Application of Mixed-Effects Modelling Approach in Tree Height Prediction Models

Arthur Mpazi Yambayamba

I56/70003/2013

A Research Project submitted to the School of Mathematics

Department of Statistics and Operations Research

in Partial Fulfilment of the

Requirement for the Degree of

Masters of Science in Biometry at the

University of Nairobi

July 2015

Declaration

Student Declaration: "This Research Project is my original work and to the best of my knowledge has not been presented for a degree in any other university. Furthermore, the work by other authors has been duly acknowledged."

Signature:....................................Date......................................

Name: Arthur Mpazi Yambayamba

Registration Number: I56/70003/2013

Supervisor Declaration: "This Research Project has been submitted for examination with my approval as supervisor."

Signature:....................................Date......................................

Dr. Nelson Owuor Onyango

School of Mathematics

University of Nairobi

Dedication

This research project is dedicated to my dearest wife Kanyanta, my daughter Utailo and my late mother Hyacinta Mwiche Chomba and all my relatives and friends, too numerous to mention, for the many blessings I have received from God through their lives!

Acknowledgements

I am deeply indebted to the following organizations and individuals for making my studies in general and this research project in particular a success:

1. Centre for International Cooperation(CIC) and VU University Amsterdam for sponsoring my studies;
2. The Copperbelt University for granting me leave to pursue the studies;
3. All the lecturers in the School of Mathematics for the knowledge imparted in me during my period of studies at the University of Nairobi;
4. Dr. Nelson Owuor Onyango, my supervisor, and Dr. Vincent Onguso Oeba for taking time to read through my research project and offering guidance; and
5. All my fellow students in the School of Mathematics for the academic years 2013/2014 and 2014/2015 in MSc Biometry, Social Statistics, Actuarial Science, Mathematical Statistics, Pure and Applied Mathematics for the assistance in areas where I was lacking in knowledge and also the friendship developed.

May the Almighty God bless you all and your families and all your endeavours.

Contents

List of Tables

List of Figures

Abstract

In routine forest inventories, total tree height and diameter at breast height are very important growth parameters assessed to describe and estimate the stand structure and volume, respectively, of the forest. Height-diameter models are often used to predict the height for trees where only diameter is measured for all trees in a plot and a few trees measured for total height. This is because, tree diameter can be determined easily and accurately at little cost and time, but total tree height is more difficult to measure, time consuming and more costly. Africa has lagged behind in adopting modelling techniques that can assist in estimating tree height with higher precision and accuracy than that obtained using ordinary least squares and ordinary nonlinear least squares (which are the commonly used approaches). A study was carried out to demonstrate the utility of mixed-effects modelling approach in tree height prediction models. The Chapman-Richards model was selected as base height-diameter model and was fitted to model data using Ordinary Nonlinear Least Squares method. Using the same base model and fit data, a mixed-effects model was constructed using mixed-effects modeling approach. The two models were then compared in terms of predictive accuracy on independent data set (as well as model fit data for comparison). The mixed-effects model had a better predictive accuracy on both data sets, especially the independent data. Superiority of the mixed-effects model was more clearer when the two models were compared on a plot-by-plot basis. Forest modelers and managers in Africa should consider using mixed-effects modelling approach in development and use of height-diameter models in order to estimate tree heights with higher precision and accuracy.

Chapter 1

Introduction

Tree height and diameter at breast height (DBH; that is, diameter outside of the back of tree taken at 1.30 metres above ground level) are fundamental tree characteristics used in forest measurements in order to come up with estimates for timber volume, site index and other important variables related to forest growth and yield, succession and carbon budget models (Peng *et al.*, 2001). Accurate determination of tree height and DBH are thus critical in routine forest activities, management and decision making; research and development of future forest management plans. DBH can be determined easily and accurately at little cost and time using such instruments as: diameter tapes, calipers, Biltmore stick and bark gauge. However, measurement of tree height is relatively difficult, time consuming and expensive. Furthermore, tree, stand and site conditions may also pose challenges in effective and efficient determination of tree height (Missanjo and Mwale, 2014). Therefore, in routine forest inventories and also forestry related research studies, DBH is measured for all tress sampled while tree height is measured only from a sub-sample of selected trees in different diameter classes from all DBH measurements taken within a plot/stand. Height- diameter models are then used to predict/estimate the height for trees where only DBH has been measured (Sharma and Parton, 2007).

1.1 Background

Traditionally, most of the height-diameter (H-D) models have been constructed with height as response variable and diameter as the only predictor variable. Furthermore, most of H-D models developed are fixed-effects models, which provide what is often termed a 'population average'estimate of height given DBH. The parameters in these fixed-effects models are assumed to be fixed, or that parameter estimates apply to every tree within a population. Thus, regardless of whether trees are located in same plot or stand, the parameter estimates are assumed to be correct (VanderSchaaf, 2014).

Although DBH is a suitable predictor of height, Calama and Montero (2004) report that the relationship between the diameter and height of a tree varies between stands because it depends on stand characteristics such as density and site index (Sharma and Zhang 2004). Furthermore, the H-D relationship also varies over time even within the same stand (Calama and Montero, 2004). Many studies have shown that factors such as growing space and stand conditions affect the H-D relationship; for a particular height, trees that grow in dense stands have smaller DBH than trees growing in less dense stand due to differences in competition among trees (Missanjo and Mwale, 2014). The variation in H-D relationship in this regard may result in H-D model parameters differing across plots/stands and thus specific plots/stands having what is generally termed 'random parameter'(VanderSchaaf, 2014). This may render the use of a base model (herein and throughout this report referred to as a fixed-effects H-D model with only DBH as predictor variable) less accurate in estimating tree height.

It is against this background that recent decades have seen wide spread development and application of generalized H-D models; that is, fixed-effects H-D models that incorporate stand characteristics as predictor variables in addition to tree DBH. Numerous studies have shown that inclusion of stand characteristics leads to improved accuracy in prediction of tree heights using H-D models. However, the inclusion of stand characteristic as predictor variables, in addition to tree DBH, leads to the hierarchical structure of the data (for example, trees grouped in plots and plots grouped in stands) resulting in lack of independence between measurements due to correlation in sampling units (Corral-Rivas *et al.*, 2014). When the measurements are correlated, the basic assumption about independent error term (identically distributed and normal random variable with mean of zero and equal variance) of the model does not hold (Juntunen, 2010) and this renders the application of basic modeling approaches such as ordinary least squares (OLS) and ordinary nonlinear least squares (ONLS) to estimate model parameters inappropriate. Fortunately, mixed-effects modelling approach can be applied to deal with this situation. Mixed-effects modelling estimates both fixed parameters (parameters that are common to the entire population) and random parameters (parameters that are specific to each plot/stand) at the same time within the same model and this ensures that variability between plots of the same population is modelled (Corral-Rivas *et al.*, 2014). It is also important to note that development and use of generalized H-D models might have practical implications in terms of costs and time for field sampling. This is as a result of the need to measure or determine other tree and/or stand variables in addition to tree DBH.

1.2 Problem Statement

Estimation of tree height where only DBH and/or stand variables have been measured is one of the biggest challenges faced by forest managers. Practically, it is often inevitable to have a trade-off between getting reliable estimates for tree heights and time and costs for field sampling. It is however expected that the standing volume, for example, of a given stand and/or plantation is estimated as precisely and as accurately as possible. Such precise or accurate estimations can only be achieved if DBH and tree heights were measured and/ or estimated with high precision in the first place because they (DBH and tree height) are the basic inputs in such computations (stand volume).

The challenge of estimating tree height with high precision and accuracy using H-D models is real in Africa. This is because the continent has lagged behind in adopting more efficient modelling approaches (such as mixed-effects) than the commonly used OLS and ONLS. This assertion is evidenced by insufficient information on published research works on development and use of mixed-effects H-D models in Africa; to the knowledge of the author, there are presently only two notable published research studies, (Eerikäinen, 2003; Missanjo and Mwale, 2014). Both models were developed using stand level and tree DBH as predictor variables for tree height. Insufficient information on utilization of mixed-effects modeling approach in Africa could also imply that there is continued use of H-D models developed using OLS and ONLS.

1.3 Objectives

1.3.1 Main Objective

The overall objective of this project was to demonstrate the development and utilization of a mixed-effects model in prediction of tree height (using only tree DBH as predictor variable).

1.3.2 Specific Objectives

1. To construct a mixed-effects model from a chosen base H-D model.

2. To compare the predictive accuracy of a base model developed using ONLS with that developed using mixed-effects modelling approach; on both model fitting and test data sets.

1.4 Significance of the Study

Tree height is an important variable which is used as an input in a number of aspects of forestry such as estimating stand volume, site quality and for describing stand structure (Adame *et al.*, 2008). Therefore, development and use of an H-D model with known and acceptable accuracy to estimate tree height is of paramount importance in routine forest management and research related activities.

Furthermore, forest managers are often times faced with situations where they have to use simple and accurate models that can enable them determine tree height in a stand with reliability, from measured DBH values (Calama and Montero, 2004). Thus, an H-D model with only DBH as predictor variable but has a high predictive ability and accuracy offers a reasonable opportunity to reduce time and costs involved in field sampling and also data management and analysis.

Chapter 2

Literature Review

2.1 Relationship between Tree Height and Diameter

The relationship between tree height and diameter is one of the most studied areas in forestry. The relationship between the two is a structural characteristic of a tree that describes key elements of stem form, and thus the volume of the harvestable stem. The diameter to height relationship also affects product quality, as it influences wood structure such as lignin and cellulose content, and thus important properties of the stem such as stiffness (Kroon *et al.*, 2008). The relationship also stems from the fact that tree diameter is designed to support the load that depends on tree height and thus, diameter explains a lot of variation in tree height (Zeide and VanderSchaaf, 2002). Using this relationship, a number of H-D models, both linear and nonlinear, with only DBH as predictor variable, have been developed (e.g. as cited in Huang *et al.*, 1992; Zhang, 1997; Tesemesen and Gadow, 2004; Calama and Montero, 2004; Sharma, 2009; Krisnawati *et al.*, 2010; Ahmadi *et al.*, 2013; Xu *et al.*, 2014). These models with many others, with some modified from original, have been successfully used in modelling H-D relationship for specific tree species and regions and subsequently used for example, in forest inventories and growth models for predicting missing height measurements where only DBH had been measured. There are a number of studies where fixed-effects H-D models, including only DBH as independent variable, have been fitted to data sets using mostly ONLS and very good values of goodness-of-fit statistics including bias, root mean square error, values of both coefficient of determination and adjusted coefficient of determination of more than 0.80 obtained (e.g. Zhang, 1997; Sharma, 2009; Aigbe and Oyebade, 2012; Oyebade *et al.*, 2012; Osman *et al.*, 2013; Stankova and Diéguez-Aranda, 2013; Lumbres *et al.*, 2013; Obeyed, 2014; Petráš, 2014).

2.2 Effects of Stand Variables on Height-Diameter Relationship

Eichhorn (1904), cited by Picard *et al.* (2012), observed and postulated that production of an even-aged, monospecific stand, for a given tree species in a given region and in broad range of silvicultures (as long as the canopy is closed) was entirely determined by its mean height or, by Eichhorns extended rule, on its dominant height (Picard *et al.*, 2012). However, many studies have questioned Eichhorns rule and have observed that height does not exclusively depend on wood volume, which is a function of tree diameter and height. Studies have shown that stand variables such as stand density, basal area, site index and age, affect the H-D relationship (Sharma and Parton, 2007; Krisnawati *et al.*, 2010; Missanjo and Mwale, 2014). For example, in dense stands, trees with the same diameter are taller than those in less dense stands (provided that the other conditions are the same) and thus, stand density helps explain variation in height (Zeide and VanderSchaaf, 2002; Vargas-Larreta *et al.*, 2009). In particular, Vanclay (2009) found that in even-aged stands, the mean diameter of forest trees (DBH) tends to remain proportional to the stand height (average height of the largest trees in a stand) divided by the logarithm of stand density(number of trees per hectare).

Temesgen and Gadow (2004) evaluated two sets of models (first set for estimating tree height as a function of individual tree DBH, and the second set for estimating tree height as a function of individual tree DBH and other stand-level attributes) to develop generalized H-D models for major tree species in British Columbia. The stand-level attributes used were basal area (BA), basal area in larger trees (BAL) and stems per hectare (SPH).

Their results showed that inclusion of stand-level attributes to the base H-D models increased the accuracy of prediction for all species. Misir (2010) used a similar approach to develop a generalized H-D model for *Populus tremula* stands in Turkey. Two sets of models were evaluated, with the first for estimating tree height as a function of individual tree DBH and the second set for estimating tree height as a function of individual tree DBH and some stand-level attributes. The findings were that inclusion of stand-level-attributes (comprising BA, BAL, dominant height, dominant diameter and number of trees) into the base H-D models increased the accuracy of prediction of tree height. Krisnawati *et al* (2010) developed generalized H-D models for *Acacia mangium* Willd plantations in South Sumatra by comparing model fits with only DBH as predictor variable with other models that incorporated stand variables. They observed that inclusion of one or two or all of stand variables that included stand age, site index and basal area improved the resulting fit and prediction of height, compared with that of using DBH alone. They however, observed that inclusion of number of stems per hectare was less significant. In the process of developing an H-D model for *Pinus kesiya* in Malawi, Missanjo and Mwale (2014) observed that models that incorporated basal area, stand age and site index had a great impact on H-D relationship. The generalized H-D models performed better than base models.

The foregoing case studies on effects of stand variables on H-D relationship indicate that additional predictor variables are required to develop generalized H-D models in order to avoid having to establish individual H-D relationships for every stand (Temesgen and Gadow ,2004; Sharma and Parton, 2007). Furthermore, the predictive accuracy increases for generalized H-D models as compared to base models because the former captures stand/plot to stand/plot variability through the addition of other tree and/or stand variables.

2.3 Mixed-Effects Modelling in Height-Diameter Models

Although other techniques have been used and are still in use especially in Africa, many recent modelling efforts have used mixed-effects approach in order to improve the accuracy of prediction of tree height using H-D models. A mixed-effects model is a model that incorporates both fixed-effects, which are parameters associated with an entire population or with certain repeatable levels of experimental factors, and random-effects, which are parameters associated with individual experimental units drawn at random from a population (Pinheiro and Bates, 2000). Thus, mixed-effects models estimate both fixed and random parameters simultaneously for the same model (Calama and Montero, 2004). In this regard, mixed-effects models offer several advantages over OLS, ONLS, and other approaches because:

1. Mixed-effects models can incorporate the hierarchical structure of data (e.g. trees, plots, stands) into the analysis and thus reduce interdependence among measurements from the same sampling unit;

2. Mixed-effects models are a compromise between fitting global models with few parameters and that do not include variability among sampling units (i.e. stand or plot), and local models specific to each sample unit that have numerous and often inter-correlated parameters;

3. Mixed-effects models provide an unbiased estimation of model parameters for sample units with small sample sizes since the variation in the parameter estimates is known at each level of the hierarchical sampling structure; and

4. Mixed-effects models can be calibrated for new, previously unsampled plots or stands quickly and effectively (Saunders and Wagner, 2008).

Sharma and Parton (2007) studied H-D equations for boreal tree species in Ontario. They compared models with and without random effects parameters and using stand level variables in addition to DBH. Their results showed that inclusion of random parameter consistently resulted in a better fit to the data and an improved tree height prediction accuracy. Budhathoki *et al* (2008) also compared mixed-effects model approach and OLS in modelling shortleaf pine growth in even-aged stands. The findings revealed that mixed-effects models, with the inclusion of stand variables, predicted the total tree height better than the similar models developed previously for the same species using OLS methods. Saunders and Wagner (2008) developed H-D models for tree species of Central Maine in northeastern United States, using both generalized nonlinear least squares (GNLS) and mixed-effects approaches. They observed that generally the mixed-effects approaches were superior to GNLS, with inclusion of site covariates (tree density and basal area) accounting for some of the variability explained by the random coefficients in the full mixed-effects models. Huang *et al* (2009a; 2009b) compared the predictive accuracy of base models and expanded models (including stand level variables) using mixed-effects modelling for both sets of models. They observed that a base model with only DBH as predictor variable had a better predictive accuracy than expanded models at plot-specific level. Paulo *et al* (2011) compared nonlinear fixed-effects model (NLFEM) and nonlinear mixed-effects model (NLMEM) approaches in developing a generalized H-D model for Portuguese cork oak stands. They concluded that even in situations where the NLMEM calibration was not possible, the model should be preferred. Huang *et al* (2013) compared base models developed using ONLS and mixed-effects modelling techniques for major Alberta tree species. They observed that base models developed using nonlinear mixed-effects technique had a better predictive accuracy both at overall and plot-by-plot basis when compared to same models developed using ONLS. Missanjo and Mwale (2014) also used mixed-effects modelling to develop an H-D model for *Pinus kesiya* in Malawi. Realistic height predictions were obtained when a mixed-effects model, which included a random parameter and stand level variables was fitted to the data collected. Similar results of better height predictions using mixed-effects models were obtained by Xu *et al* (2014) when they compared a model using only fixed-effects parameters with a nested two level (plot and stand) nonlinear mixed-effects model. Carrol-Rivas *et al* (2014) compared models with and without random parameters in developing H-D functions for mixed, uneven-aged stands in northwestern Durango (Mexico) by considering DBH and stand variables as predictors. Their results revealed that mixed-effects models performed better than same models that did not have random parameters.

2.4 Summary of Literature Review and Knowledge Gaps

The H-D relationships referred to in the literature review can be written in a general form as:

$$H = f(DHB) \tag{2.1}$$

for a base model, where H is tree height; f denotes some nonlinear (and rarely linear) function; and DBH is the tree diameter at breast height. The relationship in (2.1) can be expanded to a generalized H-D model that includes other variables as follows:

$$H = f(DBH + othervariables) \tag{2.2}$$

where other variables can mean other tree and/or stand variables such as; tree age, stand age, dominant height, site index, stems per hectare, basal area, dominant height, top height, crown ratio, species composition, bio-geo-climatic variables, and soil related factors (Huang *et al.*, 2013). The aforementioned studies show that the predictive accuracy of a base model can be improved by addition of statistically significant tree and/or stand level attributes even in cases where mixed-effects modelling approach has not been applied. However, other studies also show that mixed-effects modelling approach leads to improved prediction ability and accuracy of H-D models. This implies that the predictive ability and accuracy of an H-D model can be significantly improved by addition of statistically significant tree and stand level variables and also applying mixed-effects modeling approach to the developed generalized H-D models, as stated in some of the foregoing studies. It is however worth noting that the inclusion of other tree and stand variables to a base model might necessitate added costs and time in terms of field sampling, data management and analysis. Most of the studies outlined here considered comparison of mixed-effects models that included stand level variables in addition to DBH. This study however focused on ascertaining the extent of improvement in predictive accuracy of a base model by applying mixed-effects modeling approach with only tree DBH as a predictor variable. The study is similar in many respects to those of Huang *et al* (2009a; 2009b) except the present study compares a base model fitted to the data using ONLS to the same model fitted to same data using nonlinear mixed-effects modeling. Furthermore, Huang *et al* (2009a; 2009b) compared a base model that had only DBH as predictor variable with expanded models that included stand level variables. The present study is also similar to Huang *et al* (2013) but uses a different formulation of the mixed-effects model. It is also worth noting that only Huang *et al* (2009a; 2009b) compared the models on plot-by-plot basis, where the main interest of a mixed-effects model lies; an approach not used by previous researchers. This study is also similar to that of Carrol-Rivas *et al* (2014) in that they also compared a base model developed using ONLS with that developed using mixed-effects modelling. However, a plot-by-plot analysis was not included and only overall summary fitting statistics presented. The present study is thus among the few that also compares models on a plot-by-plot basis in addition to comparison at population level (which is the common practise used by most researchers).

Chapter 3

Methodology

3.1 Sources of Data

The data sets used for both model fitting and testing were obtained from the website of the United States Department of Agriculture (USDA) Forest Service, Forest Inventory and Analysis Database (FIADB) (US Forest Service, 2015). Model fitting data set comprised of annual surveys for the state of Arkansas, surveys conducted from 2004 to 2010, while model testing data set was for the state of South Carolina, surveys conducted from 2004 to 2013. Both states fall under the Southern Region Research Station according to the USDA FIA (Forest Inventory and Analysis) work unit classification (O'Connell *et al.*, 2013). Data is collected on a plot which is a cluster of four points approximately 0.02 ha (1/24 acre) each in size with radius of 7.32 m (24 feet). Subplot 1 is central with the rest; that is subplots 2, 3 and 4 located 36.58 m (120 feet) horizontal at azimuths of 360, 120, and 240 degrees from the center of subplot 1, respectively as shown in Figure 3.1. Thus, a plot refers to the entire set of four subplots (US Forest Service, 2011).

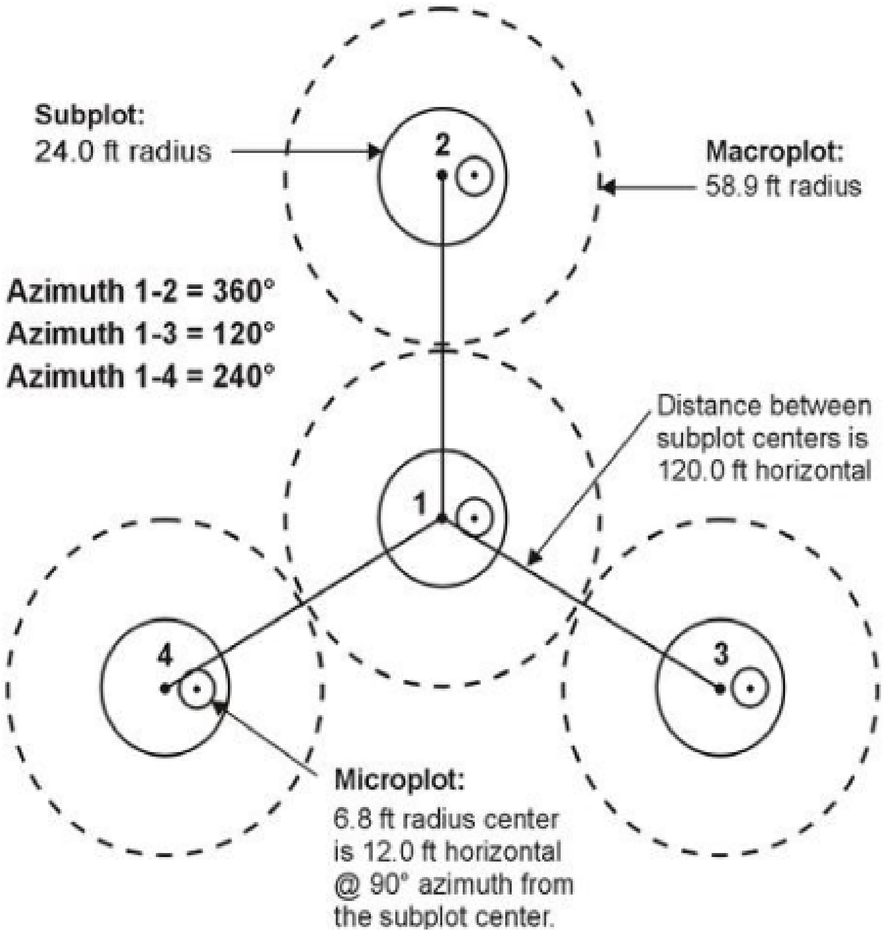

Figure 3.1: The FIA Mapped Plot Diagram
Source: US Forest Service, 2011.

Each cluster point may be surrounded by a 17.96 m (58.9 ft) fixed-radius macroplot where, generally, only trees with DBH of 60.96 cm and larger are measured. The four macroplots combined total approximately 0.4047 ha (or 1 acre). Each cluster point is surrounded by a 7.32 m (24 ft) fixed-radius subplot where trees with DBH of 12.70 cm and larger are measured. The four subplots when combined total approximately 0.0672 ha (or 1/6th acre). Each sublopt contains a 2.07 m (6.8 ft) fixed-radius microplot where only saplings with DBH ranging from 2.54 cm to 12.45 cm is measured. When combined, the four microplots total approximately 0.005ha (or 1/75th acre) (VanderSchaaf, 2014).

In terms of inclusion of trees in sample data sets, only live trees without broken tops and physically measured in the field were included as opposed to trees visually estimated or predicted using equations. Furthermore, only trees and/or plots that had been measured once were included in the sample. Lastly, only one tree species; that is Shortleaf pine (*Pinus echinata* Mill.) was considered for both model fitting and testing data sets. The structure of data used in this study, for both model fit and test, was as shown in Table 3.1.

Table 3.1: Data Structure

Plot ID(i)	Tree(j)	DBH(cm)	H(m)
1	1	17.78	8.84
1	2	14.22	9.88
$\vdots$	$\vdots$	$\vdots$	$\vdots$
2	1	13.72	14.02
2	2	20.57	17.98
2	3	14.22	12.50
$\vdots$	$\vdots$	$\vdots$	$\vdots$
m	n_i	.	.

where, m is the last plot number, i is a specific plot number, $i = 1, 2, 3, ..., m$; n_i is the number of trees in the ith plot; j is a specific tree number in the ith plot , $j = 1, 2, ..., n_i$; DBH (cm) is tree diameter at breast height measured in centimetres and H (m) is tree height measured in meters.

3.2 Selection of Base Height-Diameter Model

Most studies conducted indicate that height-DBH relationships for various tree species normally exhibit a typical sigmoidal-concave curve when total tree height to DBH is plotted. This was also evidenced in the data used for this study. Through literature review, the Chapman-Richards model, which is a nonlinear function with appropriate mathematical and biological properties, was chosen as a base model for this study. One major difference between linear and nonlinear models is that the latter are generally mechanistic in nature; that is, a nonlinear model is constructed based on the mechanism producing the response. Consequently, the parameters in a nonlinear model generally have a physical interpretation (Pinheiro and Bates, 2000). In line with this, the Chapman-Richards model defines a sigmoid curve; which means that as the size of the tree increases, the growth rate also increases from a minimum value to a maximum at a point of inflection and then declines towards zero of the upper asymptote. Furthermore, the model has three parameters representing the upper asymptote, a rate parameter and a shape parameter. These three parameters characterize the different biological processes and growth behaviours (Peng *et al.*, 2001; Lumbers *et al.*, 2013). It is also worth noting that satisfactory results, as a result of use of this model, have been reported in literature for various tree species in different regions across the globe (e.g. Zhang, 1997; Sharma and Parton 2007; Huang *et al.*, 2009a; 2009b; Lumbers *et al.*., 2013; Corral-Rivas *et al.*, 2014). The Chapman-Richards model is defined by the following expression:

$$H = 1.30 + \beta_1(1 - \exp(-\beta_2 DBH))^{\beta_3} \tag{3.1}$$

where, H= tree height (m); DBH= diameter at breast height (cm); 1.30= a constant used to account for measuring tree DBH at 1.30m above ground; β_1, β_2 and β_3= model parameters to be estimated; exp= exponential - base of the natural logarithm (≈ 2.71828).

3.3 Estimation of Parameters for the Base Model

Model (3.1) is a nonlinear function which generally may be expressed as (Rawlings *et al.*,1998; Greene, 2012; SAS Institute, 2013):

$$H_j = f(DBH_j; \boldsymbol{\beta}) + \varepsilon_j \tag{3.2}$$

where, H_j is the random variable representing the height for tree j; $f(DBH_j; \boldsymbol{\beta})$ is the nonlinear function relating the expectation of the response variable, $E(H)$, to the independent variable; DBH_j is the observed DBH for the jth tree; $\boldsymbol{\beta}$ is the vector of p parameters and ε_j is the random error associated with the jth tree; $j = 1, 2, ..., n$. The basic assumptions are made on the random errors; they are assumed to be independent random variables with mean 0 and a constant variance (σ^2).

Parameters of (3.2) are obtained as nonlinear least squares estimators that minimize the sum of squares of the residuals:

$$SSResiduals(\hat{\boldsymbol{\beta}}) = \sum_{j=1}^{n}(H_j - f(DBH_j; \boldsymbol{\beta}))^2. \tag{3.3}$$

In matrix form, (3.3) may be represented as:

$$SSResiduals(\hat{\boldsymbol{\beta}}) = (\mathbf{H} - \mathbf{f}(\hat{\boldsymbol{\beta}}))'(\mathbf{H} - \mathbf{f}(\hat{\boldsymbol{\beta}})) \tag{3.4}$$

where, $\mathbf{f}(\hat{\boldsymbol{\beta}})$ is the $n \times 1$ vector of $f(\mathbf{DBH}; \boldsymbol{\beta})$ evaluated at the n values of DBH_j.

Under the assumption that the random errors are independent variables with mean 0 and a constant variance, the nonlinear least squares estimate of $\boldsymbol{\beta}$ is also the maximum likelihood estimate of $\boldsymbol{\beta}$. The first order conditions for the minimization of (3.3) are the partial derivates of the sum of squares of the residuals with respect to each $\hat{\beta}_k$ in turn and equated to zero in order to obtain the p normal equations. The solutions to the normal equations give the nonlinear least squares of $\boldsymbol{\beta}$. The general form of each normal equation is expressed as (Rawlings *et al.*, 1998):

$$\frac{\partial SSResiduals(\hat{\boldsymbol{\beta}})}{\partial \hat{\beta}_k} = -\sum_{j=1}^{n}(H_j - f(DBH_j; \hat{\boldsymbol{\beta}}))\left(\frac{\partial f(DBH_j; \hat{\boldsymbol{\beta}})}{\partial \hat{\beta}_k}\right) = 0 \tag{3.5}$$

where, $\left(\frac{\partial f(DBH_j; \hat{\boldsymbol{\beta}})}{\partial \hat{\beta}_k}\right)$ is the partial derivative of the functional form of the model. The main difference between linear models and nonlinear models is that for the case of the latter, the partial derivatives are functions of the parameters; the resulting normal equations are not in closed form and thus cannot be simply solved by equating to zero (0) in order to get explicit solutions for $\hat{\boldsymbol{\beta}}$. For the case of (3.1) used in this study, the partial derivatives of the functional form of the model with respect to the three parameters are:

$$\frac{\partial H}{\partial \beta_1} = (1 - \exp(-\beta_2 DBH))^{\beta_3};$$

$$\frac{\partial H}{\partial \beta_2} = \beta_1 DBH \beta_3 \exp(-\beta_2 DBH)(1 - \exp(-\beta_2 DBH))^{(\beta_3 - 1)};$$

$$\frac{\partial H}{\partial \beta_3} = \beta_1 \log(1 - \exp(-\beta_2 DBH))(1 - \exp(-\beta_2 DBH))^{\beta_3}.$$

As can be observed with the three partial derivatives for (3.1), each one is a function of parameters. Therefore, the resulting normal equations do not have closed-forms and thus explicit solutions for the parameters cannot be obtained by simply equating the normal equations to zero. Instead, iterative numerical methods are used in order to get the explicit estimates for the parameters in a nonlinear model. The five commonly used iterative numerical methods are: Steepest Descent (gradient), Gauss-Newton, Marquardt, Newton and the Multivariate Secant of False Position or Derivative Free (DUD) method. The use of any of the five methods require that starting values of the parameters (herein referred to as $\boldsymbol{\beta}^0$) are specified. The starting values are then substituted for $\boldsymbol{\beta}$ to compute the residual sums of squares and also adjustments to $\boldsymbol{\beta}^0$ that will reduce the residual sums of squares. The process is repeated until there are no further adjustments on $\boldsymbol{\beta}^0$ for successive steps and the process is said to have converged (Rawlings *et al.*, 1998; SAS Institute, 2013). In the framework of nonlinear regression, convergence implies that the best estimates of the parameters have been obtained under the assumption that the model is adequate (Bates and Watts, 1988).

In this study, the Gauss-Newton and Marquardt methods were used and parameter estimates compared. Both methods regress the residuals onto the partial derivatives of the model with respect to the parameters until the estimates converge. The Marquardt method, which can be viewed as a Gauss-Newton algorithm with a ridging penalty, appears to work well in most cases (Rawlings *et al.*, 1998; SAS Institute, 2013). Since both methods gave the same solutions in this study, the estimation procedure for the Gauss-

Newton method is hereby outlined.

The Gauss-Newton method uses the Taylor series expansion of $f(DBH_j;\hat{\boldsymbol{\beta}})$ about the starting values, $\boldsymbol{\beta}^0$, in order to obtain a linear approximation of the nonlinear model in the region near the starting values. The term $f(DBH_j;\hat{\boldsymbol{\beta}})$ is replaced by a linear approximation as follows:

$$f(DBH_j;\boldsymbol{\beta}) \approx f(DBH_j;\boldsymbol{\beta}^0) + \sum_{k=1}^{p} \left(\frac{\partial f(DBH_j;\boldsymbol{\beta}^0)}{\partial \beta_k} \right) (\beta_k - \beta_k^0)$$

or in matrix form as:

$$\mathbf{f}(\boldsymbol{\beta}) \approx \mathbf{f}(\boldsymbol{\beta}^{\mathbf{0}}) + \mathbf{F}(\boldsymbol{\beta}^{\mathbf{0}})(\boldsymbol{\beta} - \boldsymbol{\beta}^{\mathbf{0}}) \tag{3.6}$$

where, $\mathbf{F}(\boldsymbol{\beta}^{\mathbf{0}})$ is an $n \times p$ matrix of partial derivatives evaluated at $\boldsymbol{\beta}^0$ and n data points, DBH_j.The matrix has the form:

$$\mathbf{F}(\boldsymbol{\beta}^{\mathbf{0}}) = \begin{bmatrix} \frac{\partial[f(DBH_1;\boldsymbol{\beta}^0)]}{\partial \beta_1} & \frac{\partial[f(DBH_1;\boldsymbol{\beta}^0)]}{\partial \beta_2} & \cdots & \frac{\partial[f(DBH_1;\boldsymbol{\beta}^0)]}{\partial \beta_p} \\ \frac{\partial[f(DBH_2;\boldsymbol{\beta}^0)]}{\partial \beta_1} & \frac{\partial[f(DBH_2;\boldsymbol{\beta}^0)]}{\partial \beta_2} & \cdots & \frac{\partial[f(DBH_2;\boldsymbol{\beta}^0)]}{\partial \beta_p} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial[f(DBH_n;\boldsymbol{\beta}^0)]}{\partial \beta_1} & \frac{\partial[f(DBH_n;\boldsymbol{\beta}^0)]}{\partial \beta_2} & \cdots & \frac{\partial[f(DBH_n;\boldsymbol{\beta}^0)]}{\partial \beta_p} \end{bmatrix}. \tag{3.7}$$

Equation (3.6) is called the linearised regression model (Greene, 2010). Linear least squares technique is then applied to (3.6) in order to get estimated amount to adjust starting values. This is done by regressing $\mathbf{H} - \mathbf{f}(\beta^{\mathbf{0}})$ on $\mathbf{F}(\boldsymbol{\beta}^{\mathbf{0}})$. New values of the parameters are obtained by adding the estimated amount of adjustment to the previous starting values. The model is then linearised about the new values of the parameters and linear least squares again applied to find the second set of adjustments. The process is repeated until the desired degree of convergence is achieved. The matrix $\mathbf{F}(\boldsymbol{\beta}^{\mathbf{0}})$, which may be written as $\mathbf{F}$ for brevity, plays the role in nonlinear least squares that $\mathbf{X}$ plays in linear least squares (Rawlings *et al.*, 1998). The adjustment or update vector of parameter estimates ($\Delta = \boldsymbol{\beta} - \hat{\boldsymbol{\beta}}$) for Gauss-Newton is given by (SAS Institute, 2013):

$$\Delta = (\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'(\mathbf{H} - \mathbf{f}(\hat{\boldsymbol{\beta}}))$$

while for the Marquardt method is given by:

$$\Delta = (\mathbf{F}'\mathbf{F} + \lambda * \mathbf{diag}(\mathbf{F}'\mathbf{F}))^{\mathbf{-1}}\mathbf{F}'(\mathbf{H} - \mathbf{f}(\hat{\boldsymbol{\beta}}));$$

where, λ is a Lagrange multiplier or algorithmic parameter ; $\lambda \geq 0$ (Marquardt, 1963; Gavin, 2013).

The Marquardt method is a compromise between the steepest descent and Gauss-Newton methods where if $\lambda \longrightarrow 0$ the method tends to approach Gauss Newton while if $\lambda \longrightarrow \infty$ the method approaches steepest descent. If $SSResiduals(\hat{\boldsymbol{\beta}})$ decreases on each iteration, then $\lambda \longrightarrow 0$ while if $SSResiduals(\hat{\boldsymbol{\beta}})$ does not improve, then λ is increased and one would be essentially moving towards steepest descent method (SAS Institute, 2013).

If the usual assumptions about the error term, ε_j, in equation (3.2) are satisfied, that is $\boldsymbol{\varepsilon} \sim N(\mathbf{0}, \mathbf{I}\sigma^{\mathbf{2}})$, $\hat{\boldsymbol{\beta}}$ is approximately normally distributed with mean $\boldsymbol{\beta}$ and var $(\hat{\boldsymbol{\beta}}) = (\mathbf{F}'\mathbf{F})^{-1}\sigma^2$:

$$\hat{\boldsymbol{\beta}} \sim N[\boldsymbol{\beta}, (\mathbf{F}'\mathbf{F})^{-1}\sigma^2].$$

In the case where normality assumptions are not satisfied for $\boldsymbol{\varepsilon}$, $\hat{\boldsymbol{\beta}}$ can still be shown to be asymptotically normally distributed as n, the data points, gets larger.

With the values $\boldsymbol{\beta}$ obtained as $\hat{\boldsymbol{\beta}}$, σ^2 is estimated as:

$$s^2 = \frac{SSResiduals(\hat{\boldsymbol{\beta}})}{n-p};$$

and thus the variance-covariance matrix for $\hat{\boldsymbol{\beta}}$ is estimated as:

$$\mathbf{s}^2(\hat{\boldsymbol{\beta}}) = (\hat{\mathbf{F}}'\hat{\mathbf{F}})^{-1}s^2. \tag{3.8}$$

The standard errors for $\hat{\boldsymbol{\beta}}$ are estimated using (3.8)(Rawlings *et al.,* 1998) for the gradient, Marquardt and Gauss-Newton methods for unconstrained estimates (SAS Institute, 2013).

The base model (3.1) was fitted to the model fitting data set and parameters estimated using ONLS using the PROC NLIN procedure in SAS. Parameter estimates obtained by other researchers were used as starting values.

3.4 Motivation for use of Mixed-Effects Modelling Approach

Quite often, the data that is used in modelling H-D relationship contain measurements of height and DBH from multiple trees from a given sample plot; although individual plots are generally located within different stands (Sharma and Parton, 2007). Such a nested structure, that is trees within plots and plots within stands, results in a lack of independence between observations since data from the same sampling unit tend to be more correlated that the average. The lack of independence between observations results in biased estimates of the confidence intervals of the parameters if ordinary least squares regression technique is used (Dorado *et al* ., 2006). Mixed-effects modeling approach can be applied to deal with such correlated observations (Dorado *et al* ., 2006; Sharma and Parton, 2007). The data used in this study comprised of tree measurements taken in different plots across the state/county and as such conformed to a nested data structure. Therefore, the chosen base model, Chapman-Richards model, was subjected to nonlinear mixed-effects (NLME) modelling.

3.4.1 Mixed-Effects Model Formulation and Assumptions

A general expression for a NLME model, at one level of grouping (plot in this study), can be defined as (Pinheiro and Bates, 2000; Dorado *et al* ., 2006, Huang *et al* ., 2009; Xu *et al* ., 2014):

$$\mathbf{H}_i = f(\mathbf{DBH}_i; \boldsymbol{\Phi}_i) + \boldsymbol{\varepsilon}_i \tag{3.9}$$

where, $\mathbf{H}_i$ is the $n_i \times 1$ vector of the n_i observations of the tree heights taken from the ith sampling unit(plot in this study); f =a general, real-valued, differentiable nonlinear function of the predictor variable and the parameter vector; $\mathbf{DBH}_i$ is the $n_i \times 1$ vector of the predictor variable (DBH) for the n_i trees taken from the ith plot; $\boldsymbol{\Phi}_i$ is a parameter vector ($r \times 1$; where r is the number of parameters in the model) which is specific for each plot; and $\boldsymbol{\varepsilon}_i$ is a $n_i \times 1$ vector for the residual terms.

The main feature of mixed-effects models is that they allow parameter vectors to vary from plot to plot; that is, regression coefficients are broken down into fixed part, common to the population, and random components, specific to each plot/stand (Calama and Montero, 2004; Dorado *et al.*, 2006). Therefore, the parameter vector $\boldsymbol{\Phi}_i$ in (3.9) can be succinctly expressed as:

$$\boldsymbol{\Phi}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i$$

where, $\boldsymbol{\beta}$ is the $p \times 1$ vector of fixed population parameters (p =number of fixed parameters in the model); $\mathbf{b}_i$ is a $q \times 1$ vector of random-effects associated with the ith plot (q is the number of random parameters in the model); $\mathbf{X}_i$ and $\mathbf{Z}_i$ are design matrices of size $r \times p$ and $r \times q$ for the fixed and random-effects specific for each plot, respectively. The elements for these design matrices are usually $0, 1$, or the value of the covariate(s) associated with the fixed and/or random-effects. The basic assumption for the nonlinear mixed-effect models theory include the asymptotic multivariate normal distribution for the random-effects vector, the residual terms vector and the observations of the response variable vector (Calama and Montero, 2004; Dorado *et al* ., 2006). Furthermore, it is assumed that $\boldsymbol{\varepsilon}_i$ and $\mathbf{b}_i$ are uncorrelated. Mathematically, the foregoing assumptions imply that:

$$E\begin{bmatrix}\mathbf{b}_i \\ \\ \boldsymbol{\varepsilon}_i\end{bmatrix} = \begin{bmatrix}0 \\ \\ 0\end{bmatrix};$$

$$Var\begin{bmatrix}\mathbf{b}_i \\ \\ \boldsymbol{\varepsilon}_i\end{bmatrix} = \begin{bmatrix}\mathbf{D} & 0 \\ \\ 0 & \mathbf{R}_i\end{bmatrix};$$

$$\mathbf{H}_i \sim MVN(\mathbf{X}_i\boldsymbol{\beta}, \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \mathbf{R}_i);$$

where, $\mathbf{D}$ is a positive-definite variance-covariance matrix of size $q \times q$ for the random-effects, representing among-plot variability and assumed to be the same for every plot; and $\mathbf{R}_i$ is the variance-covariance matrix of size $n_i \times n_i$, defining the within-plot variability. Equation (3.9) may also simply be expressed as follows:

$$\mathbf{H}_i = f(\mathbf{DBH}_i, \boldsymbol{\beta}, \mathbf{b}_i) + \boldsymbol{\varepsilon}_i \tag{3.10}$$

where, $\mathbf{H}_i = [H_{i1}, H_{i2}, ..., H_{in_i}]'$ is a vector of measurements for tree heights in the ith plot ; $\mathbf{DBH}_i$ is a known design matrix of the DBH measurements in the ith plot; $\boldsymbol{\beta} = [\beta_1, \beta_2, \beta_3]'$ is a vector of fixed-parameters; $\mathbf{b}_i = [b_{1i}, b_{2i}, b_{3i}]'$ is a vector of random parameters specific to plot i; $\boldsymbol{\varepsilon}_i = [\varepsilon_{i1}, \varepsilon_{i2}, ..., \varepsilon_{in_i}]'$ is a vector of within-plot errors; and n_i is the total number of trees in the ith plot for which corresponding individual tree height and DBH measurements had been observed.

3.4.2 Mixed-Effects Model Development

Generally, there are three steps that are necessary for constructing a mixed-effects model, once an appropriate base model has been selected (Calama and Montero, 2004; Dorado *et al.*, 2006; Sharma and Parton, 2007; Missanjo and Mwale, 2014; Xu *et al.*, 2014):

1. Specification of the nature of the parameters of the model to be treated as mixed (that is, made up of fixed and random-effects) or purely fixed-effects;
2. Determination of an appropriate within-plot variance-covariance structure ($\mathbf{R}_i$), for explaining variability among trees in the same plot; and
3. Determination of among-plot variance-covariance structure ($\mathbf{D}$).

Determination of fixed and random effects parameters in a model is a flexible decision subject to debate. All parameters in the model are first considered mixed if convergence is possible. If convergence is not achieved in first step, the number of random effects parameters is then systematically reduced until convergence is attained (Sharma and Parton, 2007; Missanjo and Mwale, 2014).

To account for within-plot heteroscedasticity in $\mathbf{R}_i$, which includes weighting factors, the approach as given by Calama and Montero(2004) and cited in Xu *et al.* (2014) is given by:

$$\mathbf{R}_i = \sigma^2 \mathbf{G}_i^{0.5} \mathbf{I}_{n_i} \mathbf{G}_i^{0.5}$$

where, σ^2 is a scaling factor for the error dispersion; $\mathbf{G}_i$ is an $n_i \times n_i$ diagonal matrix within-plot error heteroscedasticity variances and; $\mathbf{I}_{n_i}$ is an $n_i \times n_i$ matrix showing the within-plot autocorrelation structure of error.

A plot of residuals versus predicted values did not show obvious pattern of unequal error variance. Furthermore, the study only used trees and/ or plots that had been measured once, and hence autocorrelation was not considered or assumed to exist. Therefore, $\mathbf{R}_i$ was assumed to be $\sigma_\varepsilon^2\mathbf{I}_{n_i}$; where σ_ε^2 is the estimated error variance of the model and $\mathbf{I}_{n_i}$ is an $n_i \times n_i$ identity matrix. Thus:

$$\mathbf{R}_i = \sigma_\varepsilon^2\mathbf{I}_{n_i}. \tag{3.11}$$

Determination of $\mathbf{D}$ was carried out by first treating all the three parameters in (3.1) as mixed; that is, made of a fixed and random part. Unfortunately, the model did not converge. The model only converged and realistic parameter estimates were obtained when random parameters were added to β_1 and β_3. Therefore, $\mathbf{D}$ was assumed to be an unstructured covariance matrix that is the same for all i. Thus:

$$\mathbf{D} = \begin{bmatrix} \sigma_{b_1}^2 & \sigma_{b_1b_3} \\ \sigma_{b_1b_3} & \sigma_{b_3}^2 \end{bmatrix} \tag{3.12}$$

where, $\sigma_{b_1}^2$ and $\sigma_{b_3}^2$ are the variances for random parameters b_1 and b_3 respectively; and $\sigma_{b_1b_3}$ is the covariance between b_1 and b_3.

Therefore, the nonlinear mixed-effects model corresponding to (3.1) took the following form:

$$H_{ij} = 1.30 + (\beta_1 + b_{1i})(1 - \exp(-\beta_2 DBH_{ij}))^{(\beta_3 + b_{3i})} + \varepsilon_{ij} \tag{3.13}$$

where, H_{ij} and DBH_{ij} are the observed tree height and DBH for the jth tree in the ith plot; $i = 1, 2, 3, ..., m$; $j = 1, 2, 3, ..., n_i$; m is the total number of plots; n_i is the number of trees in the ith plot; β_1, β_2 and β_3 are the fixed-effects parameters common to every plot; and b_{1i} and b_{3i} are the unique random-effects parameters unique for each plot; and ε_{ij} is the within-plot error term of the model, assumed to be normally distributed with mean zero and constant variance.

3.4.3 Estimation of Parameters for the Mixed-Effects Model

The aim of mixed-effects modelling is estimation of the components:$\boldsymbol{\beta}$,$\mathbf{D}$ and $\mathbf{R}_i$. In this particular study, estimation of the parameters was based on methods based on likelihood function. Random-effects are unobserved quantities and as such maximum likelihood estimation is based on numerically maximizing an approximation to the marginal likelihood - that is, the likelihood integrated over the random effects (Pinheiro and Bates, 2000; Littell *et al* ., 2006). With reference to (3.9) and (3.10), the marginal density for the responses, $\mathbf{H}$, is given by (Pinheiro and Bates, 2000; Calama and Montero, 2004):

$$p(\mathbf{H}|\mathbf{DBH},\boldsymbol{\beta},\sigma^2,\mathbf{D}) = \prod_{i=1}^{m}\int p(\mathbf{H}_i|\mathbf{DBH}_i,\boldsymbol{\beta},\sigma^2,\mathbf{b}_i)p(\mathbf{b}_i|\mathbf{D})d\mathbf{b}_i \qquad (3.14)$$

where, $p(\mathbf{H}|\mathbf{DBH},\boldsymbol{\beta},\sigma^2,\mathbf{D})$ is the marginal density of $\mathbf{H}$; $p(\mathbf{H}_i|\mathbf{DBH}_i,\boldsymbol{\beta},\sigma^2,\mathbf{b}_i)$ is the conditional density of $\mathbf{H}_i$ given the random effects $\mathbf{b}_i$ and; $p(\mathbf{b}_i|\mathbf{D})$ is the marginal distribution of $\mathbf{b}_i$; m is the number of sampling units (plots in the case of this study); and σ^2 is the variance of the error term of the model. Maximization of expression (3.14) is not an easy task and in most cases does not have a closed form because the random-effects enter the model in a nonlinear fashion. Therefore, in order to make the numerical optimization of expression (3.14) mathematically tractable, different approximations have been proposed with the common ones being: (1) first-order Taylor series expansion of the model function around the expected value of the random effects, or (2) around the conditional modes of the random effects (first-order conditional expectation method of Lindstrom and Bates) and (3) Gaussian quadrature, which tries to solve the integral numerically (Pinheiro and Bates, 2000; Calama and Montero, 2004). In this study, the first-order Taylor series expansion approximation was adopted.

The Taylor series expanded version of equation (3.10), with the quadratics and cross-products dropped, is given by (Huang *et al* ., 2009a; 2009b):

$$\mathbf{H}_i \approx f(\mathbf{DBH}_i,\boldsymbol{\beta}^*,\mathbf{b}_i^*) + \mathbf{X}_i(\boldsymbol{\beta}-\boldsymbol{\beta}^*) + \mathbf{Z}_i(\mathbf{b}_i-\mathbf{b}_i^*) + \boldsymbol{\varepsilon}_i \qquad (3.15)$$

where, $\boldsymbol{\beta}^*$ is a vector of starting values or values close to $\boldsymbol{\beta}$ and $\mathbf{b}_i^*$ is set to zero($\mathbf{0}$), which is the expected value of the random parameters. The derivative matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$ are defined by:

$$\mathbf{X}_i = \left.\frac{\partial f(\mathbf{DBH}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \boldsymbol{\beta}'}\right| \boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}}, \mathbf{b}_i^* = \mathbf{0}; \tag{3.16}$$

$$\mathbf{Z}_i = \left.\frac{\partial f(\mathbf{DBH}_i, \boldsymbol{\beta}, \mathbf{0})}{\partial \mathbf{b}_i'}\right| \boldsymbol{\beta}^* = \hat{\boldsymbol{\beta}}, \mathbf{b}_i^* = \mathbf{0}. \tag{3.17}$$

Given that $\mathbf{b}_i^* = \mathbf{0}$, a pseudo-response function, $\mathbf{H}_i^*$, can be defined as:

$$\mathbf{H}_i^* = \mathbf{H}_i - f(\mathbf{DBH}_i, \boldsymbol{\beta}^*, \mathbf{0}) + \mathbf{X_i}\boldsymbol{\beta}^*, \tag{3.18}$$

and rearranging terms in equation (3.15), we get:

$$\mathbf{H}_i^* = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i. \tag{3.19}$$

Equation (3.19) shows that the pseudo-response vector,$\mathbf{H}_i^*$, is linear both in $\boldsymbol{\beta}$ and $\mathbf{b}_i$ and thus can be solved using linear mixed-effects model theory to get estimates for the parameters (Huang *et al.*, 2009b).

A general expression for a linear mixed-effects model can be written as (Pinheiro and Bates, 2000; Gurka, 2006):

$$\mathbf{H}_i = \mathbf{X}_i\boldsymbol{\beta} + \mathbf{Z}_i\mathbf{b}_i + \boldsymbol{\varepsilon}_i \tag{3.20}$$

where, $\mathbf{H}_i$ is a $n_i \times 1$ vector of observations(trees) on the ith subject (plot); $\mathbf{X}_i$ is a $n_i \times p$ known, constant design matrix for the ith subject with rank p (the number of fixed-effects parameters); $\boldsymbol{\beta}$ is a $p \times 1$ vector of unknown, constant population (fixed-effects) parameters; $\mathbf{Z}_i$ is a $n_i \times q$ known, constant design matrix for the ith subject with rank q corresponding to $\mathbf{b}_i$, a $q \times 1$ vector of unknown, random subject-specific parameters; and finally $\boldsymbol{\varepsilon}_i$ is a $n_i \times 1$ vector of random within-subject error terms. The linear mixed-effects model (3.20) has the same basic distributional assumptions as the NLME model given for equation (3.9), that is; $\mathbf{b}_i$ is normally distributed with mean vector 0 and variance-covariance matrix $\mathbf{D}$; and $\boldsymbol{\varepsilon}_i$ is normally distributed with mean vector 0 and variance-covariance matrix $\mathbf{R}_i$ and finally, $\mathbf{b}_i$ and $\boldsymbol{\varepsilon}_i$ are independent. The variance-covariance matrices $\mathbf{D}$ and $\mathbf{R}_i$ are characterized by unique parameters contained in the $r \times 1$ vector, $\boldsymbol{\theta}$. The total variance for the response vector is $\mathbf{V}_i = \mathbf{Z}_i\mathbf{D}\mathbf{Z}_i' + \mathbf{R}_i$. The marginal log-likelihood function for (3.20) is given by:

$$l_{ML}(\boldsymbol{\theta}) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\sum_{i=1}^{m}\log|\mathbf{V}_i| - \frac{1}{2}\sum_{i=1}^{m}(\mathbf{H}_i - \mathbf{X}_i\boldsymbol{\beta})'\mathbf{V}_i^{-1}(\mathbf{H}_i - \mathbf{X}_i\boldsymbol{\beta}) \tag{3.21}$$

where, $N = \sum_{i=1}^{m} n_i$ - the total number of observations (tree measurements) in the dataset; m and n_i are the number of plots and total number of trees in the ith plot, respectively. Maximization of $l_{ML}(\boldsymbol{\theta})$ produces maximum likelihood estimators (MLE) of the unknown parameters. When $\boldsymbol{\theta}$ is known, the MLE of $\boldsymbol{\beta}$ is given by:

$$\hat{\boldsymbol{\beta}} = \left(\sum_{i=1}^{m}\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{X}_i)^{-1}\right)\sum_{i=1}^{m}\mathbf{X}_i'\mathbf{V}_i^{-1}\mathbf{H}_i. \tag{3.22}$$

In most cases, same with this particular study, when $\boldsymbol{\theta}$ is unknown, $\mathbf{V}_i$ is replaced with its estimate $\hat{\mathbf{V}}_i$ (Gurka, 2006). Estimates for $\boldsymbol{\beta}$ and the variance components for the matrices $\mathbf{D}$ and $\mathbf{R}_i$, that is $\hat{\mathbf{D}}$ and $\hat{\mathbf{R}}_i$, respectively were obtained in SAS using the PROC NLMIXED (SAS Institute, 2014). The starting values for the parameters used were obtained from related previous studies reported by other researchers.

Finally, following the linear and normality conditions, $\mathbf{b}_i$ was approximated by the value of the empirical best linear unbiased predictor (EBLUP) for the parameters of the random-effects in plot i, $\hat{\mathbf{b}}_i$. EBLUP indicates that it is not the real best linear unbiased predictor (BLUP), since it is estimated using estimates for the variance components (Calama and Montero, 2004). The equation for the EBLUPS used was (Huang *et al* ., 2009a; Huang *et al* ., 2013; VanderSchaaf, 2014):

$$\hat{\mathbf{b}}_i \approx \hat{\mathbf{D}}\hat{\mathbf{Z}}_i'(\hat{\mathbf{Z}}_i\hat{\mathbf{D}}\hat{\mathbf{Z}}_i' + \hat{\mathbf{R}}_i)^{-1}(\mathbf{H}_i - f(\mathbf{DBH}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})); \tag{3.23}$$

which is the same as:

$$\hat{\mathbf{b}}_i \approx \hat{\mathbf{D}}\hat{\mathbf{Z}}_i'(\hat{\mathbf{Z}}_i\hat{\mathbf{D}}\hat{\mathbf{Z}}_i' + \hat{\mathbf{R}}_i)^{-1}(\mathbf{H}_i - \hat{\mathbf{H}}_{i\_fix}) \tag{3.24}$$

where, $\hat{\mathbf{H}}_{i\_fix}$ is a vector of predicted heights for the ith plot using the fixed-effects parameter estimates obtained using PROC NLMIXED with the random-effects parameters set to 0.

3.5 Prediction of Tree Height using Height-Diameter Models

The main purpose of developing a model is to use it as a predictive tool in forest management (Calama and Montero, 2004). Thus, the use of the two models in this particular study was to predict the dependent variable (total tree height) in terms of the independent variable (DBH) through the relationship specified in each model (3.1) and (3.13). Prediction of tree height was done in three ways on both model fitting and model testing data sets. The first method of prediction of tree height involved using the base model (3.1) and the estimated parameters using ONLS. Thus, the equation used was:

$$\hat{H}_{ij} = 1.30 + \hat{\beta}_1(1 - \exp(-\hat{\beta}_2 DBH_{ij}))^{\hat{\beta}_3} \tag{3.25}$$

where, $\hat{H}_{ij}$ is the predicted height for the jth tree in the ith plot, using the base model (3.1); $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ are the estimated model parameters using ONLS; and DBH_{ij} is the observed DBH for the jth tree in the ith plot.

There are two situations in which mixed-effects models are used in a predictive role (Calama and Montero, 2004; Dorado *et al.*, 2006; Sharma and Parton, 2007):

1. Fixed-effects response pattern: prediction of tree height in stands where only DBH and/or stand variables included in the model were measured and no previous height observations were made; and

2. Calibrated response pattern: prediction of tree height in stands in which a subsample of tree heights, apart from DBH and/or stand variables, are available.

The fixed-effects approach represents the pattern of predicted height representing the mean behaviour of variation in height for both the given DBH and associated stand characteristics. Therefore, the second method of prediction was based on use of model (3.13) but with the random parameters set to zero and thus involved use of fixed parameters estimated using the mixed-effects modeling approach; that is PROC NLMIXED. The predictions were done using the expression:

$$\hat{\mathbf{H}}_{i\_fix} = 1.30 + \hat{\beta}_1(1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i))^{\hat{\beta}_3} \tag{3.26}$$

where, $\hat{\mathbf{H}}_{i\_fix}$ is a vector of predicted heights for the ith plot; $\hat{\beta}_1$, $\hat{\beta}_2$ and $\hat{\beta}_3$ are the estimated model parameters using NLME modelling approach and $\mathbf{DBH}_i$ is a vector of the observed DBH for trees in the ith plot.

Prediction of height is different, however, in the case of calibrated response pattern. In this case, model parameters are localized first by using the predicted values of the random parameters for each plot. Thus, the third method of prediction of tree height was based on model (3.13) with both fixed-effects and random-effects parameters included. The expression for prediction took the form (Huang *et al.*, 2009a; 2009b):

$$\hat{\mathbf{H}}_i = \hat{\mathbf{H}}_{i\_fix} + \hat{\mathbf{Z}}_i \hat{\mathbf{b}}_i. \tag{3.27}$$

The derivatives for model function (3.13) with respect to the two random parameters took the form:

$$der_{b_1} = \frac{\partial f(\mathbf{DBH}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})}{\partial b_1} = (1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i))^{\hat{\beta}_3};$$

$$der_{b_3} = \frac{\partial f(\mathbf{DBH}_i, \hat{\boldsymbol{\beta}}, \mathbf{0})}{\partial b_3} = \hat{\beta}_1 \log(1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i))(1 - \exp(-\hat{\beta}_2 \mathbf{DBH}_i))^{\hat{\beta}_3}.$$

Calculations of the partial derivatives constituted the design matrix $\mathbf{Z}_i$ as defined in (3.17). Thus;

$$\mathbf{Z}_i = \begin{bmatrix} der_{b_{11}} & der_{b_{31}} \\ \vdots & \vdots \\ der_{b_{1n_i}} & der_{b_{3n_i}} \end{bmatrix}. \tag{3.28}$$

$\hat{\mathbf{b}}_i$ was estimated using expression (3.24).

Different researches have proposed different approaches for tree measurements to be used for calibration of a mixed-effects model (e.g Calama and Montero, 2004; Dorado *et al.*, 2006; Sharma and Parton, 2007; Huang *et al.*, 2009a; VanderSchaaf, 2013; Carrol-Rivas *et al.*, 2014). However, no attempt was made in this study to determine the appropriate number and/or approach of prior tree height measurements to be used for calibration. It is important to note however that in most forest inventories, a subsample of trees is usually measured for tree heights from each plot; and that also can be used to calibrate a mixed-effects model.

3.6 Goodness-of-Fit Measures

Determination of the accuracy of model predictions and comparison of the two models (3.1) and (3.13) was based on the following statistics (Huang *et al.*, 2009a; Juntunen, 2010; Stankova and Diéguez-Aranda, 2013; Huang *et al.*, 2013; Corral-Rivas *et al.*, 2014; Missanjo and Mwale, 2014):

1. the bias (ϵ), which reflects the deviation of the model with respect to the observed values;

$$\epsilon = H - \hat{H} \tag{3.29}$$

and the overall mean bias given by the expression:

$$\bar{\epsilon} = \frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}\epsilon_{ij}.$$

The errors (residuals) associated with predictions using (3.25) were calculated as follows:

$$\epsilon_{ij} = H_{ij} - \hat{H}_{ij};$$

while the errors(residuals) associated with predictions using (3.26) and (3.27) were calculated as:

$$\boldsymbol{\epsilon}_{i\_fix} = \mathbf{H}_i - \hat{\mathbf{H}}_{i\_fix}$$

and

$$\boldsymbol{\epsilon}_i = \mathbf{H}_i - \hat{\mathbf{H}}_i;$$

with the overall mean bias for the ith plot given by the expression:

$$\bar{\epsilon}_i = \frac{\sum_{j=1}^{n_i}(H_{ij} - \hat{H}_{ij})}{n_i}.$$

2. Percent mean bias ((Bias%) or ($\bar{\epsilon}$%)), calculated as:

$$\bar{\epsilon}\% = \frac{\bar{\epsilon}}{\bar{H}} \times 100 \tag{3.30}$$

where, $\bar{H}$ is the mean height of the observed/measured trees.

3. Standard Deviation (StdDev) of the errors given by:

$$StdDev = \sqrt{\frac{1}{N-1}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(\epsilon_{ij} - \bar{\epsilon})^2}. \tag{3.31}$$

4. The root mean square error (RMSE), which analyses the precision of the estimates:

$$RMSE = \sqrt{\frac{1}{N-p}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(H_{ij} - \hat{H}_{ij})^2} \tag{3.32}$$

   on model fitting data set (where p is the number of fixed parameters in the model) while on model testing data set was calculated as:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{m}\sum_{j=1}^{n_i}(H_{ij} - \hat{H}_{ij})^2}. \tag{3.33}$$

5. Coefficient of determination, (R^2), which is a measure of variability explained by the model and given by the expression:

$$R^2 = 1 - \left[\frac{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(H_{ij} - \hat{H}_{ij})^2}{\sum_{i=1}^{m}\sum_{j=1}^{n_i}(H_{ij} - \bar{H})^2}\right]. \tag{3.34}$$

6. Overall model Prediction Accuracy(δ) given by:

$$\delta = \bar{\epsilon}^2 + StdDev^2. \tag{3.35}$$

7. Akaikes information criterion (AIC) given by:

$$AIC = -2\log(l) + 2P. \tag{3.36}$$

8. Bayesian information criterion (BIC) given by:

$$BIC = -2\log(l) + P\log(m) \tag{3.37}$$

where, for the (AIC) and (BIC), l is the maximized value of the likelihood function of the model; P is the total number of effective parameters to be estimated in the model (for the case of mixed-effects model, it includes fixed parameters, variance-covariance components of the random parameters, plus the residual variance component); m is the number of subjects (plots) in the study; and log is the natural logarithm. For the case of base model fitted using ONLS, the $\log(m)$ for BIC used was $\log(N)$; where N is the total number of observations (trees) in model fit data.

The goodness-of-fit measures outlined by the expressions given were applied to the entire population. An examination was however carried out for the frequency distributions of plot-specific biases from the base and mixed-effects models. This was done because overall mean biases averaged over all the observations (trees) from all plots could be misleading because the positive and negative residuals (errors) from individual plots could cancel one another out. Calculations were thus also obtained on plot-by-plot basis by considering n_i observations from the ith plot for the mean bias, Bias% and $RMSE$ in order to compare the two models on plot-by-plot basis. The plots were the Bias$\%_i$ exceeded $\pm 2.5\%$ of the observed mean were identified and listed separately. The reason for this identification was based on dividing the one-sided 5% significance level commonly used in statistical inferences into two sides (Huang *et al.*, 2009a; 2009b).

Chapter 4

Data Analysis, Results and Discussion

4.1 Exploratory Data Analysis

The summary statistics of model fitting and testing data sets used in the study are shown in Table 4.1 while Figures 4.1 and 4.2 show the scatter plot of tree height versus DBH for the model fit and test data sets respectively used in the study.

Table 4.1: Summary Statistics of Model Fitting and Testing Data Sets

Data Set	Plots	No. of Trees	DBH(cm)				H(m)			
			Min.	Max.	Mean	StdDev	Min.	Max.	Mean	StdDev
Model Fit	95	3,882	2.54	68.65	20.45	9.67	1.83	32.00	14.23	5.47
Model Test	36	287	2.54	58.93	16.31	7.95	3.03	31.39	13.28	5.00

Min. = minimum; Max. = maximum; DBH=Diameter at Breast Height in centimeters; H=Height in meters; StdDev=Standard Deviation.

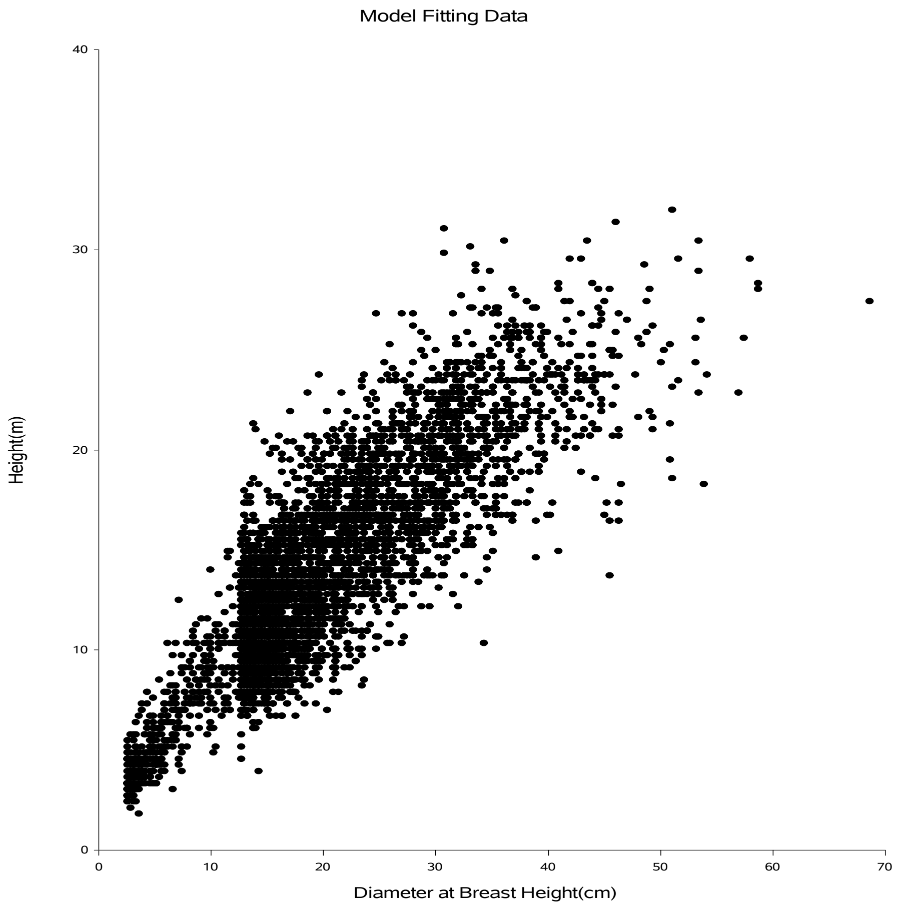

Figure 4.1: Scatter Plot of Height against Diameter at Breast Height for Model Fitting Data

Figure 4.2 shows a scatter plot for model test data set.

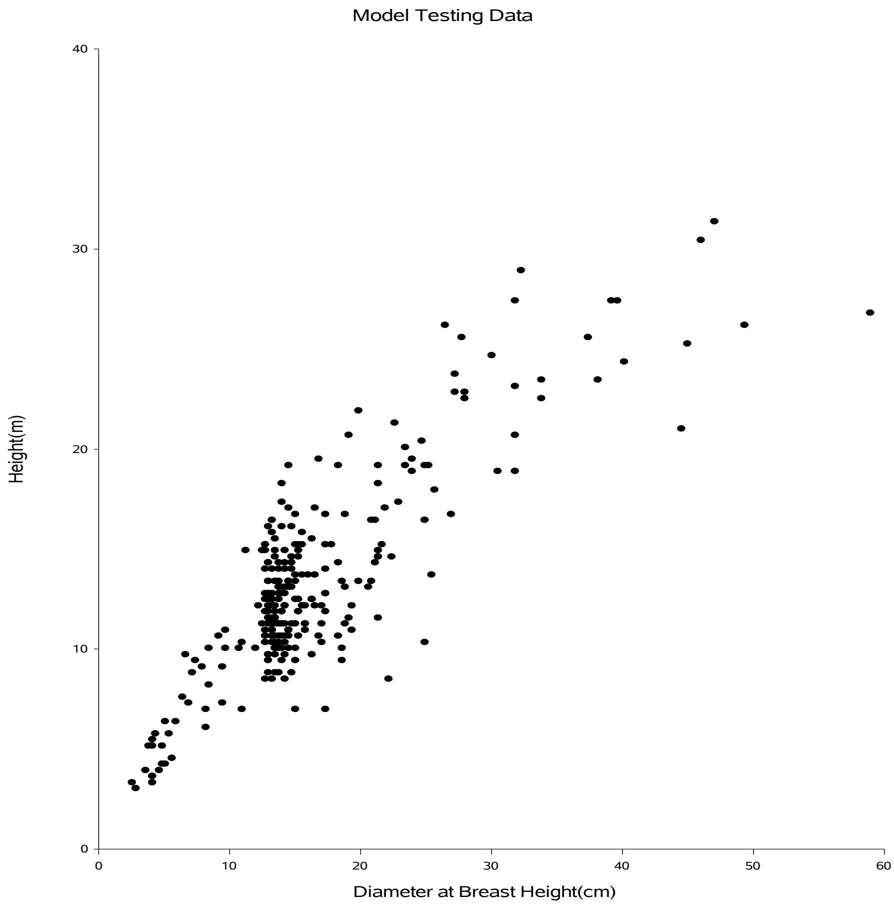

Figure 4.2: Scatter Plot of Height against Diameter at Breast Height for Model Testing Data

The plots of residuals (m) versus predicted tree heights (m) for both models, on model fitting data set, did not show obvious pattern of unequal error variance. Figures 4.3 and 4.4, show that there is no obvious violation of the assumption that the residuals (errors) are independent and identically distributed (iid) with mean 0 and a constant variance for both base and mixed-effects models.

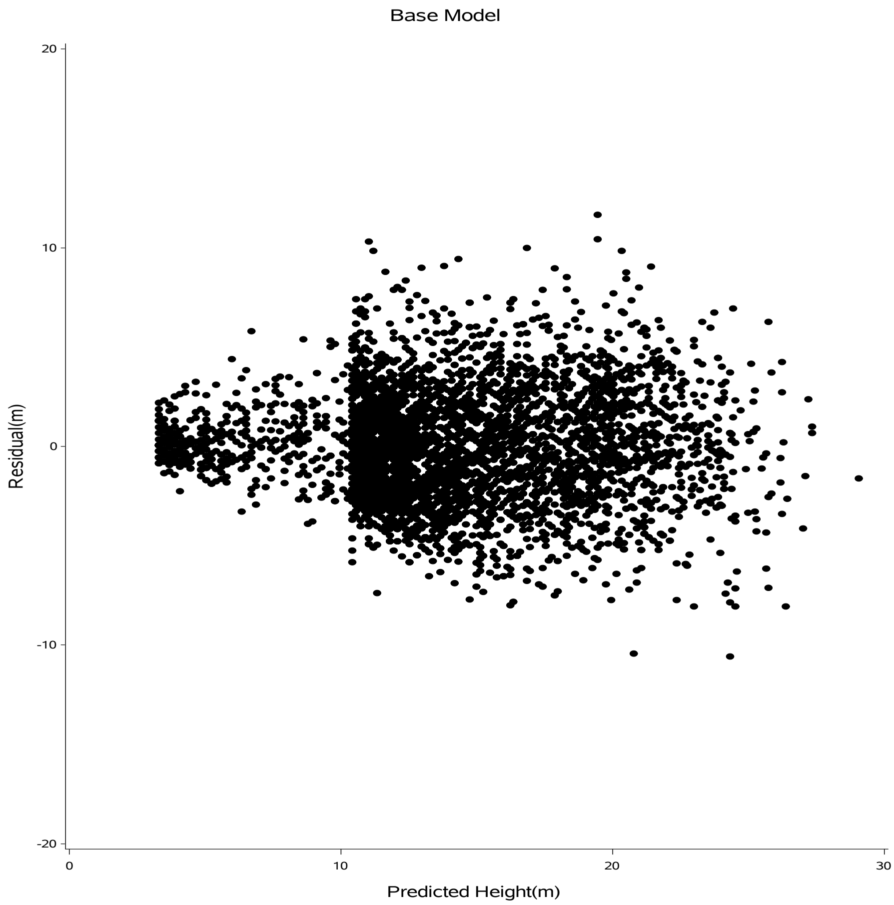

Figure 4.3: Plot of Residuals against Predicted Tree Heights for Base Model

Figure 4.4 shows the residual plot for the mixed-effects model.

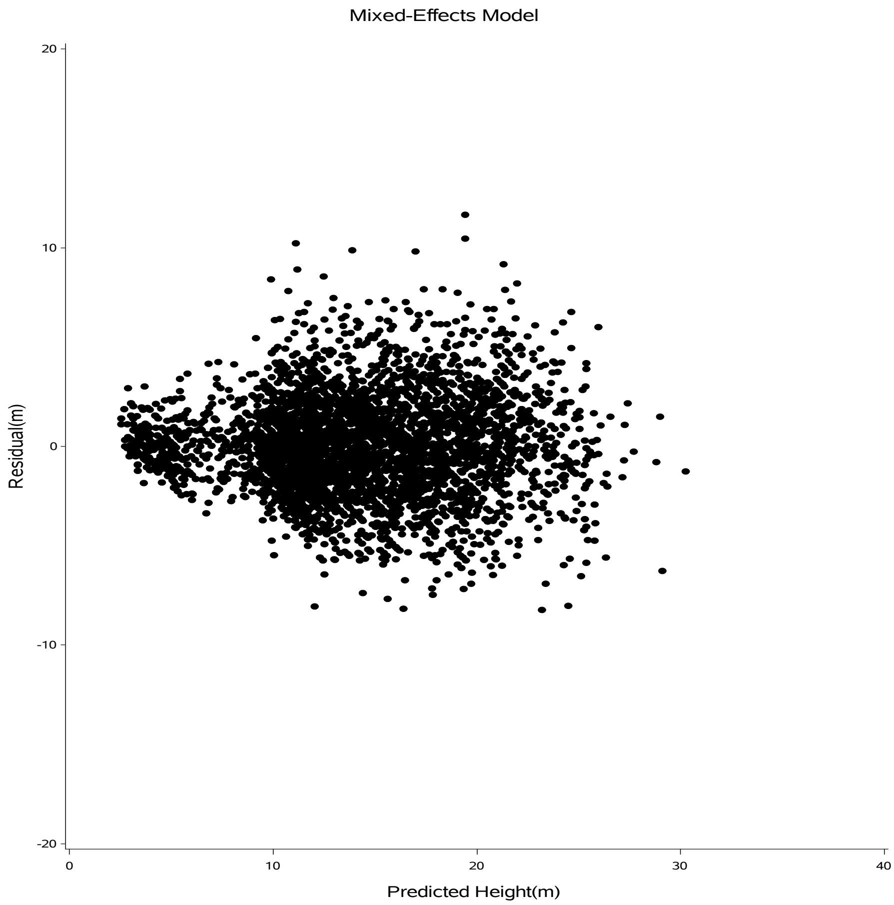

Figure 4.4: Plot of Residuals against Predicted Tree Heights for Mixed-Effects Model

4.2 Estimated Parameters and Fitting Statistics

The results obtained for parameter estimates with their associated standard errors and other statistics using SAS Proc NLIN and Proc NLMIXED procedures on model fitting data are shown in Table 4.2.

Table 4.2: Estimates of Fixed Parameters, Variance Components and their Standard Errors

Parameter	Base Model			Mixed-Effects Model			
	Estimate	Approx Std Error	Approx 95% CL	Estimate	Std Error	t-Value	Pr>$\|t\|$
β_1	33.5497	1.7981	30.0244 - 37.0244	29.7177	1.3930	21.33	<.0001
β_2	0.0259	0.0030	0.0200 - 0.0318	0.0310	0.0031	9.86	<.0001
β_3	1.0255	0.0411	0.0945 - 1.1061	1.0465	0.0474	22.08	<.0001
$\sigma^2_{b_1}$				16.2457	4.5723	3.55	0.0006
$\sigma^2_{b_3}$				0.0326	0.0086	3.18	0.0002
$\sigma_{b_1 b_3}$				0.5732	0.1664	3.44	0.0009
σ^2_{ε}				6.1796	0.1436	43.04	<.0001

All the estimated parameters for the two models were statistically significant at $\alpha = 0.05$ level as shown in Table 4.2. Furthermore, the coefficients of the two models are biologically logical; for example, β_1 represents the upper asymptotic height and was close, for both models, to the maximum height shown in Table 4.1, on model fitting data set. Based on the estimated parameters shown in Table 4.2, the two models fitted to the data can now be expressed as:

$$\hat{H}_{ij} = 1.30 + 33.5497(1 - \exp(-0.0259DBH_{ij}))^{1.0255} \tag{4.1}$$

for the base model; with the mixed-effects model without random parameters expressed as:

$$\hat{\mathbf{H}}_{i\_fix} = 1.30 + 29.7177(1 - \exp(-0.0310\mathbf{DBH}_i))^{1.0465} \tag{4.2}$$

while the mixed-effects model incorporating random effects parameters expressed as:

$$\hat{\mathbf{H}}_i = 1.30 + (29.7177 + b_{1i})(1 - \exp(-0.0310\mathbf{DBH}_i))^{(1.0465+b_{3i})}. \tag{4.3}$$

The variance-covariance matrices for among-plot and within-plot variability are given by:

$$\hat{\mathbf{D}} = \begin{bmatrix} 16.2457 & 0.5732 \\ 0.5732 & 0.0326 \end{bmatrix}; \tag{4.4}$$

$$\hat{\mathbf{R}}_i = 6.1796 \times \mathbf{I}_{n_i}. \quad (4.5)$$

Table 4.3 shows a summary of the fit statistics for the two models on model fit data set.

Table 4.3: Fit Statistics for Base and Mixed-Effects Models on Model Fit Data

Model	Fitting Method	Fit Statistics			
		R^2	$RMSE$	AIC	BIC
3.1	NLIN	0.74	2.77	18,932.2	18,957.3
3.13	NLMIXED	0.80	2.45	18,376.0	18,394.0

Table 4.3 which shows a summary of goodness-of-fit statistics indicates that model (3.13) had better results than model (3.1);R^2 was 0.80 and $RMSE$ was 2.45 for the mixed-effects model compared to 0.74 and 2.77 for R^2 and $RMSE$ respectively, for the base model. The results imply that R^2 increased by 8.12%[((0.80 − 0.74)/0.74) × 100] while $RMSE$ reduced by 11.55%[((2.77 − 2.45)/2.77) × 100] when mixed-effects modelling approach was used to fit the same data fitted using ONLS. Similarly, both AIC and BIC reduced when the same data was fitted using mixed-effects modelling approach. The increase in R^2, which measures the percentage of variation in the data explained by the model, and reduction in $RMSE$ indicates that mixed-effects modelling had a positive effect in terms of fitting the data and thus was the better model compared to the base model. Similarly, in model selection using information criterion, a model with both smaller AIC and BIC is preferred.

4.3 Predictive Performance of the Models

In order to assess the predictive performance of the two models in real-world applications, height predictions from the base model and the mixed-effects model were made based on the model testing data set (as well as model fitting data set for comparison). The summary statistics for the overall height predictive performance of the two models on both data sets are shown in Table 4.4.

Table 4.4: Summary Statistics for the Predictive Performance of the Base and Mixed-Effects Models

Data Set	Model	N	Bias(ϵ)				Accuracy(δ)	Bias%($\bar{\epsilon}$%)
			Min.	Max.	Mean	StdDev		
Fitting	3.1	3,882	-10.585	11.651	0.005	2.769	7.669	0.036
	3.13	3,882	-8.243	11.672	0.026	2.448	5.992	0.184
Testing	4.1	287	-7.085	8.932	1.128	2.540	7.721	8.494
	4.3	287	-5.688	6.077	0.362	1.916	3.802	2.725

The results as shown in Table 4.4 suggest that both models, on the average, under-predicted the height for both data sets (that is, the mean biases are all positive). The mixed-effects model (3.13) produced the larger mean bias (0.026) but lower standard deviation of the errors (2.448) on model fitting data. It (3.13) however produced an overall better predictive accuracy (5.992) than the base model (7.669). The results for model testing data reveal that the mixed-effects model outperformed the base model on all goodness-of-fit measures related to model predictive accuracy; model (4.3) had both lower mean bias (0.362) and standard deviation of the errors (1.916) and consequently a higher predictive accuracy (3.802) as compared to the base model(4.1) with mean bias of 1.128 and standard deviation of errors of 2.540 resulting in predictive accuracy of 7.721. The results suggest that the mixed-effects model was twice better in terms of predictive accuracy on model testing data set. The last column of Table 4.4 shows that the base model, on average, under-predicted tree heights by 0.036% while the mixed-effects model under-predicted by 0.184% on model fit data. The difference for under-prediction is obvious on model test data set where the base model, on the average, under-predicted by 8.494% as compared to 2.725% for the mixed-effects model.

Results in Table 4.4 are in agreement with the fit statistics in Table 4.3 which shows that the mixed-effects model had an overall better fit to the data as reflected by a higher R^2 and lower $RMSE$.

From the results in Tables 4.3 and 4.4, we can conclude that plot-by-plot variability was significant and that it was taken into account when mixed-effects modeling approach was applied. The base model fitted using ONLS however could not account for plot-by-plot variability; variability among plots was included in the error term and thus a higher $RMSE$ and lower R^2 recorded when compared to the mixed-effects model. This explains the reason for a better predictive accuracy for the mixed-effects model. It also confirms the fact that a model fitted to the data using ONLS provides population average estimate of height for given DBH observations; regardless of whether trees are measured in different plots, the parameter estimates are assumed to be correct or the same. It is common knowledge that trees, for example in a 5 years plot would not be expected to have the same maximum height as trees in a 20 years plot, all other things being equal. Thus, one would expect, for example β_1 representing the upper asymptotic height, in the Chapman-Richards model to be different for the two plots. A model fitted using ONLS however does not take that into account unless different parameter estimates are obtained for each plot! Since the mixed-effects model as constructed in this study has a random component to β_1, it means that such a difference would be taken into account and the H-D model can be quickly localized without having to fit separate models for each plot.

4.4 Prediction of Tree Height for an Example Plot

To demonstrate the prediction of tree height using equations (4.1), (4.2) and (4.3), one of the plots in model test data sets is hereby presented:

Table 4.5: Height Prediction for an Example Plot

M	Observed		Predicted						δ_i	$\bar{\epsilon}_i\%$
	DBH_{ij}	H_{ij}	$\hat{H}_{ij}$	ϵ_{ij}	der_{b_1}	der_{b_3}	$\hat{\mathbf{H}}_i$	$\boldsymbol{\epsilon}_i$		
4.1	37.34	25.60	21.84	3.7574					**32.0217**	**20.6210**
	46.99	31.39	24.71	6.6850						
	31.75	27.43	19.83	7.5979						
	2.54	3.35	3.29	0.0588						
	Mean	**21.94**	**17.42**	**4.5248**						
	StdDev			**3.3983**						
4.3	37.34	25.60			0.6742	-7.5480	25.35	0.2482	**7.9468**	**8.8354**
	46.99	31.39			0.7579	-5.9654	28.73	2.6584		
	31.75	27.43			0.6132	-8.5167	22.92	4.5119		
	2.54	3.35			0.0672	-5.1536	3.01	0.3363		
	Mean						**20.00**	**1.9387**		
	StdDev							**2.0465**		

In Table 4.5, M=Model; δ_i=Model Accuracy for the ith plot; $\bar{\epsilon}_i\%$=Bias% for the ith plot.

Predictions for tree heights using the base model were obtained using equation (4.1). The derivatives at the given DBHs, der_{b_1} and der_{b_3} , shown in Table 4.5 were calculated using equations outlined for (3.28) . They constitute the $\mathbf{Z}_i$ matrix defined in (3.28). Thus;

$$\hat{\mathbf{Z}}_i = \begin{bmatrix} 0.6742 & -7.5480 \\ 0.7579 & -5.9654 \\ 0.6132 & -8.5167 \\ 0.0672 & -5.1536 \end{bmatrix}$$

The among-plots and within-plots variability as given in (4.4) and (4.5)are given by:

$$\hat{\mathbf{D}} = \begin{bmatrix} 16.2457 & 0.5732 \\ 0.5732 & 0.0326 \end{bmatrix}$$

$$\hat{\mathbf{R}}_i = \begin{bmatrix} 6.1796 & 0 & 0 & 0 \\ 0 & 6.1796 & 0 & 0 \\ 0 & 0 & 6.1796 & 0 \\ 0 & 0 & 0 & 6.1796 \end{bmatrix}$$

With these matrices and the observed heights given in Table 4.5, the random parameters, $\hat{\mathbf{b}}_i$, are predicted using equation (3.24):

$$\hat{\mathbf{b}}_i \approx \hat{\mathbf{D}}\hat{\mathbf{Z}}_i'(\hat{\mathbf{Z}}_i\hat{\mathbf{D}}\hat{\mathbf{Z}}_i' + \hat{\mathbf{R}}_i)^{-1}(\mathbf{H}_i - \hat{\mathbf{H}}_{i\_fix})$$

$$\hat{\mathbf{b}}_i = \begin{bmatrix} 7.6893128 & 0.1555327 \end{bmatrix}'$$

with $\hat{\mathbf{H}}_{i\_fix}$ predicted using (4.2) and $\mathbf{H}_i$ are the observed tree heights as given in Table 4.5 under the third column.

Height predictions using the mixed-effects model (4.3) were obtained in a similar manner for all the other plots. The Bias% and model prediction accuracy shown in Table 4.5 were calculated using expressions (3.30) and (3.35) respectively. As can be seen in Table 4.5, it is clear that the mixed-effects model had a better predictive accuracy than the base model for the example plot. The mixed-effects model had both smaller Bias% (8.8354%) and accuracy (7.9468), which is desirable, than the base model (with Bias% = 20.6210% and accuracy = 32.0217). Furthermore, the predicted mean height (20.00m) using the mixed-effects model was much closer to the mean height observed (21.94m)for the example plot; the predicted mean height (17.42m) using base model was not as close. Summary results as shown for the example plot; that is Bias% and accuracy on a plot-by-plot basis were calculated for all the plots on both model fit and test data sets. The results for model test data are as shown in appendix 1.

4.5 Distribution of Biases for the Base and Mixed-Effects Models

It is important to note that comparison of the two models as shown in Table 4.4 could be misleading at times. This is because the biases or errors could cancel each other out when calculations are made or averaged over all observations. Thus, a plot-by-plot analysis often reveals more information about the performance of models being compared. Furthermore, the main interest of mixed-effects modeling lies in plot-by-plot performance. Therefore, summary statistics for the distribution of biases (errors) from all the 95 plots for model data and 36 plots for model test data are shown in Table 4.6.The summary in Table 4.6 is based on identified plots that had Bias$\%_i$ exceeding $\pm 2.5\%$ of the observed mean. The reason for this identification was based on dividing the one-sided 5% significance level commonly used in statistical inferences into two sides (Huang *et al* ., 2009a; 2009b).

Table 4.6: Summary of Distribution of Plot-by-Plot Biases for Model Fit and Test Data Sets

Model	Data	Variable	Model Fit Data						Model Test Data					
			Freq.	Min.	Max.	Mean	StdDev	δ	Freq.	Min.	Max.	Mean	StdDev	δ
4.1	All	$\bar{\epsilon}_i$	95	-3.233	2.700	-0.008	1.180	1.39	36	-2.126	6.168	1.243	1.960	5.39
		$\bar{\epsilon}_i\%$	95	-32.490	17.432	-1.739	10.124		36	-27.940	29.439	6.264	12.799	
		$RMSE_i$	95	0.803	6.035	2.862	0.918		36	0.786	6.329	2.583	1.450	
	$-2.5 \leq \bar{\epsilon}_i\% \geq 2.5$	$\bar{\epsilon}_i$	25	-0.394	0.346	0.027	0.220	0.05	6	-0.223	0.197	0.060	0.180	0.04
		$\bar{\epsilon}_i\%$	25	-2.231	2.450	0.201	1.514		6	-2.043	1.867	0.481	1.630	
		$RMSE_i$	25	0.803	5.652	2.682	0.975		6	0.786	1.789	1.283	0.413	
4.3	All	$\bar{\epsilon}_i$	95	-0.989	0.735	-0.006	0.270	0.07	36	-1.030	3.345	0.523	0.936	1.15
		$\bar{\epsilon}_i\%$	95	-12.052	6.529	-0.509	2.923		36	-16.243	15.964	2.411	6.444	
		$RMSE_i$	95	0.718	5.614	2.439	0.848		36	0.766	3.473	1.880	0.798	
	$-2.5 \leq \bar{\epsilon}_i\% \geq 2.5$	$\bar{\epsilon}_i$	75	-0.280	0.456	0.060	0.132	0.02	14	-0.291	0.244	-0.007	0.166	0.03
		$\bar{\epsilon}_i\%$	75	-2.233	2.235	0.349	0.865		14	-2.460	2.109	-0.141	1.435	
		$RMSE_i$	75	0.738	5.614	2.529	0.805		14	0.766	2.038	1.340	0.400	

Table 4.6 shows that the contrast in prediction accuracy between the base and mixed-effects models is obvious on both model fit and test data sets. The base model produced larger biases on more plots on both data sets than the mixed-effects model. The base model had only 25 plots out of 95 (for model data) and 6 plots out of 36(for test data) with biases not exceeding ±2.5% of the observed mean where as the mixed-effects model had 75 plots out of 95(for model data) and 14 plots(for test data) with biases not exceeding ±2.5% of the observed mean. The distribution of the biases is further illustrated in bar charts as shown in Figures 4.5 and 4.6.

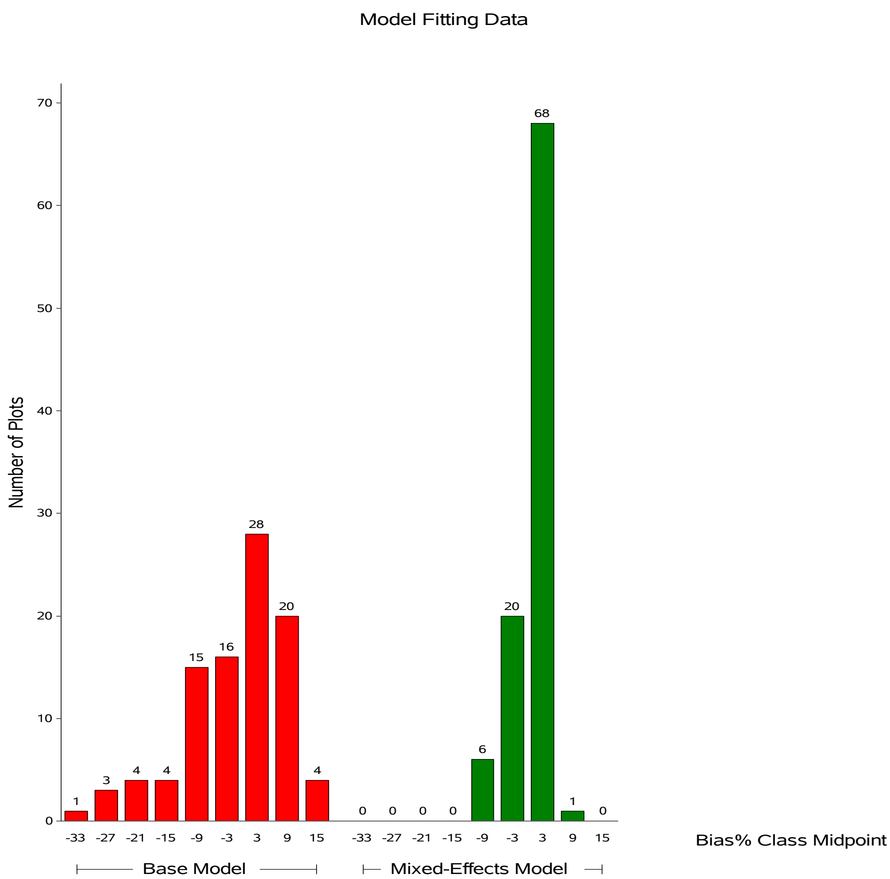

Figure 4.5: Frequency Distribution of Bias % for Model Fit Data

Figure 4.5 shows clearly that 68 plots out of 95 had bias class midpoint of 3% for the mixed-effects model compared to only 28 plots out of 95 for the base model. The results show a similar trend on the test data as can be seen on Figure 4.6.

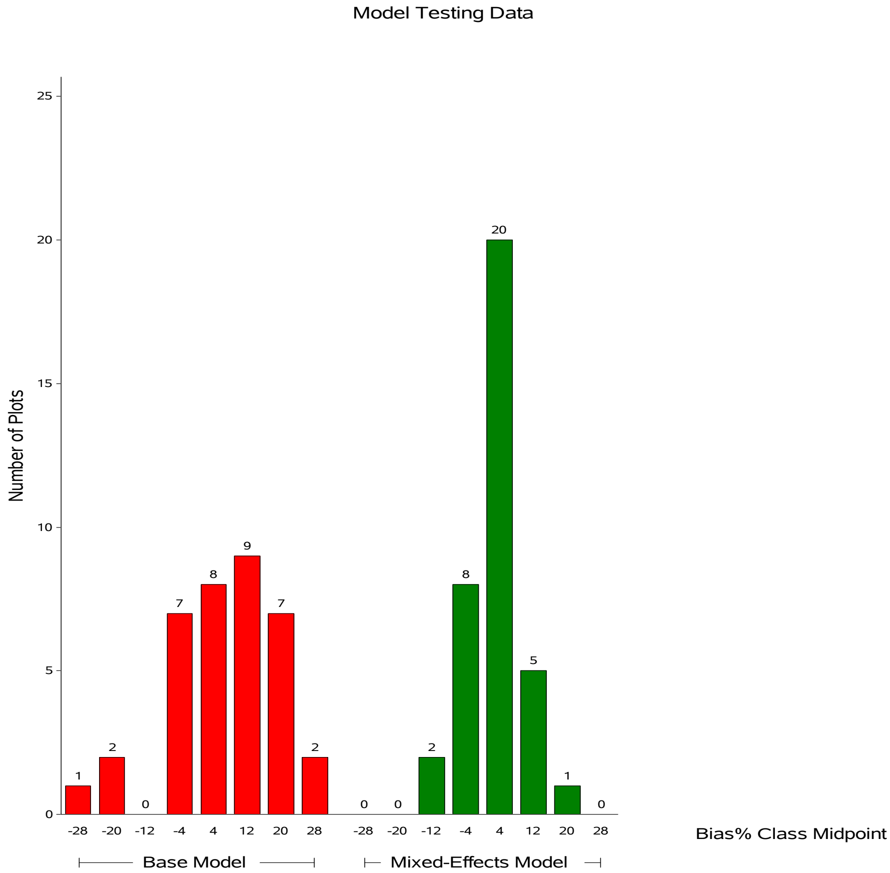

Figure 4.6: Frequency Distribution of Bias% for Model Test Data

Figure 4.6 shows clearly that more than half of biases for the mixed-effects model were centred on 4% as opposed to being spread out as can be seen for the base model. Table 4.6 and Figures 4.5 and 4.6 clearly demonstrate the superiority of the mixed-effects model over the base model when the two modes are compared on a plot-by-plot basis. The summary of fit statistics and predictive performance as shown in Tables 4.3 and 4.4 respectively, based on all observations, do not show the superiority of the mixed-effects model as revealed in Table 4.6 and Figures 4.5 and 4.6. This means that comparison of models (where mixed-effects modeling is applied) just based on all observations could be misleading and as such it is important to do a plot-by-plot comparison as was observed and recommended by (Huang *et al.*, 2009a; 2009b). A complete analysis of plot-by-plot calculation of Bias%, mean bias, standard deviation of errors and model prediction accuracy for model test data is as shown in Appendix 1.

4.6 Discussion

The results obtained in this research project are similar in many respects to those obtained by other researchers. For example, Corral-Rivas *et al* (2014) observed that a mixed-effects model had better fit statistics ($R^2 = 0.85; RMSE = 2.21; BIC = 6461$) than the base model fitted using ONLS ($R^2 = 0.73; RMSE = 2.95; BIC = 8749$). They however did not compare the two models on plot-by-plot basis as has been done in the present study. The findings in this study are also in agreement with those of Huang *et al* (2009a; 2009b) who observed that plot-by-plot comparison of models was necessary when considering mixed-effects models. In particular, Huang *et al* (2009b) recommended that frequency distribution of biases should be routinely examined in any future study involving nonlinear mixed-effects modeling approach. Their recommendation was utilized in this study and the findings are that frequency distribution of biases from individual plots are indeed more powerful in revealing the superiority of a given mixed-effects model than just relying on summary statistics averaged over all the observations. Huang *et al* (2013) also compared the predictive performance of base models with their counterpart mixed-effects models. Their results are similar to the present study in that they also concluded that mixed-effects models performed better than base models on both model and application data sets. Furthermore, Huang *et al* (2013) in their study defined the thresholds (absolute) for percent mean bias ($\bar{\epsilon}\%$) as 0.5% on model fitting data and 10% for the independent application data, respectively, for a given model to be considered generally acceptable, provided that it makes biological sense. The results obtained in this study are within these limits for both the base and mixed-effects models on both model fit and test data sets as shown in Table 4.4. Paulo *et al* (2011) also obtained similar results of better fit statistics for the nonlinear mixed effects model (NLMEM) when compared to nonlinear fixed effects model (NLFEM) approaches. They concluded that even in situations where the NLMEM calibration was not possible, the model should be preferred. The findings in this study are however in contrast to recommendations by Paulo *et al* (2011) but in agreement with De-Miguel *et al* (2013) who concluded that fixed-effects models (or base model for the case of this study) should be preferred in the absence of calibration data. This is because the base model had better fit statistics and predictive accuracy on both model fit and test data sets when compared to a mixed-effects model which did not incorporate random effects; predictions based on expressions (4.1) and (4.2). The comparison of predictive performance between base and mixed-effects (with fixed-effects only) models is presented in Appendix 2.

Chapter 5

Conclusion and Recommendations

5.1 Conclusion

The Chapman-Richards model was considered in this study to demonstrate the development and utilization of a mixed-effects model in prediction of tree height using only DBH as the predictor variable. The predictive performance of the constructed mixed-effects model was compared to the base model (fitted using ONLS) on model test data as well as model fit data set. The comparison was done based on statistics that included Bias, Bias%, $RMSE$, standard deviation of the errors, overall model accuracy, R^2,AIC and BIC. Based on these statistics, the mixed-effects model performed better than the base model on both data sets. An examination was also done on frequency distribution of the biases for individual plots and the results were more revealing on the superiority of the mixed-effects model to the base model. Mixed-effects modelling approach was therefore found to be more appropriate, for model development and use of H-D models for predicting tree height, than the commonly used ONLS.

5.2 Recommendations

1. Comparison of H-D models, where mixed-effects modelling is applied, should also be done on plot-by-plot basis and not just based on statistics averaged over all the observations.

2. Forest modelers and managers in Africa should consider using mixed-effects modelling technique in the development and utilization of H-D models.

Bibliography

[1] Adame, P., del RÍo, M., and Cañllas, I. (2008). A mixed nonlinear height-diameter model for pyrenean oak (*Quercus pyrenaica* Willd). *Forest Ecology and Management* **256**: 88 - 98.

[2] Ahmadi, K., Alavi, S.J., Kouchaksaraei, M.T., and Aertsen, W. (2013). Non-linear height-diameter models for oriental beech (*Fagus orientalis* Lipsky) in the Hyrcanian Forests, Iran. *Biotechnology, Agronomy, Society and Environment* **17**(3): 431 - 440.

[3] Aigbe, H.I and Oyebade, B. (2012). Relationship between height and stump diameter for *Terminalia ivorensis*(A.Chev) in Sokponba Forest Reserve, Edo State, Nigeria. *Journal of Agriculture and Social Research* **12**(1): 57 - 63.

[4] Bates, D.M., and Watts, D.G.(1988). *Nonlinear Regression Analysis and its Application*, 2nd edn, John Wiley and Sons Inc, New York. pp 1 - 86.

[5] Budhathoki, C.B, Lynch, T and Guldin, J. (2008). A mixed-effects model for the dbh-height relationship of Shortleaf Pine (*Pinus echinata* Mill.). *Southern Journal of Applied Forestry* **32**(1) .

[6] Calama, R., and Montero, G. (2004). Interregional nonlinear height-diameter model with random Coefficients for stone pine in Spain. *Canadian Journal of Forest Research* **34**:150 - 163.

[7] Corral-Rivas, S., Álvarez-González J.G., Crecente-Campo, F., and Corral-Rivas, J.(2014). Local and generalized height-diameter models with random parameters for

mixed, uneven-aged forests in Northwestern Durango, Mexico. *Forest Ecosystems* **1**:6.

[8] De-Miguel, S., Guzmn, G and Pukkala, T. (2013). A comparison of fixed and mixed-effects modeling and yield prediction of an indeginous neotropical species (*Centrolobium tomentosum*) in a plantation system. *Forest Ecology and Management* **291**(2013): 249 - 258.

[9] Dorado, F.C, Diéguez-Aranda, U., Anta, M., RodrÍguez and Gadow, K. (2006). A generalized height-diameter model including random components for radiate pine plantations in northwestern Spain.*Forest Ecology and Management* **229**(2006): 202 - 213.

[10] Eerikäinen, K., (2003). Predicting the height-diameter pattern of planted *Pinus kesiya* stands in Zambia and Zimbabwe. *Forest Ecology and Management* **175**: 355 - 366.

[11] Gavin, H. P. (2013). The Levenberg-Marquardt method for nonlinear least squares curve-fitting problems. pp 1 - 17.

[12] Greene, W. H. (2012). *Econometric Analysis*, 7th edn, International Edition, Pearson Education Limited, London . pp 221 - 255.

[13] Huang, S., Titus, S., J and Wiens, D. P. (1992). Comparison of nonlinear height-diameter functions for major Alberta tree species. *Canadian Journal of Forest Research* **22** : 1297 - 1304.

[14] Huang, S., Wiens, D.P, Yang, Y., Meng, S.X, and VanderSchaaf C.L. (2009a). Assessing the impacts of species composition, top height and density on individual tree height prediction of quacking aspen in boreal mixedwoods. *Forest Ecology and Management* **258** : 1235 - 1247.

[15] Huang, S., Meng, S.X and Yang, Y. (2009b). Using nonlinear mixed model technique to determine the optimal tree height prediction model for black spruce. *Modern Applied Science* **3**(4): 3 - 18.

[16] Huang, S., Yang, Y, and Aitkin, D.(2013). Population and plot-specific individual tree height-diameter models for major Alberta tree species. Technical Report Pub. No.: T/600(2013). 81pp.

[17] Juntunen, M. (2010). *Modeling tree and stand characteristics of Lodgepole pine (Pinus contorta) plantations in Iceland*. University of Eastern Finland, Faculty of Science and Forestry, Masters thesis in forest planning and economics. 68 pp.

[18] Krisnawati, H., Wang Y., and Lui L. (2010). Generalized height-diameter models for *Acacia mangium* Wild. plantations in Sumatra. *Journal of Forestry Research* **7**(1):1 - 19.

[19] Kroon, J., Andersson, B. and Mullin, T. J. (2008). Genetic variation in the diameterheight relationship in Scots pine (*Pinus sylvestris*). *Canadian Journal of Forest Research*, **38**(6): 1493 - 1503, 10.1139/X07-233.

[20] Littell, R. C., George A. M., Walter W.S., Russell D. W., and Oliver, S. (2006). *SAS®for Mixed Models*, 2nd edn. Cary, NC:SAS Institute Inc.

[21] Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* , **11**(2): 431 - 441.

[22] Misir, N. (2010). Generalized height-diameter models for *Populus tremula* L.Stands. *African Journal of Biotechnology* **9**(28): 4348 - 4355.

[23] Missanjo, E. and Mwale, G.(2014). A mixed-effects height-diameter model for*Pinus kesiya* in Malawi. *Journal of Biodiversity Management and Forestry* **3**(2).

[24] Obeyed, M.H. (2014). Predictive models between diameter, height, crown diameter and age of *Pinus brutia* ten. in Zawita and Atrush Districts. *Global Journal of Bioscience and Technology* **3**(2):203 - 210.

[25] O'Connell, B.M., LaPoint, E, B., Turner, J.A., Ridley, T., Boyer, D., Wilson, A, M., Waddell, K.L., Pugh, S.A., and Conkling, B.L. (2013). The Forest Inventory

and Analysis Database: Database Description and Users Manual Version 5.1.6 for Phase 2. U.S Department of Agriculture, Forest Service.

[26] Osman, E.M.H., Idris, E.Z.A., and Ibrahim, E.M.M. (2013). Height-diameter prediction models for some utilitarian natural tree species. *Journal of Forest Products and Industries* **2**(2): 31 - 39.

[27] Oyebade, B.A, Popo-ola, F.S and Alex, A. (2012). Height-diameters for eight Mid-western rainforest species in Nigeria using Monseruds model. *ARPN Journal of Science and Technology* **2** (5): 479 - 486.

[28] Paulo, J.A, Tomé, J and Tomé, M. (2011). Nonlinear fixed and random generalized height-diameter models for Portuguese cork oak stands. *Annals of Forest Science* **68**: 295 - 309.

[29] Peng, C., Zhang, L.,and Liu, J. (2001). Developing and validating nonlinear height-diameter models for major tree species of Ontarios boreal forests. *Northern Journal of Applied Forestry* **18**:87 - 94.

[30] Petráš, R., Bošeľa, M., Mecko, J.,Oszlányi, J., and Popa, I. (2014). Height-diameter models for mixed-species forests consisting of spruce, fir, and beech.*Folia Forestalia Polonica* , Series A **56**(2): 93 - 104.

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

[32] Pinheiro, J.C. and Bates, D.M. (2000). *Mixed-Effects Models in S and S-Plus.* Spring-Verlag, New York, 537 pp.

[33] Rawlings, J. O, Pantula, S. G and Dickey, D. A. (1998). *Applied Regression Analysis: A Research Tool* , 2nd edn. Springer Texts in Statistics. pp 485 - 511.

[34] SAS Institute Inc. (2013). The NLIN Procedure. *SAS/STAT®13.1 Users Guide.*Cary, NC: SAS Institute Inc. pp 5606 - 5709.

[35] SAS Institute Inc. (2014). The NLMIXED Procedure. *SAS/STAT®13.2 Users Guide* .Cary, NC: SAS Institute Inc. pp 5682 - 5777.

[36] Saunders, M.R., and Wagner, R.G. (2008). Heightdiameter models with random coefficients and site variables for tree species of central Maine.*Annals of Forest Science* **65**: 203 - 213.

[37] Sharma, R.P. (2009). Modeling height-diameter relationships for Chir pine trees. *Banko Janakari* **19**(2): 3 - 9.

[38] Sharma, M., and Zhang, S.Y., (2004). Height-diameter models using stand characteristics for *Pinus banksiana* and *Picea mariana*. *Scandinavian Journal of Forest Research* **19**:442 - 451.

[39] Sharma, M. and Parton, J. (2007). Height-diameter equations for boreal tree species in Ontario using a mixed modeling approach.*Forest Ecology and Management* **249** : 187 - 198.

[40] Stankova, T.V, and Diéguez-Aranda, U. (2013). Height-diameter relationships for Scots pine plantations in Bulgaria: optimal combination of model type and application.*Annals of Forest Research* **56**(1): 149 - 163.

[41] Temesgen, H., and Gadow, K.V. (2004). Generalized heightdiameter modelsan application for major tree species in complex stands of interior British Colombia. *European Journal of Forest Research* **123**:45 - 51.

[42] US Forest Service. (2015). htt://apps.fs.fed.us/fiadb-downloads/datamart.htm, version 5.0. [Accessed 20th April 2015].

[43] US Forest Service. (2011). Forestry Inventory and Analysis: National Core Field Guide Version 5.1: Volume I: Field Data Collection Procedures for Phase 2 Plots. US Department of Agriculture.

[44] Vanclay, J. K. (2009). Tree diameter, height and stocking in even-aged forests.*Annals of Forest Science* **66**(2009) 702.

[45] VanderSchaaf, C.L. (2014). Mixed-effects height-diameter models for ten conifers in the inland Northwest, USA. *Southern Forests* **76**(1): 1 - 9.

[46] Vargas-Larreta B, Castedo-Dorado F, Álvarez-González JG, Barrio-Anta M, and Cruz-Cobos F .(2009). A generalized height diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico). *Forestry* **82**(4): 445 - 462.

[47] Xu, H., Sun, Y., Wang, X., and Li, Y. (2014). Height-diameter models of Chinese fir (*Cunninghamia lanceolata*) based on nonlinear mixed-effects models in Southeast China. *Advanced Journal of Food Science and Technology* **6**(4):445 - 452.

[48] Zeide, B and VanderSchaaf, C . (2002). The effects of density on the height-diameter relationship. In Outcalt, Kenneth W., ed. 2002. Proceedings of the eleventh biennial southern silvicultural research conference Gen. Tech. Rep. SRS-48. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station. 622 pp.

[49] Zhang, L. (1997). Cross-validation of nonlinear growth functions for modeling tree heightdiameter relationships.*Annals of Botany* **79**:251 - 257.

APPENDIX 1: Plot-by-Plot Calculated Mean Bias, Bias%, Standard Deviation, RMSE and Accuracy for Model Test Data Set

	BASE MODEL					MIXED-EFFECTS MODEL				
Plot ID	Mean Bias	Bias%	RMSE	StdDev	Accuracy	Mean Bias	Bias%	RMSE	StdDev	Accuracy
1	0.614	5.556	1.334	1.233	1.896	0.136	1.231	1.183	1.223	1.514
2	-0.223	-2.043	1.654	1.795	3.272	-0.132	-1.212	1.635	1.785	3.204
3	-1.771	-27.940	3.271	3.075	12.594	-1.030	-16.243	2.621	2.694	8.320
4	-0.109	-1.079	1.789	1.996	3.996	-0.108	-1.070	1.784	1.990	3.973
5	2.343	17.672	2.890	1.854	8.927	0.977	7.370	1.954	1.853	4.389
6	-0.275	-2.931	1.289	1.336	1.861	-0.109	-1.157	1.249	1.320	1.755
7	1.089	8.933	1.936	2.263	6.309	0.703	5.765	1.750	2.266	5.630
8	-0.775	-7.754	1.257	1.027	1.655	-0.209	-2.090	0.988	1.002	1.048
9	2.274	18.650	3.321	2.794	12.978	1.174	9.631	2.679	2.780	9.108
10	0.143	1.228	0.996	1.138	1.315	0.038	0.325	0.981	1.132	1.282
11	1.731	14.256	1.856	0.732	3.532	0.733	6.039	0.983	0.718	1.053
12	0.192	1.688	0.786	0.823	0.714	0.047	0.410	0.766	0.826	0.685
13	2.652	19.956	2.950	1.361	8.886	0.808	6.080	1.521	1.358	2.497
14	4.525	20.621	5.398	3.398	32.022	1.939	8.835	2.627	2.047	7.947
15	0.884	6.855	2.194	2.200	5.622	0.359	2.782	1.989	2.143	4.722
16	2.563	13.560	2.632	0.846	7.284	1.848	9.780	1.917	0.717	3.930
17	-0.721	-6.721	3.465	3.712	14.301	-0.350	-3.264	3.364	3.665	13.554
18	-1.108	-17.316	1.108	0.001	1.228	-0.869	-13.586	0.870	0.011	0.756
19	3.589	20.419	4.719	3.200	23.120	0.973	5.534	3.090	3.063	10.328
20	0.158	1.223	1.487	1.620	2.648	0.061	0.472	1.404	1.537	2.366
21	1.577	12.617	3.073	2.849	10.605	0.659	5.271	2.633	2.753	8.013
22	1.578	12.493	1.862	1.067	3.628	0.602	4.766	1.154	1.063	1.493
23	-0.957	-8.092	1.513	1.229	2.428	-0.291	-2.460	1.178	1.197	1.518
24	-2.126	-19.648	2.959	2.200	9.359	-0.790	-7.296	2.027	1.996	4.608
25	0.955	7.794	1.961	1.742	3.945	0.132	1.075	1.612	1.633	2.684
26	1.334	7.293	1.561	0.889	2.569	0.719	3.931	1.126	0.949	1.417
27	6.168	29.439	6.233	1.037	39.122	3.345	15.964	3.473	1.079	12.352
28	1.803	13.071	2.305	1.536	5.609	0.694	5.035	1.611	1.553	2.895
29	0.512	4.425	1.304	1.340	2.059	0.244	2.109	1.045	1.136	1.350
30	6.081	26.251	6.329	1.922	40.676	2.327	10.045	2.874	1.847	8.827
31	1.736	11.360	2.883	2.461	9.070	0.648	4.241	2.354	2.420	6.275
32	1.779	11.990	2.833	2.240	8.185	0.229	1.543	2.038	2.057	4.285
33	-0.456	-4.046	2.083	2.142	4.798	-0.178	-1.578	1.937	2.033	4.165
34	3.139	17.909	3.979	3.459	21.814	2.250	12.840	3.198	3.213	15.389
35	3.663	15.963	4.796	3.344	24.597	1.201	5.234	3.104	3.092	11.003
36	0.197	1.867	0.987	1.059	1.161	0.045	0.428	0.958	1.048	1.100

APPENDIX 2: Predictive Performance of Base and Mixed-Effects (with Fixed-Effects only) Models

Table A1 Summary of Predictive Performance of the Base and Mixed-Effects (with Fixed-Effects only) Models

Data Set	Model	N	Bias (ϵ)				Accuracy	Bias%
			Min.	Max.	Mean	StdDev	(δ)	($\bar{\epsilon}$ %)
Fitting	4.1	3882	-10.585	11.651	0.005	2.769	7.669	0.036
	4.2	3882	-10.024	11.933	0.086	2.778	7.727	0.607
Testing	4.1	287	-7.085	8.932	1.128	2.540	7.721	8.494
	4.2	287	-7.046	9.260	1.122	2.574	7.883	8.453

Table A2 Height Prediction for an Example Plot

Model	Observed		Predicted				Bias%	Accuracy
	DBH	H	$\widehat{H}_{ij}$	ϵ_{ij}	$\widehat{\boldsymbol{H}}_{i\_fix}$	$\boldsymbol{\epsilon}_{i\_fix}$	($\bar{\epsilon_i}$%)	(δ)
(4.1)	37.34	25.60	21.84	3.75738			**20.6210**	**32.0217**
	46.99	31.39	24.71	6.68504				
	31.75	27.43	19.83	7.59787				
	2.54	3.35	3.29	0.05880				
	mean	**21.94**	**17.42**	**4.52477**				
	StdDev			**3.39826**				
(4.2)	37.34	25.60			21.34	4.2650	**22.5493**	**37.8319**
	46.99	31.39			23.82	7.5661		
	31.75	27.43			19.52	7.9082		
	2.54	3.35			3.30	0.0523		
	mean				**17.00**	**4.9479**		
	StdDev					**3.6538**		

Table A3 Summary of Plot-Based Height Prediction Biases from the Base and Mixed-Effects (with Fixed-Effects only) Models

Model	Data	Variable	Model Fit Data						Model Test Data					
			Freq.	Min.	Max.	Mean	StdDev	δ	Freq.	Min.	Max.	Mean	StdDev	δ
4.1	All	$\bar{\epsilon}_i$	95	-3.233	2.700	-0.008	1.180	1.39	36	-2.126	6.168	1.243	1.960	5.39
		$\bar{\epsilon}_i\%$	95	-32.490	17.432	-1.739	10.124		36	-27.940	29.439	6.264	12.799	
		$RMSE_i$	95	0.803	6.035	2.862	0.918		36	0.786	6.329	2.583	1.450	
	$-2.5 \leq \bar{\epsilon}_i\% \geq 2.5$	$\bar{\epsilon}_i$	25	-0.394	0.346	0.027	0.220	0.05	6	-0.223	0.197	0.060	0.180	0.04
		$\bar{\epsilon}_i\%$	25	-2.231	2.450	0.201	1.514		6	-2.043	1.867	0.481	1.630	
		$RMSE_i$	25	0.803	5.652	2.682	0.975		6	0.786	1.789	1.283	0.413	
4.2	All	$\bar{\epsilon}_i$	95	-2.997	2.936	0.073	1.226	1.51	36	-2.151	6.234	1.251	2.036	5.71
		$\bar{\epsilon}_i\%$	95	-32.639	17.313	-1.349	10.439		36	-28.344	29.755	6.128	13.136	
		$RMSE_i$	95	0.882	5.273	2.606	0.800		36	0.774	6.481	2.613	1.503	
	$-2.5 \leq \bar{\epsilon}_i\% \geq 2.5$	$\bar{\epsilon}_i$	20	-0.332	0.343	0.018	0.207	0.04	5	-0.176	0.119	0.050	0.128	0.02
		$\bar{\epsilon}_i\%$	20	-2.193	2.482	0.065	1.492		5	-1.739	1.129	0.387	1.204	
		$RMSE_i$	20	1.113	3.565	2.327	0.661		5	0.774	1.787	1.195	0.415	

Freq. =frequency

Table A4 Plot-by-Plot Comparison of Base and Mixed-Effects (with Fixed-Effects only) Models

	BASE MODEL					MIXED-EFFECTS MODEL (with Fixed-Effects only)				
Plot ID	Mean Bias	Bias%	RMSE	StdDev	Accuracy	Mean Bias	Bias%	RMSE	StdDev	Accuracy
1	0.614	5.556	1.334	1.233	1.896	0.541	4.903	1.299	1.229	1.803
2	-0.223	-2.043	1.654	1.795	3.272	-0.299	-2.738	1.660	1.789	3.290
3	-1.771	-27.940	3.271	3.075	12.594	-1.797	-28.344	3.262	3.043	12.492
4	-0.109	-1.079	1.789	1.996	3.996	-0.176	-1.739	1.787	1.989	3.986
5	2.343	17.672	2.890	1.854	8.927	2.265	17.081	2.827	1.854	8.566
6	-0.275	-2.931	1.289	1.336	1.861	-0.349	-3.722	1.310	1.340	1.917
7	1.089	8.933	1.936	2.263	6.309	1.011	8.295	1.895	2.267	6.160
8	-0.775	-7.754	1.257	1.027	1.655	-0.844	-8.450	1.297	1.022	1.757
9	2.274	18.650	3.321	2.794	12.978	2.196	18.011	3.265	2.790	12.609
10	0.143	1.228	0.996	1.138	1.315	0.071	0.607	0.984	1.133	1.289
11	1.731	14.256	1.856	0.732	3.532	1.659	13.663	1.787	0.727	3.281
12	0.192	1.688	0.786	0.823	0.714	0.115	1.016	0.774	0.827	0.697
13	2.652	19.956	2.950	1.361	8.886	2.573	19.361	2.878	1.360	8.470
14	4.525	20.621	5.398	3.398	32.022	4.948	22.549	5.873	3.654	37.832
15	0.884	6.855	2.194	2.200	5.622	0.828	6.420	2.154	2.179	5.432
16	2.563	13.560	2.632	0.846	7.284	2.720	14.391	2.746	0.532	7.681
17	-0.721	-6.721	3.465	3.712	14.301	-0.792	-7.385	3.470	3.700	14.320
18	-1.108	-17.316	1.108	0.001	1.228	-1.163	-18.175	1.164	0.039	1.355
19	3.589	20.419	4.719	3.200	23.120	3.624	20.618	4.725	3.166	23.157
20	0.158	1.223	1.487	1.620	2.648	0.119	0.921	1.455	1.589	2.538
21	1.577	12.617	3.073	2.849	10.605	1.546	12.371	3.086	2.885	10.715
22	1.578	12.493	1.862	1.067	3.628	1.506	11.926	1.800	1.064	3.400
23	-0.957	-8.092	1.513	1.229	2.428	-0.996	-8.416	1.529	1.217	2.473
24	-2.126	-19.648	2.959	2.200	9.359	-2.151	-19.876	2.945	2.151	9.252
25	0.955	7.794	1.961	1.742	3.945	0.912	7.446	1.923	1.721	3.795
26	1.334	7.293	1.561	0.889	2.569	1.695	9.268	1.898	0.936	3.750
27	6.168	29.439	6.233	1.037	39.122	6.234	29.755	6.289	0.957	39.785
28	1.803	13.071	2.305	1.536	5.609	1.780	12.903	2.319	1.590	5.694
29	0.512	4.425	1.304	1.340	2.059	0.566	4.888	1.383	1.411	2.312
30	6.081	26.251	6.329	1.922	40.676	6.221	26.856	6.481	1.989	42.661
31	1.736	11.360	2.883	2.461	9.070	1.744	11.412	2.865	2.431	8.949
32	1.779	11.990	2.833	2.240	8.185	1.790	12.064	2.789	2.173	7.927
33	-0.456	-4.046	2.083	2.142	4.798	-0.518	-4.596	2.073	2.115	4.744
34	3.139	17.909	3.979	3.459	21.814	3.165	18.058	3.939	3.318	21.022
35	3.663	15.963	4.796	3.344	24.597	4.160	18.130	5.167	3.310	28.262
36	0.197	1.867	0.987	1.059	1.161	0.119	1.129	0.974	1.059	1.136