Height – Diameter allometry in South Africa’s indigenous high forests: Assessing generic models performance and function forms

Sylvanus Mensah\textsuperscript{a,b,c,*}, Otto L. Pienaar\textsuperscript{d}, Anton Kunneke\textsuperscript{b}, Ben du Toit\textsuperscript{b}, Armin Seydack\textsuperscript{d}, Enno Uhl\textsuperscript{e}, Hans Pretzsch\textsuperscript{e}, Thomas Seifert\textsuperscript{b}

\textsuperscript{a} Department of Forest and Wood Science, Stellenbosch University, Matieland, South Africa
\textsuperscript{b} Laboratory of Biomathematics and Forest Estimations, University of Abomey-Calavi, Benin
\textsuperscript{c} Regional Universities Forum for Capacity Building in Agriculture, Makerere University, Wandegeya, Kampala, Uganda
\textsuperscript{d} Scientific Services Knysna, South African National Parks, South Africa
\textsuperscript{e} Chair of Forest Growth and Yield Science, Technical University of Munich, Germany

ARTICLE INFO

Keywords:
- Allometric equation
- Estimation error
- Height growth model
- Power law

ABSTRACT

Height-diameter equations are essential to understand forest dynamics and estimate forest biomass and carbon stocks. Most existing large scale height-diameter equations in Africa are based on data from rain forests, and their application to species from southern and eastern parts of Africa can result in large estimation error. Using a dataset of 1130 trees measured for their diameter and height from four forest sites with varying environmental characteristics across South Africa, we (1) evaluated the deviations in height estimated from existing generalized height-diameter equations; (2) compared the predictive ability of eight function forms applied to develop height-diameter models; (3) tested for sites and species effects on tree height-diameter allometries; and (4) developed country scale and site-specific height-diameter models in South Africa natural forests. The existing continental height-diameter equations significantly overestimated tree height in South Africa. The deviations associated with these equations, though varied with sites, remained substantially large and increased with increasing tree diameter. The power function outperformed the other theoretical functions forms and proved to be the most suitable for height-diameter allometry at country scale. As expected, forest sites and species respectively had significant effects on height-diameter allometry, suggesting further need for site and functional groups-specific height-diameter relationships. The effect of site was shown by higher scaling allometric exponents at warmer and wetter sites. On the other hand, species potentially occupying same canopy niche seem to have similar allometric relationships. Our results reveal that tree height in South Africa is more accurately predicted using locally developed models. Site-specific and country scale allometric models were thus documented for future use.

1. Introduction

Due to the complex nature of tropical evergreen or closed-canopy forest ecosystems, height measurements are often influenced by visual obstructions and observer error (Larjavaara and Muller-landau, 2013). This often limits the accuracy of height measurements in those forests even if modern laser hypsometers or LiDAR are used. As a result, forest ecologists measure tree diameter in most forest inventories in tropical forests at national or regional scale, which they rely on as a main input variable. This requires that tree height be estimated from tree diameter, and be accounted for in growth and yield models, and analyses of stand dynamics and ecosystem functioning, including biomass and carbon estimation.

Estimating tree height from trunk diameter (Curtis, 1967; Winsor, 1932) has important implications for forest management, through understanding of tree architecture and forest stand dynamics (Aiba and Kohyama, 1996; King, 1986; Pretzsch, 2010, 2009; Sokpon and Biaou, 2002), calibration of remote sensing techniques (Colgan et al., 2013; Kunneke et al., 2014) and estimation of timber volume, forest biomass and carbon (Brown, 1997; Chave et al., 2014; Colgan et al., 2013; Garber et al., 2009; Kunneke et al., 2014; Mensah et al., 2016b). Forest growth simulators use tree height-diameter models to understand competition dynamics and predict stand growth in both commercial and natural forests (Gebakken et al., 2008; Pretzsch et al., 2002; Seifert et al., 2014, Vanclay, 1994). In addition, there is mounting evidence...
that accounting for tree height in allometric biomass models leads to significantly reduced biomass estimation error (Chave et al., 2014; Mensah et al., 2017, 2016c; Rutishauser et al., 2013). Similarly, the impact of height on volume estimation (Garber et al., 2009) showed the need to develop accurate height-diameter models. While this is critical for yield modelling and management decision-making, it also has implications for tree volume and carbon stock accounting. For instance, biased estimation of tree height may result in large uncertainties of carbon stock estimates, and may also jeopardize outputs from ecological and physiological processes based forest growth models. Furthermore, height – diameter equations are also relevant for mechanical stability and wood quality. Consideration of both diameter and height offer flexible perspectives for understanding species- or functional groups-specific differences in resources allocations and growth.

The study of height – diameter relationships has attracted many research questions exploring different modelling approaches for height prediction and testing sites or environmental conditions, resource availability, taxonomic and phylogenetic effects (Banin et al., 2012; Pretzsch, 2010; Pya and Schmidt, 2016; Schmidt et al., 2011; Sumida et al., 1997; Temesgen and Gadow, 2004; Temesgen et al., 2014; Tewari and Gadow, 1999; Zucchini et al., 2001). The specific function form in the height diameter model has also been of interest for the last fifty decades (Curtis, 1967; Pretzsch et al., 2013). Substantially, the relationship between height and diameter can be expressed using linear and non-linear models, which are based on mean regression, and can be regarded as the most commonly used. Simple linear models suggest that relative tree height should scale with a constant proportion of relative tree diameter, which is not realistic, because trees reach maximum/ asymptotic heights. On the other hand, non-linear functions assume that growth in tree height is a multiplicative process through exponential scaling with diameter growth. Accordingly, several theoretical distributions were proposed to improve height-diameter models. For instance, bivariate height – diameter distributions, including the $S_{B}$ distribution, and mixture of two bivariate normal distributions (Zucchini et al., 2001) were used to model height - diameter relationship (Tewari and Gadow, 1999). Similarly, the two and three-parameter exponential and Weibull distributions, as well as many others such as the power law, Gompertz and logistic distributions, and the Chapman–Richards function were tested (Curtis, 1967; Fang and Bailey, 1998; Huang et al., 1992; Sánchez et al., 2003; van Laar and Akça, 2007; Zeide, 1993). While some authors found the power law model (or its linearized form, log-log model) useful to normalize the data and suitable to use (Mensah et al., 2017; Motatleli and Kangur, 2016), fitting the three-parameter exponential and Weibull functions also reduced errors in height estimations (Kearsley et al., 2017; Ledo et al., 2016). Furthermore, spatially explicit mixed modelling approaches were also proposed, based on stand quadratic mean diameter and spatial information (plot geographical coordinates), allowing for high accuracy prediction from a minimum set of predictor variables (Schmidt et al., 2011).

Both generalized and generic models were developed to estimate height-diameter equations (Banin et al., 2012; Chave et al., 2014; Feldpausch et al., 2011; Temesgen and Gadow, 2004). Generic models are simple general equations that lack species-specific coefficients and describe the relationship between height and diameter from global data sets across continents. Some account for environmental effects, but a few also take into account stand structure effects (Feldpausch et al., 2011). These are regional, continental and pan-tropical height – diameter models applied for tree height estimation in regions where height are difficult to measure (Chave et al., 2014; Lewis et al., 2013). Generalized models, on the other hand, are extended forms that incorporate in addition to tree diameter, stand-level variables such as stem density, basal area, quadratic mean diameter (Li et al., 2015; Temesgen et al., 2014), and relative position of trees (Temesgen and Gadow, 2004), thus accounting for species competition, spatial and temporal ecological environments (Forrester et al., 2017). They may also incorporate the between-habitat type variability of height – diameter relationship. The reason for developing generalized models is to avoid having to establish individual height–diameter relationships for each stand (Temesgen and Gadow, 2004).

The generality in these generic regional, continental and pan-tropical height – diameter models is ensured by the use of data spanning a variety of species and a wide range of environmental conditions; and hence they offer flexible perspectives for distinguishing between effects of environments, sites and different groups of species at large scale (Banin et al., 2012; Feldpausch et al., 2012, 2011). In addition these generic models offer a cost effective (less ground-based measurements and inventory efforts) and relatively accurate (for large scale application) approach of estimating tree heights. However these models can be very unsatisfactory for local or fine scale application (Kearsley et al., 2017, 2013), especially for species in environments other than those where these equations have been developed. This is likely because height – diameter relationships are inherently species-specific, and determined by plant growth through effects of environment, site quality and site productivity (Skovsgaard and Vanclay, 2013). Thus, the application of these models without consideration of site conditions and/or species information might result in large systematic errors (Kearsley et al., 2017, 2013).

Pan-tropical model parameters were originally developed without data from Africa (Chave et al., 2005), and the recent expansions of the spatial coverage in Africa to improve height estimates (Feldpausch et al., 2011, 2012; Banin et al., 2012; Chave et al., 2014) were mostly limited to the tropical central (Gabon and Congo Basin), eastern and western regions. Data from southern Africa is still largely under-represented in continental and regional scale analyses of these studies. Such a gap has been somewhat filled with a recent study that developed generic tree height-diameter using additional environmental stress variables (Chave et al., 2014). However, the validity of these continental height-diameter allometric equations has rarely been tested in southern Africa, while comparatively, more research effort has gone into that same aspect in central Africa and other parts of the world (Kearsley et al., 2013, Rutishauser et al., 2013).

Although indigenous high forests cover only about 0.1% of land area of South Africa, they have a high ecological and conservation value. The structure and functioning in these forests have been extensively studied to understand their stand dynamics (Gadow et al., 2016; Mensah et al., 2016b; Seifert et al., 2014; Seydack et al., 2012, 2011), yet there is still a research gap on height – diameter allometries in South African high forests. In particular, there is a need for validation of height – diameter models, using locally available data in southern Africa.

Therefore, the aim of this paper is to develop models for estimation of tree height in natural forest systems in South Africa. Because generic models developed for Africa in Feldpausch et al. (2012) and Banin et al. (2012) did not consider data from South Africa, we suspect that the application of these models to local situations will result in significant errors. Our study was therefore built on that hypothesis and addressed the following specific objectives: (1) evaluate the potential deviations in tree height estimation when using existing continental height – diameter models; (2) compare commonly used height-diameter function forms to determine the best model fit for South Africa based on local data; (3) use the best selected function to test for site and species effects on height-diameter relationship; and (4) use the best selected model to develop a country scale and site-specific height-diameter equations for natural forests.

2. Materials and methods

2.1. Study sites and height – diameter data

We used tree diameter and corresponding height data from four natural forest sites spanning a considerable range of environmental and
climatic conditions in South Africa (Table 1). They include data from 45 species from some of the high forest types with the largest representation in South Africa, i.e. the afrotemperate forests (Diepwalle and Groenkop) in the Southern Cape, the Southern Mistbelt forest in KwaZulu Natal and the Northern Mistbelt forest in Limpopo province (Fig. 1). Trees were selected to represent a diameter-height matrix typical for the sampled stands. The tree selection followed a random process. Tree height and diameter data for the Diepwalle research site were obtained from the French Volume Curve (FVC) research areas established in the Diepwalle forests managed by SANPARKS. These data were monitored in a context of long-term stand dynamic assessment and experimental management. Diepwalle forest is part of the Southern Cape Forests group, mostly found scattered on the coastal strip and foothill zone of the Outeniqua and Tsitsikamma mountains (between 190 and 520 m a.s.l.) (Seydack et al., 2012). The climate can be described as in a transitional stage between tropical/sub-tropical and temperate (Seifert et al., 2014). Average annual minimum and maximum temperatures in the region are 11.1 °C and 19.2 °C, respectively. Rain falls throughout the year, with annual values ranging from 700 mm on the coast to over 1200 mm at higher altitudes. Recent studies on the Diepwalle research site showed a multi-storey type of vegetation, structurally characterized by dominance of emergent species such as *Afrocarpus falcatus* and *Olinia ventosa*, with canopy and subcanopy species such as *Curtisia dentata*, *Podocarpus latifolius*, *Olea capensis* subsp. *macrocarpa*, *Gonioma kamassi* and *Elaeodendron croceum* (Gadow et al., 2016; Seifert et al., 2014).

The Groenkop forest site (33°56.5′S, 22°33′E) is located near George at about 260 m a.s.l and in the western part of the Southern Cape Forests group. Groenkop also falls within the Southern Afrotemperate forest vegetation type (Mucina and Rutherford, 2006). Annual rainfall varies between 500 mm and 1200 mm, with mean precipitation of about 850 mm (Geldenhuys, 1998) and mean annual temperature of

<table>
<thead>
<tr>
<th>Variables</th>
<th>Diepwalle afrotemperate forest</th>
<th>Groenkop afrotemperate forest</th>
<th>Southern Mistbelt forest</th>
<th>Northern Mistbelt forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coordinates (longitude, latitude)</td>
<td>33°56’S, 23°09'E</td>
<td>33°56.5’S, 22°33'E</td>
<td>29°48’38”S, 30°13’33”E</td>
<td>23°50’S, 29°59’E</td>
</tr>
<tr>
<td>Temperature (°C)</td>
<td>11.1–19.2</td>
<td>18</td>
<td>19.4–26</td>
<td>20</td>
</tr>
<tr>
<td>Precipitation (mm)</td>
<td>700–1200</td>
<td>500–1200</td>
<td>850</td>
<td>1200–1800</td>
</tr>
<tr>
<td>Altitude (m a.s.l.)</td>
<td>190–520</td>
<td>260</td>
<td>964–1554</td>
<td>1050–1800</td>
</tr>
<tr>
<td>No. of species</td>
<td>9</td>
<td>9</td>
<td>5</td>
<td>37</td>
</tr>
<tr>
<td>No. of trees</td>
<td>252</td>
<td>238</td>
<td>179</td>
<td>461</td>
</tr>
<tr>
<td>Mean diameter (cm)</td>
<td>33.0</td>
<td>36.4</td>
<td>19.7</td>
<td>24.1</td>
</tr>
<tr>
<td>Diameter range (cm)</td>
<td>9–116.4</td>
<td>5.7–165.3</td>
<td>1.0–97.0</td>
<td>0.7–94.5</td>
</tr>
<tr>
<td>Mean height (m)</td>
<td>19.0</td>
<td>18.9</td>
<td>9.8</td>
<td>13.3</td>
</tr>
<tr>
<td>Height range (m)</td>
<td>8.8–31.2</td>
<td>6.3–32.4</td>
<td>2.2–27.3</td>
<td>2.3–28.0</td>
</tr>
</tbody>
</table>

Fig. 1. Maps showing (A) the location of South Africa in Africa and (B) the location of the studied forest sites (NMF: Northern Mistbelt Forest; SMF:Southern Mistbelt Forest; Gk: Groenkop forest; and Dp: Diepwalle) in South Africa.

Table 1
Environmental characteristics of the four forest sites in South Africa and descriptive statistics of number of species, number of trees, mean, minimum and maximum values of tree diameter and height.
18°C. Its floristic composition is slightly similar to the one of the Diepwalle forest, with dominant species such as *Olea capensis* subsp. *macrorapus*, *Podocarpus latifolius* and *A. falcatus*. Studies on recruitment, growth and mortality (Geldenhuys, 1998), as well as on sustainable harvesting of non-timber forest products (Ngubeni et al. 2017, Vermeulen et al., 2012) were conducted in Groenkop Forest.

Our study site in the Southern Mistbelt Forest group was located within a plantation forest estate (Enon), managed by NCT Forestry Group. The site is situated more or less 10 km from the town of Richmond within the KwaZulu Natal Midlands (29°48′30″S, 30°13′33″E). The natural forests lie within a mosaic of exotic *Pinus*, *Eucalyptus* and *Acacia* plantations and natural Fynbos bushland vegetation, and cover an altitudinal range of 964–1554 m a.s.l. (Pienaar, 2016). Rainfall occurs throughout the year, with mean value of 850 mm and peak values during the summer months. Average temperature varies between 19.4°C and 26°C. In terms of structure, the Southern Mistbelt Forests are multi-layered, composed of trees of 10–30 m height, and occur mainly on mountain foothills, scarp slopes and gullies (Pienaar, 2016). The vegetation is dominated by species such as *Xy malos monospora* (Harv.) Baill. and *Celtis Africana* Burm.f. in the canopy layer, and *Rinorea angustifolia* (Thouars) Baill. and *Kraussia floribunda* Harv. in the subcanopy layer.

The Woodbush – De Hoek natural forest (23°50′S, 29°59′E), near Magokoskloof in the Limpopo province (South Africa) was the fourth study and covered the Northern Mistbelt forest type, classified as part of the Afromontane Archipelago in Africa (White, 1983). The annual rainfall varies between 1200 mm and 1800 mm, with the highest peak during summer. The natural vegetation occurs at altitudes up to 1800 m a.s.l, and is made up by tall and evergreen species such as *Xy malos monospora* (Harv.) Baill. ex Warb., *Syzygium gerrardii* Burtt Davy, *Cus sonia sphaerocarpa* Strey, *Combretum kraussii* Hochst., *Cassipourea malosana* (Baker) Alston, and *Trichilia dregeana* Sonn. Tree diameter and height data in the Woodbush – De Hoek natural forest were obtained from recent studies by the authors on biomass and carbon densities (Mensah et al., 2017, 2016c). Descriptive summaries of the height–diameter data (number of species covered by sites, trees measured, diameter and mean values and range) are presented in Table 1.

### 2.2. Data analyses

#### 2.2.1. Evaluating potential deviations in tree height estimation from existing continental height-diameter equations
Using measured tree diameter from South African origin, we estimated tree height from three well-known and frequently used height-diameter equations in Africa. These equations included Feldpausch et al. (2012)’s Weibull form, Banin et al. (2012)’s three-parameter exponential form and Banin et al. (2012)’s power form (see Table 2). We calculated the relative error of estimation for all the trees from the four sample sites resulting from the application of each of the three continental height-diameter equations, by comparing the estimated tree height value to its respective empirical value. The relative error for the ith tree was defined as follows:

$$\text{Relative Error}_i = \frac{\text{Estimated height}_i - \text{Observed height}_i}{\text{Observed height}_i}$$

#### Table 2

<table>
<thead>
<tr>
<th>Function</th>
<th>p</th>
<th>Model</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Weibull</td>
<td>3</td>
<td>$h_i = 50.096(1-\exp(-0.0371 \cdot dbh^{0.8291}))$</td>
<td>Feldpausch et al. (2012)</td>
</tr>
<tr>
<td>Exponential</td>
<td>45.08–42.8 $\exp(-0.025 \cdot dbh)$</td>
<td>Banin et al. (2012)</td>
<td></td>
</tr>
<tr>
<td>Power</td>
<td>3.21 $dhb^{5.99}$</td>
<td>Banin et al. (2012)</td>
<td></td>
</tr>
</tbody>
</table>

#### 2.2.2. Comparing the predictive ability of commonly used height-diameter theoretical function forms
There are several theoretical models on how to describe height-diameter relationship (Hulshof et al., 2015; Ledo et al., 2016; Li et al., 2015; Temesgen et al., 2014). Here, eight commonly used linear and non-linear function forms (Table 3) were compared for their predictive ability to identify the best model fit. For modelling and validation, the data set was randomly split into two partitions: 80% (910 individual trees) and 20% (220 trees), regardless of sites and species, set for model calibration and validation, respectively. Diameter and height values for the calibration data set ranged from 0.7 cm and 2.3 m to 165.3 cm and 32.4 m, respectively. For the validation data, dbh and height value ranges were 1–103.8 cm and 2.2–28 m, respectively.

Model parameters and regression statistics of the tested non-linear functions were estimated using the “fits” function of the *lmfit* package in the R software version 3.3.2. The performance of fitted linear and non-linear models was evaluated based on (i) numerical analysis of statistical fits such as Adjusted Coefficient of Determination (Adjusted $R^2$), Residual Standard Error (RSE), and the Mean Relative Error (MRE); and (ii) graphical analysis of the predicted vs. empirical values of tree height, which consisted of comparing the trend of predicted vs. empirical height values to the linear y:x trend. As a general rule, models with higher Adjusted $R^2$, lower RSE and lower MRE are preferred. The RSE was computed as the standard deviation of the residual errors, using the formula below:

$$\text{RSE} = \sqrt{\frac{1}{(n-p)} \sum_{i=1}^{n} (\text{Estimated height}_i - \text{Observed height}_i)^2}$$

where n is the sample size, i stands for the ith tree and p the number of parameters in the model. MRE was obtained by averaging the relative error (as calculated above) by the total number of observations:

$$\text{MRE} = \frac{1}{n} \sum_{i=1}^{n} \frac{\text{Estimated height}_i - \text{Obseved height}_i}{\text{Observed height}_i}$$

#### 2.2.3. Testing for site and species effects on height – diameter allometries
Height – diameter allometries are assumed to be species-sensitive; however, our data presents 45 species, and it would be impracticable to evaluate the exact nature of the effect of each species so a generic function was fitted. Considering the best selected model, we tested for tree size, sites and species effects on height models, using mixed-effects models (Zuur et al., 2009). We specified species as a random effect and site (Diepwalle, Groenkop, Northern Mistbelt and Southern Mistbelt) as fixed effect. The general equation of the mixed-effects model is as follows:

$$h_i = f(dbh,x_{site}) + \beta_{species} + \varepsilon$$

where $f$ is the best selected function, $x_{site}$ represents the fixed effect...
3. Results

3.1. Deviations in tree height estimation due to continental height-diameter models

Generic existing height-diameter models showed significantly higher asymptotes as compared to the trend in the observed data (Fig. 2). The three continental models ("Feldpausch et al. Weibull"; “Banin et al.3p_Exponential” and “Banin et al.Power”; see legend description in Fig. 2) estimated significantly higher tree heights compared to the empirical findings in South Africa. For the three models, we found increased relative error of height estimation with increasing tree diameter, i.e., significant overestimation of tree heights. Similar trends were also observed on each of the four study sites (Supplementary data; Fig. A). These results indicate that tree height-diameter allometry in South Africa may differ from existing tree allometries developed for Africa.

3.2. Performance of commonly used height-diameter models

The results of the statistic fits and model parameters for the tested linear and non-linear models are presented in Table 4. All the parameter estimates were significant (p < .05). Adjusted R² and RSE values ranged from 62 to 80% and 0.22 to 3.72, respectively. When comparing the eight height-diameter function forms, we found that the linear function with just two parameters showed the poorest fits as expected (lowest R², and highest values of MRE and RSE). The adjusted R² and the RSE were higher and approximately similar (76–77% for R² and 2.85–2.90 for RSE) for the Curtis, Wykoff, Richards-Chapman, Ratkowsky, Gompertz and Logistic function forms (Table 4). Compared to

\[ \text{CF} = \exp(\text{RSE}^2/2) \]
all these function forms, the power function had the best statistic fits (highest $R^2$ and lowest values of MRE and RSE), and could thus be considered as the most accurate for tree height estimation. Furthermore, the trend in the observed vs. predicted values of height showed good coincidence to the linear equation $y = x$ when the power function was used (Fig. 3).

3.3. Site and species effects on height-diameter allometric equation

The results of the linear mixed-effect models (Table 5) revealed that 90% of the variance in tree height was explained by the fixed effects of tree dbh and sites, and the random effect of species. The effect of sites was shown by significantly lower coefficients for Groenkop ($\beta = -0.03 \pm 0.014; p = .013$), Northern Mistbelt ($\beta = -0.21 \pm 0.036; p < .001$) and Southern Mistbelt ($\beta = -0.28 \pm 0.041; p < .001$), as compared to the one for Diepwalle considered as baseline. Both dbh and sites explained 83% of the variance of tree height. Conditional and marginal $R^2$ values were 90% and 83%, respectively, suggesting that little variation was caused by the random effects of species. The plot of the random intercept suggests that the species contribute differently to the general height-diameter equation (Fig. 4).

3.4. Regional and site-specific height-diameter allometric equations for tree species in natural forests of South Africa

The results of fitted power law based height-diameter equations showed highly significant model coefficients ($P < .001$; Table 6). The variance explained by the model at country scale was 80%, and ranged from 55% on Diepwalle forest site to 87% in Southern Mistbelt forest. The country scale model can be expressed as $h = 1.025 \times \exp (0.93 + 0.055 \times \ln (dbh))$. Using the fitted models, the diagnostic plots of the residuals vs. predicted values of height did not show any heteroscedastic trend; in addition, the Durbin–Watson statistic values were within the acceptable range of an upper limit of 2, suggesting no enough evidence to reject the residuals independence hypothesis. Finally, the plots of the observed versus predicted values of height showed good coincidence to the 1:1 linear trend (Figs. 5 and 6).

4. Discussion

The use of different height-diameter allometric models and parameters would likely entail significant differences in height estimates and thus affect the precision in estimation. The choice of an appropriate height-diameter model is therefore crucial for accurate analysis of forest dynamic and functions. Within this study, the deviations in tree height estimation in South Africa were analyzed when applying existing continental scale height-diameter models, and newly developed site-specific and country scale allometric models for accurate estimation of tree height in natural forests systems. It was found that (1) continental height-diameter equations significantly overestimated tree height in South Africa; (2) tree height was most accurately predicted with the power law model; (3) height-diameter allometries varied with sites and species; and (4) higher scaling height-diameter allometric exponents were found at warmer and more watered sites.

4.1. Existing generic height-diameter models induce large deviations in tree height estimates in South Africa natural forests

As pointed out existing generic height-diameter models are widely used for plant community and structure studies, application of remote sensing methods, and further for stand biomass and carbon stocks, at local, regional and global scale. For example, Lewis et al. (2013) used diameter measurement and the recommended regional height-diameter equations (Feldpausch et al., 2012) for biomass and carbon measurement in West, Central, and East Africa. Similarly, Rutishauser et al. (2013) and Kearsley et al. (2017) applied the continental model proposed by Feldpausch et al. (2012) in South East Asia and a regional model in Central Africa, respectively for estimation of biomass and comparison purposes. Clearly, the application of different published equations will affect the precision in height estimates, and this is even a more critical issue if these models are not applied to their specific region. While the use of these models at larger (regional, continental and pan-tropical) scale can be justified with regards to inventory cost-precision analysis trade-offs, their application at smaller scale, and to species and sites other than those originally used for their calibration, can entail considerable deviations between modelled outcome and reality. Therefore, country scale assessment and validation of these published larger scale equations should be done prior to application.

Our results showed that the application of three well-known generic models: Feldpausch et al.’s (2012)’s continental model for Africa; Banin et al. (2012)’s three-parameter exponential model for Africa; Banin et al. (2012)’s power model for Africa, significantly overestimated tree height for each of our four sites, and further for all sites combined (county scale). Furthermore, the relative error was found to increase with increasing tree diameter, which suggests that height asymptotic values were unrealistically overestimated, despite the fact that the range of diameters in this study was within the range of diameter data used to develop these equations. The results were consistent with other published studies. For instance, the application of Feldpausch et al.’s (2012)’s continental height-diameter models in Asia did not yield accurate estimations of tree height (Rutishauser et al., 2013). Similarly, when assessing performance of models used for height-diameter relationships, Kearsley et al. (2017) found that the use of regional models proposed in central Africa (Feldpausch et al., 2012; Banin et al., 2012) and of Chave et al. (2014)’s pan-tropical model resulted in an overestimation of tree heights in the central Congo. The present study provided evidence that existing continental height-diameter allometric models are not suitable for any of the four sites, nor for South Africa in general. The overall deviations observed may well be attributed to the difference in environmental variations (site quality and productivity, soil properties), and available species pools. The sites used in this study do have a certain degree of seasonality in the rainfall regimes, which
could reduce the maximum achievable tree heights of large diameter trees. Wind could also be a factor in limiting tree height growth by abrasion of crown tips.

Considering the fact that tree height was found to improve the performance of both generic biomass equations (Chave et al., 2014) and site-specific biomass models in South Africa (Mensah et al., 2017, 2016c), respectively and that height-diameter models are suggested to be incorporated into biomass equations (Banin et al., 2012; Feldpausch et al., 2012; Rutishauser et al., 2013), care should be given to the specific equation to use (Kearsley et al., 2017; Picard et al., 2015a). In this study, the deviations in tree height estimates, would likely propagate into estimation of standing biomass and carbon stocks. Height-diameter relationships are frequently considered as basic input for growth and yield models (Saunders and Wagner, 2008). Thus growth models and remote sensing approaches that rely on these models for calibration would result in bias outcomes if applied to our study area. Altogether, the results of this study combined with evidences of others (Kearsley et al., 2017, 2013; Rutishauser et al., 2013), reinforce the need for developing and applying locally developed height-diameter equations. Key advantages of using locally developed (site-specific) equation are that, they outperform in most cases, regional and pan-tropical equations, and account for site-related factors (e.g. competition, disturbances, soil properties, seasonality in rainfall, available species) that are known to strongly explain forest structures and tree allometry.

4.2. Power law as the most suitable theoretical function form in South Africa natural forests

Our results showed that the linear function exhibited the poorest statistical fits (lowest R square, and highest values of relative error and
R square), suggesting that stem diameter does not scale in a similar proportion with tree height, as also reported in previous studies (Motallebi and Kangur, 2016; Pretzsch et al., 2013). Compared to all the other (nonlinear) function forms, the allometric power function exhibited the best statistical fits for our data in South Africa. The suitability of the power law model to our data set is supportive of the allometric theory that assumes that the allometric exponent describes the allocation between diameter and height growth (Pommerening and Muszta, 2016; Pretzsch, 2010; Pretzsch et al., 2013).

However, if our finding accords with few studies that found the power law model suitable (Feldpausch et al., 2011; Mensah et al., 2017; Motallebi and Kangur, 2016), it also runs contrary to many previous studies that evaluated the performance of theoretical functions forms (Banin et al., 2012; Hulshof et al., 2015; Kearsley et al., 2017; Picard et al., 2015b). On the one hand, authors reported that the three-parameter exponential function, also known as the Mitscherlich model (van Laar and Akça, 2007) provided the best statistical fits for Central Africa forests (Kearsley et al., 2017, 2013; Picard et al., 2015b). On the other hand, the three-parameter Weibull model was revealed as the least biased function form (Ledo et al., 2016), and more useful to reduce uncertainties in tree height estimations especially for the smallest diameter size classes (Feldpausch et al., 2012). Furthermore, both the 3-parameter exponential and Weibull functions forms were reported to have additional biologically meaningful parameter for tree height estimation, yet, unrealistic asymptotic maximum height was obtained in the Brazilian Shield of Amazonia (Feldpausch et al., 2012). Besides, tree height was found to be best predicted in the United States when applying the Gompertz equation (Hulshof et al., 2015). This lack of convergence in the choice of the model function, indicates that some influential factors have not been incorporated in the model. This may be partly due to the environmental variation among studies and should be further investigated. A major conclusion that could thus emerge from these findings is that the suitability of the theoretical function may vary with the scale of the data and the environment.

### Table 5
Results of the mixed-effect model testing the effects of sites and species as fixed and random factor, respectively on tree height; Est.: coefficient estimates; SE: standard error.

<table>
<thead>
<tr>
<th>Model: log (h) = f (log (dbh), ωsite) + ωsp + ε</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fixed effects (Intercept)</td>
</tr>
<tr>
<td>log (dbh)</td>
</tr>
<tr>
<td>Groenkop</td>
</tr>
<tr>
<td>Northern Mistbelt</td>
</tr>
<tr>
<td>Southern Mistbelt</td>
</tr>
<tr>
<td>Random effects (Species)</td>
</tr>
<tr>
<td>Residual variance</td>
</tr>
<tr>
<td>Marginal R² (%)</td>
</tr>
<tr>
<td>Conditional R² (%)</td>
</tr>
<tr>
<td>Shapiro-Wilk normality</td>
</tr>
<tr>
<td>Number of observations</td>
</tr>
</tbody>
</table>

### Table 6
Fitted height-diameter equations (power model) with coefficient estimates and statistic fits: SE: Standard Error, R²: Adjusted R Square; RSE: Residual Standard Error and CF: Correction Factors; DW: Durbin-Watson statistic testing for autocorrelation between residuals.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Est.</th>
<th>SE</th>
<th>R²</th>
<th>RSE</th>
<th>CF</th>
<th>DW</th>
</tr>
</thead>
<tbody>
<tr>
<td>All sites</td>
<td>ln (α)</td>
<td>0.93</td>
<td>0.03</td>
<td>&lt; .001</td>
<td>80.01</td>
<td>0.224</td>
</tr>
<tr>
<td>β</td>
<td>0.55</td>
<td>0.01</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Diepwalle</td>
<td>ln (α)</td>
<td>1.88</td>
<td>0.06</td>
<td>&lt; .001</td>
<td>55.62</td>
<td>0.139</td>
</tr>
<tr>
<td>β</td>
<td>0.31</td>
<td>0.02</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Groenkop</td>
<td>ln (α)</td>
<td>1.69</td>
<td>0.06</td>
<td>&lt; .001</td>
<td>65.25</td>
<td>0.157</td>
</tr>
<tr>
<td>β</td>
<td>0.36</td>
<td>0.02</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Northern Mistbelt</td>
<td>ln (α)</td>
<td>1.01</td>
<td>0.03</td>
<td>&lt; .001</td>
<td>83.81</td>
<td>0.184</td>
</tr>
<tr>
<td>β</td>
<td>0.51</td>
<td>0.01</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Southern Mistbelt</td>
<td>ln (α)</td>
<td>0.69</td>
<td>0.04</td>
<td>&lt; .001</td>
<td>87.04</td>
<td>0.218</td>
</tr>
<tr>
<td>β</td>
<td>0.56</td>
<td>0.02</td>
<td>&lt; .001</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Fig. 4
Random intercept for each species showing the specific random predicted effect of each species on height variation; the value and the sign of the intercept for each species is indicative of the size and direction of the random effect.

### Fig. 5
Residuals and observed vs. predicted values of tree height for all sites; values were predicted from the power function.
Variations in height-diameter allometries among sites have been reported (Assogbadjo et al., 2017; Temesgen et al., 2014; Wang et al., 2006). Our results are also consistent with many other previous global and regional-scale studies (Banin et al., 2012; Chave et al., 2014; Marshall et al., 2012; Motallebi and Kangur, 2016), and therefore support the general hypothesis that tree height-diameter relationships is environmentally influenced. The variation of the height-diameter allometric equation was supported by significantly higher allometric coefficients in the Northern and Southern Mistbelt forests, as compared to the ones of Diepwalle and Groenkop forest sites. This suggests that for a same diameter, tree canopy was higher in the Northern and Southern Mistbelt forests. One potential explanation that could be offered is the marked environmental variation between these sites; Diepwalle and Groenkop forest sites are relatively close to the coast, and characterized by remarkably lower temperature (11 °C–20 °C) and annual rainfall (500 mm–1200 mm), while comparatively, species and trees in the Northern and Southern Mistbelt forests, are exposed to higher altitude, temperature and precipitation and more subtropical conditions. Altitude, water availability and increased temperature were previously reported to induce notable differences in forest structures and dbh-height relationship (Marshall et al., 2012; Wang et al., 2006). In addition, growth in tree height and diameter are modulated by trade-off in resources investment (Aiba and Kohyama, 1996), which in turn, is an adaptive strategy to the local environment. Accordingly, our results (higher allometric scaling exponents of tree diameter with height at warmer and better water supplied sites) suggest that increased temperature combined with reduced water stress likely leads to more allocated resources to height growth rather than diameter growth (Aiba and Kitayama, 1999). This apparent effect of climate on height-diameter allometry in this study, supports the hydraulic limitation theory (Ryan and Yoder, 1997). This finding also agrees with intra-specific metabolic trade-off modes with higher vertical growth priority predominating in trees of the moist, less seasonal quasi-tropical Tsitsikamma forests; with lateral growth priority prevailing in the cool, moist quasi-temperate Knysna forests (Seydack et al., 2011).

Though the results of this study suggested that species have significant effects on height-diameter allometry, we did not develop species-specific allometric equations due to the low individual sampling size by species and of the species across the forest sites. However, because site-specific models are potentially superior to generic models (as demonstrated in this study), country scale and site-dependent height-diameter models were documented for future use. Accordingly, relatively lower explained variances were found mostly in the studied Southern Cape forest sites, suggesting that additional stand related predictors might be useful in improving the models fits. Site-specific models in these areas could be extended by incorporating stand density attributes such as basal area and tree relative position, thereby accounting for inter- and intraspecific competition which strongly governs these forests (Seifert et al., 2014).

5. Conclusion

There is an increased importance of understanding tree height-diameter relationships to facilitate biomonitoring and decision making for sustaining and improving ecological functions in natural forest systems. This study presents an assessment of height-diameter relationships in relation with existing generic models and functions forms across four forest sites in South Africa. The application of regional/continental models to the data set results in considerable errors that might undermine further stand level analyses. It is therefore suggested that similar investigations be undertaken in other regions of the world lacking extensive height data to test the validity of these generic models prior to their application.

There is mounting evidence that no general function form fits best to height – diameter allometry across scales and under different environment. Here, the power model was found to be more realistic than

4.3. Site and species effects on tree height-diameter allometries

As expected, both species and sites respectively showed significant effects on tree height-diameter allometries. The effect of species on high-diameter allometry is in agreement with earlier studies (Mugasha et al., 2013; Temesgen et al., 2014). Different species have different physiological characteristics and functional traits related to wood density, specific radial variation, light requirement (shade-tolerant vs. light-demanding), that probably determine species-specific growth rate (Mensah et al., 2016a). Thus, species-dependent height-diameter allometry, as observed in this study, reflects difference in architecture and traits among coexisting species. For example, light-demanding and lower wood density species are known to be fast-growing, contrary to shade-tolerant and higher wood density species. While the effect of species was only studied across forest sites, it is important to note that species potentially occupying similar canopy niches seem to equally influence height-diameter allometry. In particular, a general increasing trend in random intercept was observed with increasing gradient in a species potential social position (understory-canopy-above canopy). This result shows the importance of revealing the complex competitive processes behind tree growth in multispecies forests as indicated by Seydack et al. (2012) and Seifert et al. (2014) in South African stands and is in line with previous studies that emphasized the need for grouping species when assessing tree allometry (Manuri et al., 2017; Mugasha et al., 2013).

Fig. 6. Residuals and observed vs. predicted values of tree height for (a) Diepwalle; (b) Groenkop; (c) Southern Mistbelt forest; and (d) Northern Mistbelt forest. Values were predicted from the power function.
of the other function forms, and the locally developed tree height-diameter allometric equation provided more statistical precision, as compared to the three tested continental height – diameter models. Height – diameter allometry also varied among study site and thus along climatic and altitudinal gradients. Both site-specific and country scale height-diameter allometric models were documented for future use in South Africa.

The present study covered a total of 45 tree species within the natural forests, and despite the considerable species random variance, and the idea that height – diameter equations are inherently species-specific, it was not possible to fit height – diameter model by tree species due to the low individual sampling size by species and of the species across the forest sites. Furthermore, height – diameter allometries for individual species or groups of species may be influenced by the locality, as shown for scots pine (Pinus sylvestris) in mixed-species stands in Estonia (Schmidt et al., 2011). Thus there is need for more extensive data on individual species and across different sites to take upfront such investigation.

Our results further suggest that the developed site-specific models might be extended to incorporate stand attributes and account for inter- and intraspecific competition. Therefore future studies should endeavor to investigate the effect of stand characteristics, tree damages, wood properties and relative position on the height-diameter relationships. Due to the need for diameter and height measurements and models that enable estimation of tree heights from measured diameter in natural forests of Southern Africa, we also suggest these aspects be extended to other forest sites in South Africa such as the forests in the Eastern Cape and in Mambalanga region as well as the coastal dune forests along the KwaZulu Natal coastline. Another meaningful research perspective, which however would require extensive data collection, would be to develop an extended set of height – diameter allometries for most important tree species in Southern Africa.

Acknowledgement

The authors are grateful to SANPARKS, NCT and Komantland Forests for granting us access to their protected areas. The first author thanks the research division of Stellenbosch University for the financial support through the “consolidor” award. We would also like to mention the EU CARE4C RISE project which enabled the fruitful exchange of ideas at an international level, and enabled our collaboration across countries. Finally, we thank the two anonymous reviewers for the valuable comments that greatly improved this article.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foreco.2017.12.030.

References

Assogbadjo, A.E., Mensah, S., Glèlè Kakaï, R., 2017. The relative importance of climatic and altitudinal gradients. Both site-specific and altitudinal height – diameter models provided more statistical precision, as compared to the three tested continental height – diameter models. Height – diameter allometry also varied among study site and thus along climatic and altitudinal gradients. Both site-specific and country scale height-diameter allometric models were documented for future use in South Africa.

The present study covered a total of 45 tree species within the natural forests, and despite the considerable species random variance, and the idea that height – diameter equations are inherently species-specific, it was not possible to fit height – diameter model by tree species due to the low individual sampling size by species and of the species across the forest sites. Furthermore, height – diameter allometries for individual species or groups of species may be influenced by the locality, as shown for scots pine (Pinus sylvestris) in mixed-species stands in Estonia (Schmidt et al., 2011). Thus there is need for more extensive data on individual species and across different sites to take upfront such investigation.

Our results further suggest that the developed site-specific models might be extended to incorporate stand attributes and account for inter- and intraspecific competition. Therefore future studies should endeavor to investigate the effect of stand characteristics, tree damages, wood properties and relative position on the height-diameter relationships. Due to the need for diameter and height measurements and models that enable estimation of tree heights from measured diameter in natural forests of Southern Africa, we also suggest these aspects be extended to other forest sites in South Africa such as the forests in the Eastern Cape and in Mambalanga region as well as the coastal dune forests along the KwaZulu Natal coastline. Another meaningful research perspective, which however would require extensive data collection, would be to develop an extended set of height – diameter allometries for most important tree species in Southern Africa.

Acknowledgement

The authors are grateful to SANPARKS, NCT and Komantland Forests for granting us access to their protected areas. The first author thanks the research division of Stellenbosch University for the financial support through the “consolidor” award. We would also like to mention the EU CARE4C RISE project which enabled the fruitful exchange of ideas at an international level, and enabled our collaboration across countries. Finally, we thank the two anonymous reviewers for the valuable comments that greatly improved this article.

Appendix A. Supplementary material

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.foreco.2017.12.030.

References

Aiba, S., Kitayama, K., 1999. Structure, composition and species diversity in an altitude-