	23.2572	0.4517							
		σ_v^2	0.00006	0.00078							
		σ^2	4.5904	0.2007							
		σ_{uv}	−0.9981								
		δ	0.3769								
DLA	100 neurons in 9 hidden layers					1.9391	0.0079	4.6993	0.0976	165.0	
						(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	5.00

SE standard error, values in parenthesis are relative ranks

by the root mean square (RMSE), the mean relative error (MRE) and critical error (Ecrit) (Table 6). The results show that using tree heights predicted by DLA into the AGB Eq. (12) yielded the smallest RMSE (0.1931 Mg), MRE (0.0353) and critical error (0.4511 Mg) values. It brought about more than 30% reduction in the indices relative to NLS and NLME. The graph of relative error (%) also show that minimal error was inserted into the estimation of AGB using predicted heights by DLA compared with those of NLS and NLME models (Fig. 4). The DLA produced a near perfect smooth spline regression with little tendency toward overestimation and underestimation of aboveground biomass, whilst those of NLS and NLME were more irregular.

Discussion

This research developed models for predicting tree heights in the complex rain forest ecosystems of Nigeria using classical methods (nonlinear least square and nonlinear mixed effect) and a robust AI technique, i.e., a deep learning algorithm (DLA) with a view to improving aboveground biomass estimations. The DLA models produced the smallest evaluation statistics and, as such, were more suitable in predicting tree heights in complex tropical rain forests. Parallel observation was reported in Ercanli (2020a) who applied the DLA technique to predict tree heights of even-aged pure Anatolian Crimean pine in Turkey. The author found the DLA model 100 neurons in 9 hidden layers to be the best for predicting tree heights compared with nonlinear regression and nonlinear mixed-effect models. Similarly, Ercanli (2020b) observed that a DLA model with 100 neurons in 8 hidden layers produced the best height predictions in even-aged pure Turkish pine. In the case of the complex tropical rain forest ecosystems, DLA with 100 neurons in six hidden layers was more accurate for predicting tree heights in SG1. Species group 1 contains more than 60 tree species. For SG2 (25 tree species) and SG3 (23 tree species), 100 neurons in nine hidden layers and 100 neurons in seven hidden layers, respectively, produced the best predictions of tree height.

The DLA models trained for the tropical rain forests resulted in more than 20% and 50% reduction in the RMSE and BIC values relative to NLS and NLME models across the species groups. As a rule of thumb, a minimum ΔBIC $\leq$ 2 is required for two models to be similar (Gorgoso-Varela et al. 2019). In addition, Temesgen et al. (2014) noted

Table 4 Species-group 3: Information on parameters of models, root mean square error (RMSE), mean relative error (MRE), mean absolute percentage error (MAPE) critical error (Ecrit), Bayesian information criterion (BIC) and relative rank sum ($\sum$R)

| Method	Eqn	Parameter	Estimate	SE	p-value	RMSE	MRE	MAPE	Ecrit	BIC	$\sum$R
NLS	2	b_0	33.3667	1.8174	0.0000	4.9308	0.0547	19.0593	0.3821	1083	
		b_1	0.0499	0.0053	0.0000	(5.00)	(5.00)	(5.00)	(5.00)	(5.00)	25.00
	3	a_0	13.517	4.8776	0.0062	4.7180	0.0498	18.6285	0.3656	1077	
		a_1	0.4674	0.1166	0.0000	(4.68)	(4.54)	(4.85)	(4.68)	(4.97)	23.72
		a_2	0.0137	0.0055	0.0137						
		b_1	0.0535	0.0056	0.0000						
NLME	13	b_0	32.1665	1.8012	0.0000	4.5584	0.0515	17.7168	0.3532	1055	
		b_1	0.0542	0.0059	0.0000	(4.44)	(4.70)	(4.52)	(4.44)	(4.86)	22.96
		σ_u^2	5.9993	0.1841							
		σ^2	4.0685	0.1516							
		δ	0.2714								
	14	a_0	13.1208	6.1663	0.0350	4.3462	0.0476	17.2072	0.3368	1048	
		a_1	0.4601	0.1509	0.0027	(4.12)	(4.33)	(4.33)	(4.12)	(4.82)	21.72
		a_2	0.0128	0.0064	0.0477						
		b_1	0.0567	0.006	0.0000						
		σ_u^2	23.5421	0.3647							
		σ_v^2	0.0002	0.0009							
		σ^2	5.1425	0.1705							
		σ_{uv}	0.9998								
		δ	0.2188								
DLA	100 neurons in 7 hidden layers					2.2883	0.0123	7.9234	0.1773	309	
						(1.00)	(1.00)	(1.00)	(1.00)	(1.00)	5.00

SE Standard error, values in parenthesis are relative ranks

that the extension of a model is only necessary if the difference in RMSE is > 5%. Beside the evaluation statistics, only in the DLA models were the null hypothesis (H_0) of dissimilarity for intercept (b_0) and slope (b_1) parameters rejected. The performance of the DLA models in predicting tree heights could be attributed to the complex network of neurons with different numbers of hidden layers. The DLA models are multi-layered ANNs with at least 3 hidden layers and hundreds to thousands of neurons (Ercanli 2020a). This is the first attempt to apply DLA techniques to model height-diameter relationships in complex tropical rain forests. Although Hamidi et al. (2021) used two ANNs, i.e., multilayer perception (MLP) and radial basis function (RBF) to model height-diameter relationship and other dendrometric variables in complex Hyrcanian forests of northern Iran, few species composition exist compared to those of tropical rain forests. Moreover, the MLP and RBF contain fewer networks than those of DLA models. Ercanli (2020a) also reported better performance with DLA models compared with ANNs in pure pine stands.

Bayat et al. (2020) used the ANNs and adaptive neuro-fuzzy inference system (ANFIS) to provide better estimation of tree heights in uneven-aged, mixed stands in Iran

compared with regression analysis. Similar observation was reported by Vieira et al. (2018) for eucalyptus species. Özçelık et al. (2013) also showed that the use of ANNs improved height prediction of Crimean juniper. The ANNs model resulted in 20% reduction in RMSE compared to 13% by NLME. In addition, they noted that using ANNs is more advantageous than NLME because no height measurements are required for its application. In contrast, prior information is needed for mixed-effect model calibrations. Saudi et al. (2016) also asserted that random parameters in NLME may not be applicable for most prediction purposes except that calibration data are readily available. Data availability remains a limiting factor in complex tropical rain forests.

One important limitation of artificial intelligence is model transferability to other users (Hamidi et al. 2021). To ensure efficient transferability, the R syntax files of the DLA models was provided for the three species group in downloadable links via google drive (SG1: https://drive.google.com/file/d/1faIwy3ndBBCm39GNpxxKG2wXY_UqiT0E/view?usp=sharing; SG2: https://drive.google.com/file/d/13p9yW36_73M6U0PY42cxqWwFKNd5MwOU/view?usp=sharing; SG3: https://drive.google.com/file/d/1-bgIOsP8o25_HL-d6m2GpxNZ5tNMKwh5/view?usp=sharing). A
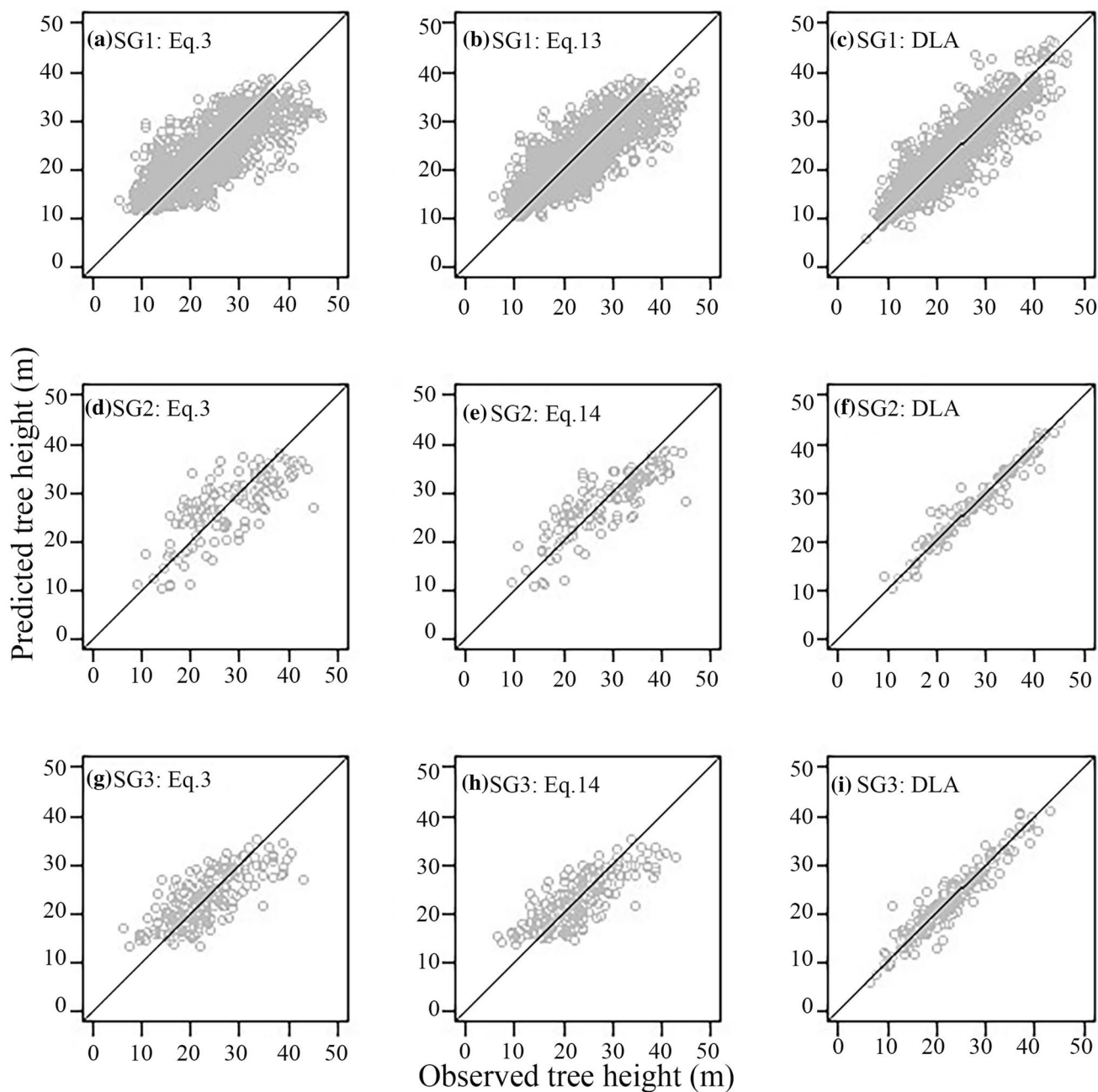
Fig. 2 Relationship between observed (x-axis) and predicted (y-axis) tree height by the three best models for each species group (SG1, SG2 and SG3)

step-by-step guide for uploading the R syntax files of DLA models in R for tree height prediction purposes can be found in the appendix of Ercanli (2020b). This ensures accessibility so that forest practitioners can use the predicted heights to estimate other dendrometric variables like tree biomass and volume.

Estimation of aboveground biomass of forest ecosystems is relevant, especially in the context of climate change. Accurate tree height predictions are required to improve

AGB estimation (Kearsley et al. 2017). Using predicted tree heights by DLA in AGB equations resulted in a 30% reduction in the root mean square error, mean relative error and critical error. This implies that the number of errors introduced into the estimation of aboveground biomass is small. In contrast, errors produced by NLS and NLME in predicting tree heights of complex tropical rain forests are brought about in AGB estimations. Because tree diameters and wood density are fixed variables, i.e., the same for DLA,
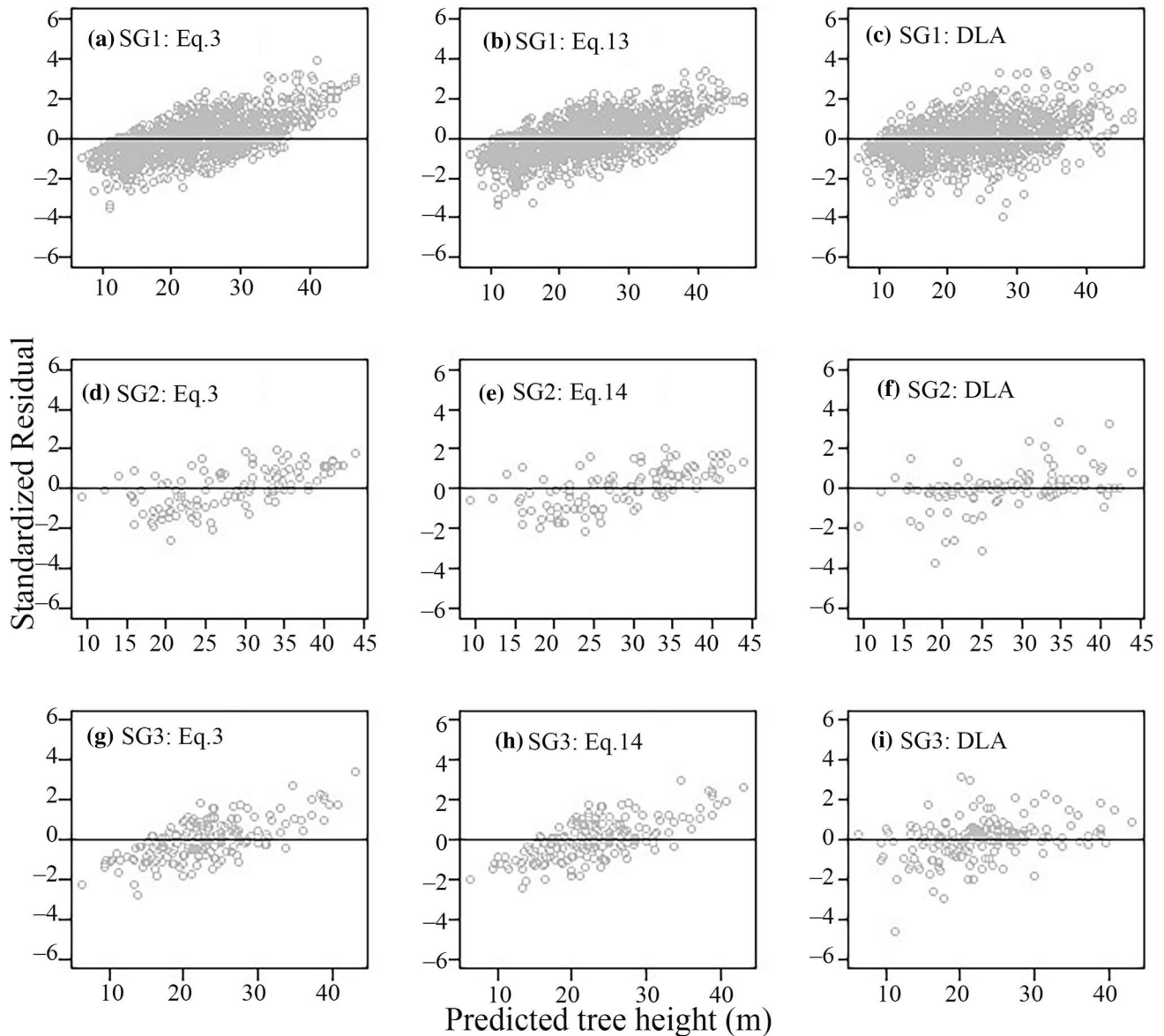
Fig. 3 Residual plots of the three best models for each species group (SG1, SG2 and SG3)

NLS and NLME, tree heights are the only source of variability. Several studies (Chave et al. 2014; Popkin 2015; Kearsley et al. 2017; Foyolle et al. 2018) have supported the use of local height-diameter model in generalised pan-tropical AGB models to minimise error in biomass estimations. Kearsley et al. (2017) quantified the size of error from using heights predicted by pan-tropical height-diameter values for aboveground estimation for the central Congo Basin. They reported a significant overestimation of tree heights which resulted in significant overestimation of AGB.

Besides the estimation of aboveground biomass, tree height predictions by DLA models could be applied to quantify the volumes of important timber species of the region. Volume equations developed for these species in the tropical rain forest by Akindele and LeMay (2006) require information on tree height as input variables. The predicted height by DLA models will improve the accuracy of estimated tree volumes, which could be scaled up to stand level.

Table 5 Equivalence tests of the models (equations [2], [3], [13], [14] and DLA) using the 15% validation data set from the species-group (SG1, SG2 and SG3)

Species-group (SG)	Eqn	b_0 limits		Bootstrap b_0 limits			b_1 limits		Bootstrap b_1 limits		
		Lower	Upper	Lower	Upper	H_0: not Equivalent	Lower	Upper	Lower	Upper	H_0: not Equivalent
SG1	2	20.7803	21.8172	15.8325	26.3875	Rejected	0.5587	0.6659	0.7500	1.2500	Not rejected
	3	20.6555	21.7476	15.8325	26.3875	Rejected	0.5636	0.6713	0.7500	1.2500	Not rejected
	13	20.7796	21.7937	15.8325	26.3875	Rejected	0.5850	0.6910	0.7500	1.2500	Not rejected
	14	20.7754	21.8061	15.8325	26.3875	Rejected	0.5836	0.6936	0.7500	1.2500	Not rejected
	DLA	20.9555	21.3236	15.8325	26.3875	Rejected	0.9690	1.0307	0.7500	1.2500	Rejected
SG2	2	22.4171	26.3201	19.8643	33.1071	Rejected	0.3618	0.8017	0.7500	1.2500	Not rejected
	3	22.2301	26.0857	19.8643	33.1071	Rejected	0.3481	0.7934	0.7500	1.2500	Not rejected
	13	22.2875	27.0757	19.8643	33.1071	Rejected	0.2999	0.8292	0.7500	1.2500	Not rejected
	14	22.2755	27.2147	19.8643	33.1071	Rejected	0.2991	0.8307	0.7500	1.2500	Not rejected
	DLA	25.4213	26.5076	19.8643	33.1071	Rejected	0.9055	1.0093	0.7500	1.2500	Rejected
SG3	2	22.9379	24.7649	17.5078	29.1797	Rejected	0.4759	0.7398	0.7500	1.2500	Not rejected
	3	22.8664	25.0090	17.5078	29.1797	Rejected	0.4747	0.7640	0.7500	1.2500	Not rejected
	13	22.5427	24.4876	17.5078	29.1797	Rejected	0.4226	0.7064	0.7500	1.2500	Not rejected
	14	22.2839	24.3139	17.5078	29.1797	Rejected	0.4086	0.7074	0.7500	1.2500	Not rejected
	DLA	23.2583	24.1429	17.5078	29.1797	Rejected	1.0090	1.1373	0.7500	1.2500	Rejected

Conclusions

The complexity of tropical rain forest ecosystems requires innovative techniques to improve the prediction of important dendrometric variables such as tree heights for

Table 6 The root mean square error (RMSE), mean relative error (MRE) and critical error (Ecrit) of models based on aboveground biomass (AGB, Mg) estimation

Method	Equation	RMSE	MRE	Ecrit
NLS	2	0.3558	0.0601	0.8311
	3	0.3578	0.0560	0.8357
NLME	13	0.3125	0.0523	0.7301
	14	0.3114	0.0518	0.7275
DLA		0.1931	0.0353	0.4511

aboveground biomass estimation. This study has shown the relevance of artificial intelligence (e.g., deep learning algorithm [DLA]) in addressing the problem of modelling tree height in complex tropical rain forest ecosystems. The DLA models outperformed other classical modelling techniques (nonlinear least square and nonlinear mixed-effects) in predicting tree heights in these ecosystems, consequently, minimizing the amount of error in aboveground biomass estimation. The input variables for the DLA models included diameter at breast height quadratic mean diameter and number of trees per ha. To facilitate the application of the DLA models by other users, a link is provided where the models can be downloaded and reused for tree height prediction.
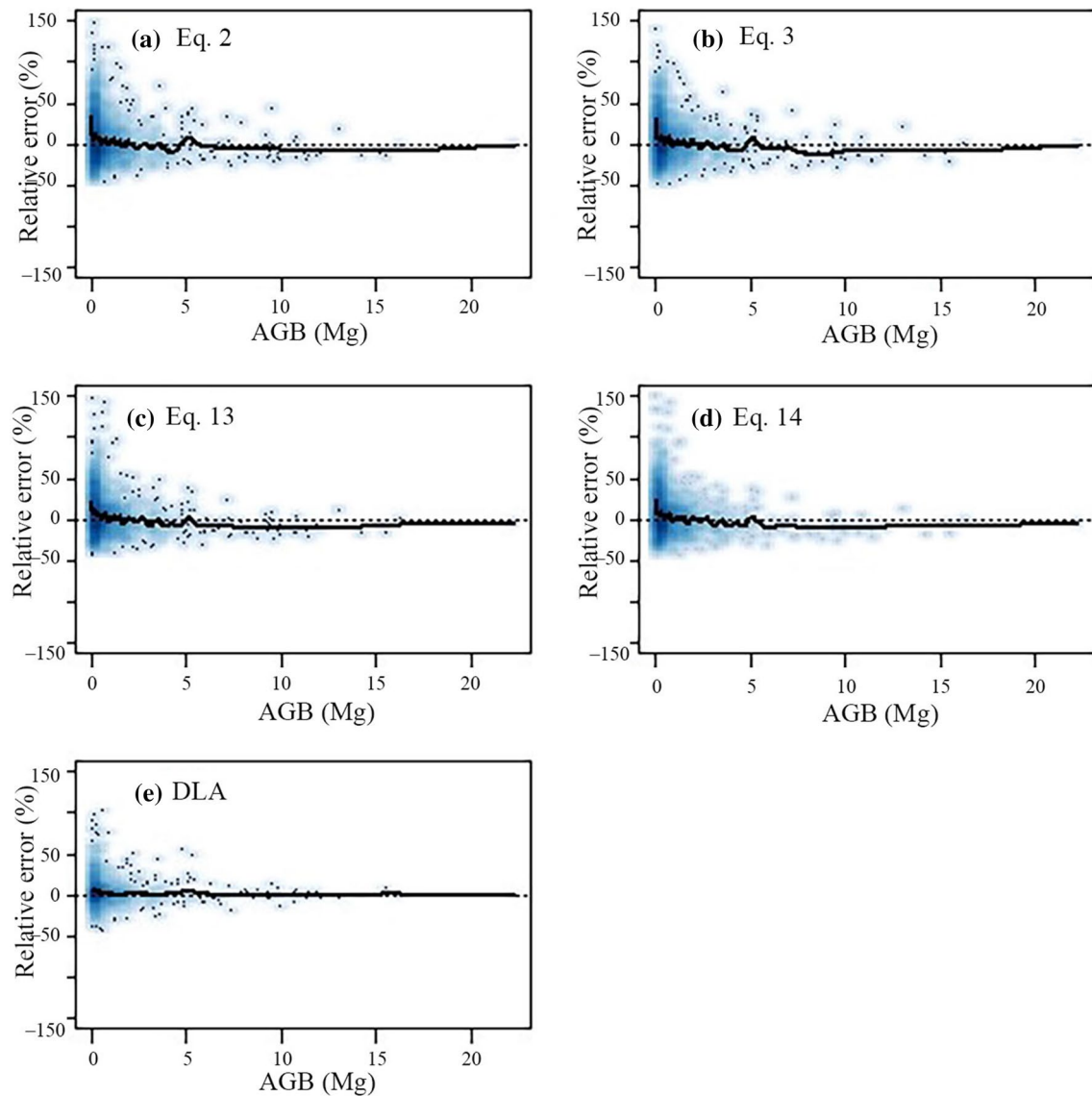
Fig. 4 Relative error (%) in the predicted AGB from the five h–d models; background and black lines represent data-point density and a spline regression data point, respectively

References

Aertsen W, Kint V, Van Orshoven J, Özkan K, Muys B (2010) Comparison and ranking of different modelling techniques for prediction of site index in Mediterranean mountain forests. Ecol Modell 221(8):1119–1130

Akindele SO, LeMay VM (2006) Development of tree volume equations for common timber species in the tropical rain forest area of Nigeria. For Ecol Manage 226:41–48

Ashraf MI, Zhao Z, Bourque CPA, MacLean DA, Meng FR (2013) Integrating biophysical controls in forest growth and yield predictions with artificial intelligence technology. Can J for Res 43(12):1162–1171

Bayat M, Bettinger P, Heidari S, Henareh-Khalyani A, Jourgholami M, Hamidi SK (2020) Estimation of tree heights in an uneven-aged, mixed forest in northern Iran using artificial intelligence and empirical models. Forests 11(3):324. https://doi.org/10.3390/f11030324

Bravo F, Fabrika M, Ammer C, Barreiro S, Bielak K, Coll L, Fonseca T, Kangur A, Löf M, Merganičová K, Pach M, Pretsch H, Stojanović D, Schiler L, Peric S, Rötzer T, del Río M, Dodan M, Bravo-Oviedo A (2019) Modelling approaches for mixed forest dynamics prognosis. Research gaps, and opportunities. For Syst 28(1):eR002. https://doi.org/10.5424/fs/2019281-14342

Bronisz K, Mehtätalo L (2020) Mixed-effects generalized height-diameter model for young sliver birch stands on post-agricultural lands. For Ecol Manag 460:117901. https://doi.org/10.1016/j.foreco.2020.117901

Burkhart HE, Strub MR (1974) A model for simulation of planted loblolly pine stands. *In*: Growth Models for Tree and Stand Simulation. Royal College of Forestry Stockholm, 379 p.

Chave J, Coomes DA, Jansen S, Lewis SL, Swenson NG, Zanne AE (2009) Towards a worldwide wood economics spectrum. Ecol Lett 12(4):351–366

Chave J, Réjou-Méchain M, Búrquez A, Chidumayo E, Colgan MS, Delitti WBC, Duque A, Eid T, Fearnside PM, Goodman RC, Henry M, Martínez-Yrízar A, Mugasha WA, Muller-Landau HC, Mencuccini M, Nelson BW, Ngomanda A, Nogueira EM, Ortiz-Malavassi E, Pélissier R, Ploton P, Ryan CM, Saldarriaga JG, Vieilledent G (2014) Improved allometric models to estimate the aboveground biomass of tropical trees. Glob Chang Biol 20:3177–3190

Chenge IB (2021) Height-diameter relationship of trees in Omo district nature forest reserve, Nigeria. Trees For People 3:100051. https://doi.org/10.1016/j.tfp.2020.100051

Ciceu A, Garcia-Duro J, Seceleanu L, Badea O (2020) A generalised nonlinear mixed-effects height-diameter model for Norway spruce in mixed-uneven aged stands. For Ecol Manage 477:118507. https://doi.org/10.1016/j.foreco.2020.118507

Corral-Rivas S, Antuna SAM, Quinonez-Barraza G (2019) A generalized nonlinear height-diameter model with mixed-effects for seven Pinus species in Durango Mexico. Revista Mexicana de Ciencias Forestales 10(53):86–117

Curtis RO (1967) Height-diameter and height-diameter-age equations for second-growth douglas-fir. For Sci 13(4):365–375

Ercanli I (2020a) Innovative deep learning artificial intelligence applications for predicting relationships between individual tree height and diameter at breast height. For Ecosyst 7:12. https://doi.org/10.1186/s40663-020-00226-3

Ercanli I (2020b) Artificial intelligence with deep learning algorithms to model relationships between total tree height and diameter at breast height. For Syst 29(2):e014. https://doi.org/10.5424/fs/2020292-16393

Fang ZX, Bailey RL (1998) Height-diameter models for tropical forests on Hainan Island in Southern China. For Ecol Manag 110:315–327

Fayolle A, Ngomanda A, Mbasi M, Barbier N, Bocko Y, Boyemba F, Couteron P, Fonton N, Kamdem N, Katembo J, Kondaoule HJ, Loumeto J, Maïdou HM, Mankou G, Mengui T, Mofack GI, Moundounga C, Moundounga Q, Nguimbous L, Nsue Nchama N, Obiang D, Ondo Meye Asue F, Picard N, Rossi V, Senguela YP, Sonké B, Viard L, Yongo OD, Zapfack L, Medjibe VP (2018) A regional allometry for the Congo basin forests based on the largest ever destructive sampling. For Ecol Manag 430:228–240. https://doi.org/10.1016/j.foreco.2018.07.030

Feldpausch TR, Banin L, Phillips OL, Baker TR, Lewis SL, Quesada CA, Affum-Baffoe K, Arets EJMM, Berry NJ, Bird M, Brondizio ES, de Camargo P, Chave J, Djagbletey G, Domingues TF, Drescher M, Fearnside PM, França MB, Fyllas NM, Lopez-Gonzalez G, Hladik A, Higuchi N, Hunter MO, Iida Y, Salim KA, Kassim AR, Keller M, Kemp J, King DA, Lovett JC, Marimon BS, Marimon-Junior BH, Lenza E, Marshall AR, Metcalfe DJ, Mitchard ETA, Moran EF, Nelson BW, Nilus R, Nogueira EM, Palace M, Patiño S, Peh KSH, Raventos MT, Reitsma JM, Saiz G, Schrodt F, Sonké B, Taedoumg HE, Tan S, White L, Wöll H, Lloyd J (2011) Height-diameter allometry of tropical forest trees. Biogeosciences 8:1081–1106. https://doi.org/10.5194/bg-8-1081-2011

Ferraz-Filho AC, Mola-Yudego B, Ribeiro A, Scolforo JRS, Loos RA, Scolforo HF (2018) Height-diameter models for *Eucalyptus* spp. Plant Brazil Cerne 24(1):9–17

Gomez-Garcia E, Dieguez-Aranda U, Castedo-Dorado F, Crecente-Campo F (2014) A comparison of model forms for the development of height-diameter relationships in even-aged stands. For Sci 60:560–568

Gompertz B (1825) On the nature of the function expressive of the law of human mortality, and on a new mode of determining the value of life contingencies. Philos Trans R Soc Lond B Biol Sci 115:513–585

Gorgoso-Varela JJ, Ogana FN, Alonso-Ponce R (2019) Evaluation of direct and indirect methods of modelling the joint distribution of tree diameter and height data with bivariate Johnson's S_{BB} function to forest stands. For Syst 28(1):e004. https://doi.org/10.5424/fs/2019281-14104

Hamidi SK, Weiskittel A, Bayat M, Fallah A (2021) Development of individual tree growth and yield model across multiple contrasting species using non-parametric and parametric methods in the Hyrcanian forests of northern Iran. Eur J for Res. https://doi.org/10.1007/s10342-020-01340-1

Hartigan JA, Wong MA (1979) A K-means clustering algorithm. Appl Stat 28:100–108

Kalbi S, Fallah A, Bettinger P, Shataee S, Yousefpour R (2018) Mixed-effects modelling for tree height prediction models of Oriental beech in the Hyrcanian forests. J for Res 29(5):1195–1204. https://doi.org/10.1007/s11676-017-0551-z

Kassambara A (2017) Practical guide to cluster analysis in R. STHDA (http://www.sthda.com), 1st edn. 187 p (accessed on 7 March 2019)

Kassambara A, Mundt F (2020) Factoextra: extract and visualize the results of multivariate data analyses. R package version 1.0.7. https://CRAN.R-project.org/package=factoextra. (accessed on 13 August 2020)

Kearsley E, Mooen PCJ, Hufkens K, Doetterl S, Lisingo J, Bosela FB, Boeckx P, Beeckman H, Verbeeck H (2017) Model performance of tree height-diameter relationship in the central Congo Basin. Ann for Sci 74:7. https://doi.org/10.1007/s13595-016-0611-0

Larjavaara M, Muller-Landau HC (2013) Measuring tree height: a quantitative comparison to two common filed methods in moist tropical forest. Methods Ecol Evol 4:793–801

LeDell E, Gill N, Aiello S, Fu A, Candel A, Click C, Kraljevic T, Nykodym T, Aboyoun P, Kurka M, Malohlava M (2020) h2o: R interface for the 'H2O' Scalable Machine Learning Platform. R package version 3.30.0.1. https://CRAN.R-project.org/package=h2o. (Accessed on 2 September 2020)

Lundqvist B (1957) On the height growth in cultivated stands of pine and spruce in Northern Sweden. Medd, Frstatens skogforsk, pp 133

MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Le Cam LM, Neyman JB (Eds) Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability. University of California Press, CA, pp. 281–297

Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K (2019) Cluster: cluster analysis basics and extensions. R package version 2.1.0. https://CRAN.R-project.org/package=cluster. (accessed on 13 August 2020)

Magnussen S, Kleinn C, Fehrmann L (2020) Wood volume errors from measured and predicted heights. Eur J for Res 139:169–178

Mehtätalo L, de-Miguel S, Gregoire TG (2015) Modelling height-diameter curves for prediction. Can J for Res 45:826–837

Meyer HA (1940) A mathematical expression for height curves. J for 38:415–420

Michaelis M, Menten ML (1913) Die kinetik der invertinwinkung. [The kinetics of invertase action.]. Biochemische Zeitung 49:333–369

Michailoff I (1943) Zahlenmassiges verfahren fur die ausfuhrung der bestandeshohenkurven forstw. Forstwissenschaftliches Centralblatt Und Tharandter Forstliches Jahrbuch 6:273–279 (**In German**)

Miguel EP, Mota FCM, Téo SJ, Nascimento RGM, Leal FA, Pereira RS, Rezende AV (2016) Artificial intelligence tools in predicting the volume of trees within a forest stand. Afr J Agric Res 11:1914–1923

Näslund M (1937) Skogsförsöksanstaltens gallringsförsök I tallskog (Forest research institute's thinning experiments in Scots pine forests). Meddelanden frstatens skogsförsöksanstalt Häfte 29. (In Swedish).

Nunes MH, Görgens EB (2016) Artificial intelligence procedures for tree taper estimation within a complex vegetation mosaic in Brazil. PLoS One 11:e0154738

Ogana FN (2019) Tree height prediction models for two forest reserves in Nigeria using mixed-effects approach. Trop Plant Res 6(1):119–128

Ogana FN, Gorgoso-Varela JJ (2020) A nonlinear mixed-effects tree height prediction model: application to Pinus pinaster Ait. in Northwest Spain. Trees For People 1:100003. https://doi.org/10.1016/j.tfp.2020.100003

Ogana TE, Ogana FN (2019) Quantification of the effect of agriculture on forest carbon stock: case study of a Nigerian forest reserve. Trop Plant Res 6(1):106–114

Ogana FN, Corral-Rivas S, Gorgoso-Varela JJ (2020) Nonlinear mixed-effect height-diameter model for Pinus pinaster Ait. and Pinus radiata D. Don. Cerne 26(1):150–161

Özçelık R, Diamantopoulou MJ, Crecente-Campo F, Eler F (2013) Estimating Crimean juniper tree height using nonlinear and artificial neural network models. For Ecol Manag 306:52–60

Özçelık R, Diamantopoulou MJ, Eker M, Gürlevık N (2017) Artificial neural network models: an alternative approach for reliable aboveground pine tree biomass prediction. For Sci 63:291–302

Özçelık R, Cao QV, Trincado G, Nilsum G (2018) Predicting tree height from tree diameter and dominant height using mixed-effect and quantile regression models for two species in Turkey. For Ecol Manage 419(420):240–248

Pearl R, Reed LJ (1920) On the rate of growth of the population of the United States since 1970 and its mathematical representation. Proc Natl Acad Sci U S A 6:275–288

Pinheiro J, Bates D (2013) Mixed-effects models in S and S-PLUS. Springer, New York, USA, p 537

Popkin G (2015) Weighing the world's trees. Nature 523:20–22

Poudel KP, Cao QV (2013) Evaluation of methods to predict Weibull parameters for characterising diameter distributions. For Sci 59(2):243–252

R Core Team (2020) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/ (accessed on 13 August 2020)

Ratkowsky DA (1990) Handbook of nonlinear regression. Marcel Dekker Inc, New York, p 19

Rejou-Mechain M, Tanguy A, Piponiot C, Chave J, Herault B (2017) BIOMASS: an R package for estimating above-ground biomass and its uncertainty in tropical forests. Methods Ecol Evol 8(9):1163–1167

Reyes G, Brown S, Chapman J, Lugo AE (1992) Wood densities of tropical tree species. Southern Forest Experiment Station, New Orleans, Louisiana

Richards FJ (1959) A flexible growth function for empirical use. J Exp Biol 10:290–300

Richards PW (1996) The tropical rain forest, 2nd edn. Cambridge University Press, Cambridge, p 599

Robinson A (2016) equivalence: provides tests and graphics for assessing tests of equivalence. R package version 0.7.2. https://CRAN.R-project.org/package=equivalence.

Robinson AP, Duursma RA, Marshall JD (2005) A regression-based equivalence test for model validation: shifting the burden of proof. Tree Physiol 25:903–913

Sarumi MB, Ladipo DO, Denton L, Olapade EO, Badaru K, Ughasoro C (1996) NIGERIA: Country Report to the FAO International Technical Conference on Plant Genetic Resources, Leipzig, Germany, 17–23 June 1996, 108 p.

Saudi P, Lynch TB, Anup KC, Guldin JM (2016) Using quadratic mean diameter and relative spacing to enhance height-diameter and crown ratio models fitted to longitudinal data. Forestry 89:215–229

Schenute J (1981) A versatile growth model with statistically stable parameters. Can J Fish Aquat Sci 38(9):1128–1140

Sharma M, Parton J (2007) Height–diameter equations for boreal tree species in Ontario using a mixed-effects modelling approach. For Ecol Manag 249:187–198

Stoffels A, van Soest J (1953) The main problems in sample plots. Ned Boschb Tijdschr 25:190–199

Strand L (1959) The accuracy of some methods for estimating volume and increment on sample plots. Medd Norske Skogfors 15(4):284–392 **(in Norwegian)**

Strobl RO, Forte F (2007) Artificial neural network exploration of the influential factors in drainage network derivation. Hydrol Process 21:2965–2978

Temesgen H, Zhang CH, Zhao XH (2014) Modelling tree height-diameter relationships in multi-species and multi-layered forests: a large observation study from Northeast China. For Ecol Manage 316:78–89

Turner IM (2001) The Ecology of trees in the tropical rain forest. Cambridge University Press, Cambridge, p 298

Vieira GC, de Mendoça AR, da Silva GF, Zanetti SS, da Silva MM, dos Santos AR (2018) Prognoses of diameter and height of trees of eucalyptus using artificial intelligence. Sci Total Environ 619:1473–1481

West PW (2015) Tree and forest measurement, 3rd edn. Springer Cham Heidelberg, New York, p 218

Wykoff WR, Crookston NL, Stage AR (1982) User's guide to the stand prognosis model. USDA For. Serv. Gen. Tech. Rep. INT-133

Xie LF, Widagdo FRA, Dong LH, Li FR (2020) Modelling height-diameter relationships for mixed-species plantations of Fraxinus

mandshurica Rupr. and Larix olgensis A. Henry in Northeastern China. Forests 11:610. https://doi.org/10.3390/f11060610

Yang RC, Kozak A, Smith JHG (1978) The potential of Weibull-type functions as flexible growth curves. Can J for Res 8:424–431

Zanne AE, Lopez-Gonzalez G, Coomes DA, Ilic J, Jansen S, Lewis SL, Miller RB, Swenson NG, Wiemann MC, Chave J (2009) Global wood density database. Dryad. Available at: https://hdl.handle.net/10255/dryad.235 (accessed 1 September 2020)

Zeiler MD (2012) ADEDELTA: an adaptive learning rate method. ArXiv-Machine Learning. arXiv:1212.5701 [cs.LG], p 6. https://arxiv.org/abs/1212.5701

Zhang XQ, Wang HC, Chhin S, Zhang JG (2020) Effects of competition, age and climate on tree slenderness of Chinese fir plantations in southern China. For Ecol Manag 458:117815. https://doi.org/10.1016/j.foreco.2019.117815

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.