Environ Monit Assess (2016) 188: 245
DOI 10.1007/s10661-016-5252-0

Strategies for the use of mixed-effects models in continuous forest inventories

James A. Westfall

Received: 20 August 2015 / Accepted: 16 March 2016 / Published online: 24 March 2016

Abstract Forest inventory data often consists of measurements taken on field plots as well as values predicted from statistical models, e.g., tree biomass. Many of these models only include fixed-effects parameters either because at the time the models were established, mixed-effects model theory had not yet been thoroughly developed or the use of mixed models was deemed unnecessary or too complex. Over the last two decades, considerable research has been conducted on the use of mixed models in forestry, such that mixed models and their applications are generally well understood. However, most of these assessments have focused on static validation data, and mixed model applications in the context of continuous forest inventories have not been evaluated. In comparison to fixed-effects models, the results of this study showed that mixed models can provide considerable reductions in prediction bias and variance for the population and also for subpopulations therein. However, the random effects resulting from the initial model fit deteriorated rapidly over time, such that some field data is needed to effectively recalibrate the random effects for each inventory cycle. Thus, implementation of mixed models requires ongoing maintenance to reap the benefits of improved predictive behavior. Forest inventory managers must determine if this gain in predictive power outweighs the additional effort needed to employ mixed models in a temporal framework.

J. A. Westfall (✉)
U.S. Forest Service, Northern Research Station, Newtown Square, PA, USA
e-mail: jameswestfall@fs.fed.us

Keywords Forest monitoring · Subsampling · Tree height model · Random parameter · Prediction bias

Introduction

Statistical models are often used in forestry applications for attributes that are difficult and/or costly to measure directly, e.g., tree biomass. For large-area forest inventories, models are often developed and employed at regional or subregional scales for efficiency (Woodall et al. 2011). Estimates of population parameters derived from the predicted values of these models can be unbiased if the model predictions are unbiased for the population of interest. However, these models are susceptible to prediction bias when applied to smaller areas (Westfall 2015). Thus, there is a need to localize prediction accuracy such that some assurance is obtained that subpopulation estimates are unbiased. However, fitting models for application at small spatial scales is often problematic due to associated small sample sizes; for large-area forest inventory applications, the number of models needed to cover the landscape would likely be substantial. Thus, methods for local calibration of models developed for large areas are of particular relevance to this issue.

Various techniques have been suggested to obtain better local prediction accuracy including the following: (1) using latitude, longitude, and/or elevation as predictor variables (Bechtold 2003; Westfall 2006); (2) geographically weighted regression (Zhang and Shi 2004); and (3) mixed-effects models (Trincado and Burkhart

2006; de-Miguel et al. 2014). Due to the readily available spatial information provided by global positioning system (GPS) and global navigation satellite system (GNSS), the direct use of spatial predictors may be the most appealing solution operationally. However, unless the attribute of interest varies systematically with these locational covariates, their effectiveness in localizing predictions can be poor (Westfall 2015). Geographically weighted models produce numerous vectors of parameter estimates, each associated with a particular geographic point, e.g., individual trees. Further, the models may need to be refitted over time to account for changes specific to the local environment of an individual tree (Zhang and Gove 2005). Mixed models can provide a local calibration to individual observations via random-effects (REs) model parameters that modify the fixed parameter values (Jayaraman and Zakrzewski 2001). Similar to the geographic weighting methods, the efficacy of the REs over time is unknown. An advantage of the mixed-model approach is that a mean response model can be fitted over a relatively large spatial domain to describe general relationships between the response and predictor variables, which can then be used as the basis for as long as those underlying relationships are considered to hold. Localized predictions would then arise via the influence of RE parameters. For trees in the model fitting data that are present in subsequent inventory cycles, it is possible to use the existing REs. However, prediction of random effects (REs) is necessary for any new trees appearing in later cycles. Methods for obtaining these new REs include subsampling to determine local allometric relationships (Temesgen et al. 2008) or utilizing existing REs from neighboring trees.

Numerous studies have reported the improved predictive ability of mixed models and the appeal of generating REs for new observations; however, these evaluations are often presented in the context of independent validation data for a one-time assessment (Adame et al. 2008; Jiang and Li 2010). Practical implementation of such models requires further evaluation in a continuous forest inventory context where a time component must also be considered. In this study, tree height prediction models for red maple (*Acer rubrum*) trees in PA, USA, were used to explore issues related to implementation of a mixed-model framework to improve local prediction accuracy for large-area forest inventories. Specifically, the objectives of the analysis are to (1) determine the magnitude and statistical significance of reduction in model prediction bias and residual variance for subpopulations, (2) quantify the effects on model prediction accuracy resulting from transferring existing REs to subsequent inventory cycles, and (3) evaluate the performance of two strategies for calculating REs needed for new observations.

Methods

Data

This study focused on live red maple (*A. rubrum*) trees measured between 2000 and 2013 in the state of PA, USA, by the forest inventory and analysis (FIA) program of the U.S. Forest Service. All data were collected using the annualized FIA sample and plot design, where the sampling intensity is approximately one plot for every 2400 ha of area (Reams et al. 2005). Each sample plot contains four 7.3-m (24 ft) radius subplots (Bechtold and Scott 2005). All trees within forested areas of the plot having a diameter at breast height (dbh) of 12.7 cm (5.0 in.) and larger were measured, and variables used in this study included dbh, total tree height, crown ratio and crown class (a measure of social position) (U.S. Forest Service 2007). Each tree is also associated with a plot condition; conditions delineate different forest stands within the plot based on attributes such as forest type, stand size, and tree density. In most cases, a plot will be entirely within a single stand; thus stand- and plot-level attributes are often the same. The data were divided into three parts, each corresponding to a 5-year inventory cycle in which all FIA plots in the state were measured (approximately 20 % of the plots are measured each year). The cycles were as follows: (1) 2000–2004, (2) 2005–2009, and (3) 2010–2013 (incomplete cycle, missing 2014 data). Data summary statistics are provided in Table 1.

Analysis

Statistical models relating tree heights to tree-size attributes (e.g., dbh) have been extensively reported in the forestry research literature over the last several decades (Gregoire 2012) due to the importance of these relationships in many aspects of forestry (Feldpausch et al. 2011). More recently, a number of studies have examined these allometric relationships using mixed-model techniques (Castedo-Dorado et al. 2006; Vargas-Larreta et al. 2009; Coble and Lee 2011). The issues outlined

Table 1 Summary statistics for red maple tree attributes (H = height (m); DIA = diameter at breast height, dbh (cm); CR = crown ratio (%); CC = crown class indicator) measured across three inventory cycles in PA, USA

Cycle[a]	Number	Mean			Std. dev.			Min.			Max.			No. trees	
		H	DIA	CR	H	DIA	CR	H	DIA	CR	H	DIA	CR	CC = 0	CC = 1
1	15,460	17.6	22.5	30.8	4.9	8.8	11.6	3.0	12.7	1	37.5	108.2	95	9940	5501
2	14,298	18.0	23.2	32.6	4.9	9.3	11.9	3.0	12.7	1	37.8	93.7	95	8350	5933
3	8820	18.7	23.9	35.0	5.0	9.8	12.2	2.1	12.7	0	36.0	92.7	90	5237	3578

[a] 1 = 2000–2004; 2 = 2005–2009; 3 = 2010–2013 (incomplete cycle; missing 2014)

above were investigated via analysis of a simplified version of the tree height model from Westfall and Laustsen (2006) fitted separately to data from each inventory cycle *l*:

$$H_{ijkl} = \left(\beta_{0l} + \beta_{1l}CC_{ijkl} + \beta_{2l}CR_{ijkl}\right) \times \left(1-\exp\left(-\beta_{3l}DBH_{ijkl}\right)\right)^{\beta_{4l}} + \varepsilon_{ijkl} \quad (1)$$

where:

H_{ijkl} total height of tree *i* in stand *j* on plot *k* in cycle *l* (m)
DBH_{ijkl} diameter at 1.37 m of tree *i* in stand *j* on plot *k* in cycle *l* (cm)
CC_{ijkl} 1 if crown class of tree *i* in stand *j* on plot *k* in cycle *l* is dominant/codominant, = 0 otherwise
CR_{ijkl} crown ratio of tree *i* in stand *j* on plot *k* in cycle *l* (%)
β_{0l}-β_{4l} fixed-effects parameters estimated from the data in cycle *l*
ε_{ijkl} random residual error of tree *i* in stand *j* on plot *k* in cycle *l* (m)

Two mixed-model specifications of [1] were evaluated. The first was designed to maximize prediction accuracy for each tree by incorporating tree-level REs:

$$H_{ijkl} = \left(\beta_{0l} + \beta_{1l}CC_{ijkl} + \left(\beta_{2l} + \theta_{1ijkl}\right)CR_{ijkl}\right) \times \left(1-\exp\left(\left(-\beta_{3l} + \theta_{2ijkl}\right)DBH_{ijkl}\right)\right)^{\beta_{4l}} + \varepsilon_{ijkl} \quad (2)$$

where:

θ_{hijkl} predicted random effect parameter for tree *i* in stand *j* on plot *k* in cycle *l* ($h = 1, 2$); $\theta_{hijkl} \sim N(0, \sigma^2_{hl})$.

A second formulation was also developed to evaluate the potential use of stand-level REs, i.e., a single value for each RE would apply to all trees in the stand:

$$H_{ijkl} = \left(\beta_{0l} + \beta_{1l}CC_{ijkl} + \left(\beta_{2l} + \theta_{1jkl}\right)CR_{ijkl}\right) \times \left(1-\exp\left(\left(-\beta_{3l} + \theta_{2jkl}\right)DBH_{ijkl}\right)\right)^{\beta_{4l}} + \varepsilon_{ijkl} \quad (3)$$

The placement of the REs was contingent upon σ^2_{hl} being statistically different from zero. As with model [1], models [2] and [3] were fitted independently to data from each cycle *l*.

It should be noted here that the intent of specifying the random-effects parameters in [2] and [3] is to improve the predictive accuracy of the models. Often, the use of mixed models is motivated by a need to account for correlated observations; however, this situation is not a requirement for the use of mixed models. In this specific case, there are no correlated observations at the tree level, i.e., model [2]. At the stand level, the mixed-model parameterization [3] could be considered as addressing correlations among trees within a stand, but the general rationale was to simply allow for comparisons of random-effect efficacy at the tree and stand levels.

Model fit was assessed via concordance correlation and mean squared error (MSE) statistics. Vonesh et al. (1996) describe the concordance correlation as:

$$r_c = 1 - \frac{\sum_{ijkl}\left(H_{ijkl} - \hat{H}_{ijkl}\right)^2}{\sum_{ijkl}\left(H_{ijkl} - \overline{H}\right)^2 + \sum_{ijkl}\left(\hat{H}_{ijkl} - \widehat{\overline{H}}\right)\left(H_{ijkl} - \overline{H}\right) + n\left(\overline{H} - \widehat{\overline{H}}\right)^2} \quad (4)$$

where $\hat{H}_{ijkl}$ is the model prediction, $\widehat{\overline{H}}$ is the mean model prediction, $\overline{H}$ is the mean observed height, and *n* is the number of observations. The r_c statistic spans the interval between −1 and +1, with $r_c = 1$ indicating a perfect fit to the data and $r_c \leq 0$ suggesting considerable lack of fit. MSE is estimated during the iterative model fitting

process, which is too protracted for presentation in this paper; however, the interpretation is consistent with other types of models in that it represents the variance of the residuals when the model is correctly specified.

The ability of mixed models to ameliorate prediction biases for subpopulations within the study area was assessed by examination of means and standard deviations of residuals by county.[1] Specifically, comparisons of mean residuals between fixed-effects model [1] and mixed-effects models [2] and [3] were made to determine the extent to which mixed models may be advantageous. The magnitude and statistical significance of bias reduction was evaluated via ratio estimates of mean absolute residuals:

$$\hat{\mu}_{c(\bullet)} = \sum_{ijkl \in c} \left| H_{ijkl} - \hat{H}_{ijkl} \right| / n_c \tag{5}$$

$$R_b = \frac{\sum_{c=1}^{m} \hat{\mu}_{c(ME)}/m}{\sum_{c=1}^{m} \hat{\mu}_{c(FE)}/m} = \frac{\widehat{\overline{\mu}}_{c(ME)}}{\widehat{\overline{\mu}}_{c(FE)}} \tag{6}$$

$$V(R_b) = \frac{1}{m\widehat{\overline{\mu}}^{2}_{c(FE)}} \left(V\left(\widehat{\overline{\mu}}_{c(ME)}\right) + R_b^2 V\left(\widehat{\overline{\mu}}_{c(FE)}\right) - 2R_b Cov\left(\widehat{\overline{\mu}}_{c(ME)}, \widehat{\overline{\mu}}_{c(FE)}\right) \right) \tag{7}$$

Where c indexes counties, n_c is the number of observations in county c, $\hat{\mu}_c$ $(\bullet)$ is the county-level mean absolute residual (from the mixed models [2] or [3] ($\hat{\mu}_c$ (ME)); from the fixed model [1] ($\hat{\mu}_c$ (FE)), R_b is the ratio of county-level mean biases for mixed-effects and fixed effects models, m is the number of counties, $V(\bullet)$ is the variance of the specified estimate, and $Cov(\bullet)$ is the covariance between the specified estimates. The result of primary interest is whether the estimated ratio is statistically smaller than one, which would indicate small area prediction bias is reduced via the mixed model. Table 2 provides summary statistics for numbers of trees at both the plot- and county-level aggregations.

Changes in variability of residuals between fixed-effects and mixed-effects models were also analyzed. Similar to the above, county-level standard deviations of residuals were analyzed in a ratio context:

$$R_s = \frac{\sum_{c=1}^{m} \hat{\sigma}_{c(ME)}/m}{\sum_{c=1}^{m} \hat{\sigma}_{c(FE)}/m} = \frac{\widehat{\overline{\sigma}}_{c(ME)}}{\widehat{\overline{\sigma}}_{c(FE)}} \tag{8}$$

$$V(R_s) = \frac{1}{m\widehat{\overline{\sigma}}^{2}_{c(FE)}} \left(V\left(\widehat{\overline{\sigma}}_{c(ME)}\right) + R_s^2 V\left(\widehat{\overline{\sigma}}_{c(FE)}\right) - 2R_s Cov\left(\widehat{\overline{\sigma}}_{c(ME)}, \widehat{\overline{\sigma}}_{c(FE)}\right) \right) \tag{9}$$

Where $\hat{\sigma}_c$ $(\bullet)$ is the county-level residual standard deviation (from the mixed models [2] or [3] ($\hat{\sigma}_c$ (ME)); from the fixed model [1] ($\hat{\sigma}_c$ (FE))) and R_s is the ratio of county-level mean standard deviations of residuals for mixed-effects and fixed effects models.

Impacts on predicted values by using both fixed- and mixed-effects models in subsequent cycles were also assessed, e.g., by applying the regression results from cycle 1 to make predictions for cycles 2 and 3. For example, predicted values of model [2] for cycle 2 using model parameters estimated from cycle 1 would be accomplished via:

$$\hat{H}_{ijk2(1)} = \left(\beta_{01} + \beta_{11} CC_{ijk2} + \left(\beta_{21} + \theta_{1ijk1}\right) CR_{ijk2}\right) \times \left(1 - \exp\left(\left(-\beta_{31} + \theta_{2ijk1}\right) DBH_{ijk2}\right)\right)^{\beta_{41}} \tag{10}$$

The key analytical statistics were:

$$\overline{\varepsilon}_{Hl(l')} = \sum_{i=1}^{n_{l(l')}} \left(H_{ijkl} - \hat{H}_{ijkl(l')} \right) / n_{l(l')} \tag{11}$$

$$Var(\varepsilon)_{Hl(l')} = \sum_{i=1}^{n_{l(l')}} \left(H_{ijkl} - \hat{H}_{ijkl(l')} \right)^2 / n_{l(l')} \tag{12}$$

$$RMSE_{Hl(l')} = \sqrt{Var(\varepsilon)_{Hl(l')} + \overline{\varepsilon}^{2}_{Hl(l')}} \tag{13}$$

where $\overline{\varepsilon}_{Hl}\left(l'\right)$ is the mean residual (difference between observed and predicted heights) calculated from estimated parameters from cycle (l') model fit applied to the cycle l data, $Var(\varepsilon)_{Hl}\left(l'\right)$ is the residual variance,

[1] Counties are administrative subdivisions of states having the primary purpose of providing local governance. The study area (Pennsylvania) has 67 counties ranging in area from 337 to 3182 km^2 (mean = 1730 km^2).

Table 2 Summary statistics for number of red maple trees by plot and county across three inventory cycles in PA, USA

	Plot				County			
Cycle	Mean	Std. dev.	Min.	Max.	Mean	Std. dev.	Min.	Max.
1	6.99	6.22	1	56	237.85	183.08	5	733
2	6.90	5.98	1	55	219.97	172.38	3	747
3	6.79	5.82	1	50	135.69	111.06	5	453

$n_l\left(l'\right)$ is number of trees common to both cycles l and (l'), and $RMSE_{Hl}\left(l'\right)$ is the root mean squared error (RMSE) for height predictions.

In practice, a mixed model would be developed using a specific data set and then applied in subsequent inventory cycles. Trees remaining in the inventory in subsequent cycles would already have predicted REs resulting from the initial regression analysis. However, new trees will also enter the inventory sample in subsequent cycles and values for REs will need to be generated. In this study, approximately 10 % of the trees in subsequent cycles were not previously measured. Two options to be evaluated are as follows: (1) adopting the RE from a previously existing tree(s) in the stand, and (2) direct empirical calculation using a subsample of trees in the current inventory (Trincado et al. 2007). To evaluate option 1 (using existing REs), assignment of the mean REs from up to seven of the nearest trees on the plot was investigated. In this case, nearest is defined as the shortest "distance" between tree attributes. For new trees appearing in cycle 2, this attribute distance would be:

$$Dist_{ijk2} = \left(DBH_{\bullet jk2} - DBH_{ijk2}\right)^2 + \left(CR_{\bullet jk2} - CR_{ijk2}\right)^2 + \left(CC_{\bullet jk2} - CC_{ijk2}\right)^2 \quad (14)$$

Where $Dist_{ijk2}$ is the attribute distance from tree i to the subject tree (denoted with a $\bullet$ subscript) and others as previously defined. During the analyses, it was noted that changes in RE were more correlated with changes in CR than the other predictors, and thus, CR largely drives the distance calculation. The REs from the selected number of trees with the closest distance were averaged to obtain the RE for the new tree. When the previously existing red maple trees on the plot were insufficient for the number of trees to be selected, the distances were computed based on all red maples in the county in order to assign REs to the subject tree. Generally, this process results in tree-level REs for all new trees as the selection of nearest trees with existing REs based on [14] changes for each new tree. All trees surviving from the initial inventory (cycle 1) retain their REs from the model fitting procedure.

For option 2, the new REs can be obtained via this equation from Vonesh and Chinchilli (1997):

$$\theta = DZ'\left(ZDZ' + R\right)^{-1}(y - Xb) \quad (15)$$

where θ = vector of predicted random-effects parameters, $Z = FB$ (B = regression design matrix for random-effects parameters, F = matrix of partial derivatives with respect to each fixed parameter evaluated at DBH_{ijk} and CR_{ijk} for each calibration tree), R = predicted variance/covariance matrix of residual errors, D = variance/covariance matrix of random effects, y = vector of observed tree heights, X = regression design matrix for fixed-effects parameters, and b = vector of fixed-effects parameters. This analysis examined the use of up to five trees to estimate the REs for new trees. This methodology produces a single estimate for each RE that would be applied to all new trees appearing in the stand (essentially stand-level REs for new trees). This approach differs from option 1 in that the data arise from a subsample of trees taken during the current inventory cycle.

Results

The regression analyses across all tree inventory cycles for model [1] produced statistically significant ($\alpha = 0.05$) estimates for all parameters, with ranges for mean squared error (σ_ε^2) and concordance correlation (r_c) statistics being (8.52–9.66) and (0.75–0.79) respectively. Regression analyses of models [2] and [3] showed similar results across the three inventory cycles

with all fixed-effects parameter estimates and random-effects variance estimates being statistically significant (Table 3). With the exception of cycle 1, σ_{ε}^2 was smaller for model [2], generally indicating that tree-level REs provide better predictions for individual trees than stand-level effects. Similarly, although the model fits exhibited r_c above 0.90, the correlations were higher for model [2]. Examination of residual plots did not reveal any indication of systematic model prediction misbehavior.

Due to the inclusion of random effects when making model predictions for models [2] and [3], prediction bias and residual variance in comparison to the mean response model [1] for smaller areas (counties) showed substantial reductions across all three inventory cycles (Table 4). With the exception of cycle 1, model [2] mean absolute residuals were reduced by approximately 60–70 %. The reduction for cycle 1 was smaller; however, reductions near 50 % were attained. Notable improvements were also found for model [3], where reductions were in the range of 35–45 %. In all cases, the ratios were statistically different from 1 with 95 % confidence. The statistics for model [2] were superior to those for model [3], again indicating that REs at the tree level are preferred in comparison to stand-level aggregation.

One of the primary motivations of this paper was to consider the use of mixed models in a continuous forest inventory context. One evaluation was to examine the use of regression results for models [2] and [3] in subsequent inventory cycles for trees remaining in the sample. The results in Table 5 suggest that this approach generally performed poorly, as mean residuals were commonly in the 0.5–1.0 m range. For both models [2] and [3], the variance of the residuals was slightly higher for model [3], particularly when applied to trees measured in cycle 3. The resultant RMSE statistics were approximately 2.6–2.9 m, with model [2] exhibiting slightly smaller values. The RMSE from the original regression based on cycle 1 data were 1.0–1.2 m for model [2] and 1.8–1.9 m for model [3]. Due to the superior performance of model [2] in all the results examined thus far, model [3] will no longer be considered for further analysis.

The effort to predict REs for new trees appearing in subsequent cycles by selecting a specified number of nearest trees (based on [14]) from cycle 1 and averaging their REs was marginally successful when compared to the alternative of using the mean response model [1]. When applying this method to new trees in cycle 2, the mean residual was reduced from −0.25 to ≈0, regardless of the number of trees selected from cycle 1 (Table 6). However, in the case where only the nearest single tree was used, the residual RMSE was larger (2.98 m) than found with the mean response model (2.94 m). As the

Table 3 Results of regression analysis of models [2] and [3] for three inventory cycles in PA, USA

Parameter	Tree-level random effects (model 2)						Stand-level random effects (model 3)					
	Cycle 1		Cycle 2		Cycle 3		Cycle 1		Cycle 2		Cycle 3	
	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)
β_0	36.8551	(0.9151)	38.1379	(0.8978)	40.9365	(1.1529)	28.1135	(0.4390)	30.9837	(0.6388)	33.3112	(0.8108)
β_1	2.6678	(0.1335)	2.5840	(0.1296)	2.2865	(0.1524)	3.7309	(0.1023)	3.8152	(0.1169)	3.8854	(0.1470)
β_2	−0.1534	(0.0062)	−0.1604	(0.0056)	−0.2025	(0.0075)	−0.0740	(0.0042)	−0.0855	(0.0046)	−0.0949	(0.0059)
β_3	0.0267	(0.0020)	0.0232	(0.0017)	0.0235	(0.0021)	0.0310	(0.0020)	0.0210	(0.0018)	0.0183	(0.0018)
β_4	0.7758	(0.0228)	0.6960	(0.0186)	0.6994	(0.0231)	0.6463	(0.0207)	0.5171	(0.0156)	0.4907	(0.0160)
σ_{ε}^2	5.0075	(0.3463)	3.6222	(0.3094)	2.9662	(0.3924)	4.4643	(0.0575)	4.2998	(0.0577)	3.9202	(0.0672)
σ_1^2	3.10E-05	(0.0000)	3.10E-05	(0.0000)	2.60E-05	(0.0000)	7.60E-05	(0.0000)	4.20E-05	(0.0000)	3.60E-05	(0.0000)
σ_2^2	7.35E-03	(0.0012)	4.41E-03	(0.0006)	3.50E-03	(0.0008)	5.79E-03	(0.0005)	6.79E-03	(0.0008)	7.04E-03	(0.0010)
σ_{12}	1.87E-04	(0.0001)	9.30E-05	(0.0000)	−4.30E-05	(0.0000)	1.95E-04	(0.0000)	2.16E-04	(0.0000)	2.11E-04	(0.0000)
r_c	0.937		0.969		0.978		0.913		0.916		0.929	

Table 4 Ratios (and standard errors) of means (R_b from [6]) and standard deviations (R_s from [8]) of county-level absolute residuals between fixed-effects model [1] and mixed-effects models [2] and [3] for three inventory cycles in PA, USA

Statistic	Tree-level random effects (model 2)						Stand-level random effects (model 3)					
	Cycle 1		Cycle 2		Cycle 3		Cycle 1		Cycle 2		Cycle 3	
	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)	Estimate	(Std. err.)
R_b	0.524	(0.005)	0.382	(0.005)	0.352	(0.003)	0.575	(0.014)	0.573	(0.016)	0.567	(0.016)
R_s	0.497	(0.012)	0.353	(0.009)	0.332	(0.007)	0.633	(0.017)	0.631	(0.014)	0.622	(0.016)

number of selected trees increased to three, the RMSE became smaller (2.89) than the mean response model and continued to decrease with selections of five and seven trees, respectively. The results were not as good when extending the cycle 1 model fit results to cycle 3. In this case, the mean residuals were larger than for cycle 2, were generally invariant to the number of selected trees, and were slightly larger than the mean residual from the mean response model. The RMSE statistics for the cycle 3 analysis followed the same pattern as for cycle 2. In comparison to the mean response model, RMSE was higher when using model [2] and only one tree was selected; however RMSE was smaller when three or more trees were selected.

Evaluation of the protocol where a selected number of trees on the plot would be subsampled for height measurements to predict REs for the unmeasured trees produced better results than taking averages of existing RE (Table 6). Regardless of the number of trees selected and the inventory cycle used, mean residuals from model [2] were always closer to zero than mean residuals from the mean response model [1]. Similarly, RMSE values were smaller in all scenarios and the differences in RMSE between models [1] and [2] were greater than the differences obtained using mean RE values.

Discussion

The results of the model fitting exercise were consistent with many of the studies referenced earlier in this paper in that superior predictive ability of mixed-effects models compared to fixed-effects models was demonstrated. A particularly noteworthy outcome was that the benefit was realized at various spatial scales, including for the model fitting data (statewide) and also for subpopulations (counties) (Table 4). This alleviates some of the concern of model prediction bias for these smaller areas, which is commonly found when applying fixed-effects only models. Thus, these results substantiate the speculation in Westfall (2015) that using mixed-effects models may be a viable alternative to fitting fixed-effects models to relatively small spatial domains to avoid localized bias issues. However, further testing of other species and geographic locations are needed to substantiate the global applicability of this result. In the context of this paper, this outcome only pertains to the current inventory data to which the model was fitted and does not address the issues of model application in other inventory cycles.

Application of the mixed models to data collected on the same trees in successive forest inventory cycles indicated quick deterioration in model

Table 5 Mean residual, variance of residuals, and root mean squared error in predicted heights for trees in cycle l using model parameters from cycle (l'). $RMSE_{Hl}$ is the root mean squared error between observed and predicted heights using model parameters from cycle l

$l(l')$	Tree-level random effects (model 2)				Stand-level random effects (model 3)			
	$RMSE_{Hl}$	$\bar{\varepsilon}_{Hl}\left(l'\right)$	$Var(\varepsilon)_{Hl}\left(l'\right)$	$RMSE_{Hl}\left(l'\right)$	$RMSE_{Hl}$	$\bar{\varepsilon}_{Hl}\left(l'\right)$	$Var(\varepsilon)_{Hl}\left(l'\right)$	$RMSE_{Hl}\left(l'\right)$
2(1)	1.17	0.30	6.50	2.57	1.90	0.51	6.71	2.64
3(1)	1.00	0.81	6.50	2.68	1.81	0.99	7.24	2.87
3(2)	1.00	0.74	5.24	2.41	1.81	0.85	5.95	2.58

Table 6 Mean residual, variance of residuals, and root mean squared error in predicted heights. The analysis of mean REs uses fixed-effects coefficients from cycle *l'* applied to trees in cycle *l* for model [1] and fixed-effects coefficients from cycle *l'* combined with mean RE from CYCLE *l'* for model [2]. The analysis of subsample trees uses fixed-effects coefficients from cycle *l'* and the stand-level RE predicted from subsampled trees in cycle *l*

	No. trees	Cycle	Fixed-effects only (model 1)			Tree-level random effects (model 2)		
			$\bar{\varepsilon}_{Hl}\left(l'\right)$	$Var(\varepsilon)_{Hl}\left(l'\right)$	$RMSE_{Hl}\left(l'\right)$	$\bar{\varepsilon}_{Hl}\left(l'\right)$	$Var(\varepsilon)_{Hl}\left(l'\right)$	$RMSE_{Hl}\left(l'\right)$
Mean of cycle 1 RE	1	2	−0.25	8.61	2.94	0.03	8.85	2.98
		3	0.92	9.15	3.16	1.08	9.19	3.22
	3	2	−0.25	8.61	2.94	0.00	8.35	2.89
		3	0.92	9.15	3.16	1.02	8.65	3.11
	5	2	−0.25	8.61	2.94	−0.03	8.30	2.88
		3	0.92	9.15	3.16	1.00	8.51	3.08
	7	2	−0.25	8.61	2.94	−0.05	8.26	2.87
		3	0.92	9.15	3.16	0.98	8.38	3.06
Predict RE from subsample	1	2	0.16	9.60	3.10	0.11	7.73	2.78
		3	0.84	9.20	3.15	0.41	7.10	2.70
	2	2	0.19	9.50	3.09	0.16	6.82	2.62
		3	0.86	9.13	3.14	0.29	6.30	2.53
	3	2	0.22	9.35	3.07	0.19	6.34	2.53
		3	0.87	9.09	3.14	0.26	5.96	2.46
	4	2	0.27	9.34	3.07	0.24	6.06	2.47
		3	0.89	9.00	3.13	0.23	5.55	2.37
	5	2	0.27	9.48	3.09	0.15	5.94	2.44
		3	0.92	8.98	3.13	0.23	5.39	2.33

performance. Predictive bias (as measured by the mean residual) increased as the time interval lengthened, suggesting a shift in the empirical relationships described by the initial regression analysis. It was somewhat unexpected to find that the residual variance was rather invariant to the time interval for model [2] (Table 5). The overall uncertainty quantified by RMSE suggests that only marginal differences are attributable to whether the REs are specified at the tree or stand level. The primary finding of this analysis was that applying the fixed and random parameters resulting from cycle 1 to the same trees surviving to subsequent cycles may be unwise. This conclusion is also supported by the regression analysis results shown in Table 3, where systematic trends in fixed effects coefficients and random effects variances are often found across inventory cycles. Due to the lack of independence among some observations across inventory cycles, it is difficult to ascertain if any of these differences are statistically significant; however, the deterioration of model predictive ability over time is clear evidence of practical importance.

While the concept of fitting a mixed model and using it in subsequent inventories to supplant costly field measurements is appealing, employing such a practice requires consideration of future implementation issues. Clearly, REs could be assigned their expected value (0) for subsequent inventories; however, the advantage of improved predictive ability is lost and the implementation essentially reverts to current practices of using fixed-effects models. Maintaining the use of RE in the prediction model requires additional analytical work and possibly additional field work. In this analysis, it was shown that taking mean values of REs from the initial model fit and applying them to new inventory trees in subsequent cycles provided only marginal gains (providing at least three trees were selected) over the fixed-effects model [1] for subsequent cycles (Table 6). The most promising method evaluated was to

subsample some trees on the plot in each inventory cycle and use that information to predict REs for the remaining unmeasured trees. The analysis revealed that RMSE can be reduced by 10–25 % in comparison to model [1], depending on the number of subsampled trees. Sampling more trees corresponded to larger decreases in RMSE. These results are consistent with outcomes reported previous studies (Castedo-Dorado et al. 2006; Crecente-Campo et al. 2010). Thus, it is recommended that long-term implementation of mixed models in forest inventory applications be supported by taking some field measurements to support the prediction of REs for unmeasured trees needing model-predicted attribute values, e.g., tree height. It should be noted here that this method will result in stand-level RE for these trees, in that the predicted RE arising from the subsampled trees will be the same for all the remaining unmeasured trees on the plot. In this sense, the actual implementation is similar to model [3], but only for the portion of trees not measured for height.

In regards to measuring a subsample of trees in each cycle, there is a key point to consider. It was shown that more subsampling decreased the RMSE for the remaining trees to which the model was applied. However, the number of trees subsampled should be in context to the number of trees on the plot. For example, if the protocol was to subsample five trees on the plot and the plot had seven trees total, more than 70 % of the trees on the plot were measured to support the prediction of REs for the remaining 30 %. In these cases, it may be more sensible to simply measure all trees and avoid the complexity of predicting REs. Gómez-García et al. (2014) suggest that fixed-effects models can be fitted for individual plots if the subsample contains at least 12 trees; however, their data were from even-aged, birch-dominated stands. Additional evaluation of this recommendation would be needed for large-area forest inventories where a wide range of species and stand conditions are likely to be encountered.

An analogous concern is the level of species aggregation established in the model fitting process. It is often the case that models are fitted separately by tree species or species groups, resulting in numerous sets of coefficients that are species (group) specific. Thus, in the context of the five-tree subsample protocol, five trees need to be subsampled for each species (group) occurring on the plot. This could again result in most trees on the plot being subsampled and the value of the REs prediction for the few remaining trees becoming questionable. An alternative protocol would be to subsample only a few trees in each plot and then expand the area from which the nearest tree(s) were selected, e.g., countywide instead of within plot. However, such an approach would rely on there being a sufficient numbers of trees measured on other plots, which may not be the case for species of uncommon occurrence. In this context, determination of an appropriate subsampling protocol should be carefully considered, and it may be prudent to modify the resulting database such that differentiation between observed and model-predicted tree height values is possible.

The underlying premise of this study is that a model is fitted to a cycle of inventory data, and that model must then be used as for prediction in subsequent inventories where model performance may be enhanced using either prior information or newly collected data (Russell 2015). Thus, the analysis differs from other research that uses longitudinal data from several repeated measurements. The seminal paper by Lappi (1997) examined the influence of time (and stand variables as well) on height-diameter model parameters. In application, the model parameters vary with stand age; which could be problematic in areas where uneven-aged stands are often encountered or stand age is unknown. A conclusion similar to that of this study was that model calibration over time was best accomplished by measuring heights of some sample trees at each remeasurement. Mehtätalo (2004) and Schmidt et al. (2011) took a similar approach to incorporating time as a modifier of model parameters, but used central measures of stand diameter instead of age. The use of stand information derived from tree diameters seems generally more viable than age, assuming the stand attributes can be accurately assessed for all species occurring on sample plots (including those with very few trees). A comparison of these methods with those evaluated is this paper is warranted in cases where longitudinal data are available.

Lastly, some discourse regarding the use of model predictions instead of observed field data is warranted. Observed field data is often hailed as the gold standard, providing a direct and current observation of the phenomena at hand. However, field data are subject to bias and variability in the same context as model predictions; although the field data bias is routinely (but often incorrectly) considered to be zero and the variation negligible. In continuous forest inventories, consistency in measurement over time is critical for assessing forest resource trends. However, such consistency is difficult

to ensure as personnel, training practices, and field methods change over time. Substituting model predictions for field data is also subject to bias and variation. The residual error is often quantified when the model is developed, but prediction bias and uncertainty for subpopulations is unknown. The temporal concern is whether the relationships described by the model are still valid in subsequent inventory cycles. Given that both field measures and models are subject to bias, variability, and temporal inconsistency, models may be preferred to reduce field work costs and to provide consistent predictions, i.e., the same result is obtained for a given set of input values. In this context, the concepts presented can be extended to other models used to provide predicted values in a temporal framework. This includes models employed in continuous forest inventory settings or in growth-and-yield model systems.

Conclusions

Improvements in predictive ability associated with mixed-effects models as compared to fixed-effects models has been well documented for both the model fitting data as well as new observations. However, the practical implementation of mixed models into continuous forest inventory efforts has largely been overlooked, primarily due to familiarity with the use of fixed-effects models and concern regarding the complexity of maintaining/predicting REs parameters. The primary advantage of mixed models is the considerable reduction in bias and variance of model predictions, particularly for subpopulations where fixed-effects models may perform poorly. From an implementation standpoint, the disadvantage is that mixed-effects models must be continually maintained, whereas in contrast, fixed-effects models require no ongoing maintenance effort as long as the underlying relationships remain valid. The most promising long-term maintenance strategy for mixed models requires that some field data be collected in order to obtain reasonably accurate REs. This, in turn, results in more accurate estimates of the attribute of interest (e.g., tree height). In summary, effective implementation of mixed models requires an initial investment in developing the algorithms to predict REs and making associated database modifications, with an ongoing commitment to implement a field protocol for subsampling. The efficacy of employing mixed models must be evaluated by forest inventory managers in light of the costs/benefits to the overall inventory program.

Acknowledgments The author would like to thank Mahadev Sharma, Felipe Crecente-Campo, John Stanovick, Matt Russell and an anonymous reviewer for providing valuable comments that substantially improved the manuscript.

References

Adame, P., del Río, M., & Cañellas, I. (2008). A mixed nonlinear height–diameter model for pyrenean oak (*Quercus pyrenaica* Willd.). *Forest Ecology and Management, 256*, 88–98.

Bechtold, W. A. (2003). Crown-diameter prediction models for 87 species of stand-grown trees in the eastern United States. *Southern Journal of Applied Forestry, 27*, 269–278.

Bechtold, W. A., & Scott, C. T. (2005). The forest inventory and analysis plot design. In W. A. Bechtold, & P. L. Patterson (Eds.). *The enhanced forest inventory and analysis program – national sampling design and estimation procedures* (pp. 27–42). General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Southern Research Station.

Castedo-Dorado, F., Diéguez-Aranda, U., Barrio Anta, M., Sánchez Rodríguez, M., & von Gadow, K. (2006). A generalized height–diameter model including random components for radiata pine plantations in northwestern Spain. *Forest Ecology and Management, 229*, 202–213.

Coble, D. W., & Lee, Y.-J. (2011). A mixed-effects height-diameter model for individual loblolly and slash pine trees in East Texas. *Southern Journal of Applied Forestry, 35*, 12–17.

Crecente-Campo, F., Tomé, M., Soares, P., & Diéguez-Aranda, U. (2010). A generalized nonlinear mixed-effects height-diameter model for *Eucalyptus globulus* L. in northwestern Spain. *Forest Ecology and Management, 259*, 943–952.

de-Miguel, S., Mehtätalo, L., & Durkaya, A. (2014). Developing generalized, calibratable, mixed-effects meta-models for large-scale biomass prediction. *Canadian Journal of Forest Research, 44*, 648–656.

Feldpausch, T. R., Banin, L., Phillips, O. L., et al. (2011). Height-diameter allometry of tropical forest trees. *Biogeosciences, 8*, 1081–1106.

Gómez-García, E., Diéguez-Aranda, U., Castedo-Dorado, F., & Crecente-Campo, F. (2014). A comparison of model forms for the development of height–diameter relationships in even-aged stands. *Forest Science, 60*, 560–568.

Gregoire, T. G. (2012). Height-dbh bibliography: 1932 – present. http://environment.yale.edu/content/documents/00001660/Height-DBH.pdf. Accessed 14 May 2015.

Jayaraman, K., & Zakrzewski, W. T. (2001). Practical approaches to calibrating height–diameter relationships for natural sugar maple stands in Ontario. *Forest Ecology and Management, 148*, 169–177.

Jiang, L., & Li, Y. (2010). Application of nonlinear mixed-effects modeling approach in tree height prediction. *Journal of Computers, 5*, 1575–1581.

Lappi, J. (1997). A longitudinal analysis of height/diameter curves. *Forest Science, 43*, 555–570.

Mehtätalo, L. (2004). A longitudinal height–diameter model for Norway spruce in Finland. *Canadian Journal of Forest Research, 34*, 131–140.

Reams, G. A., Smith, W. D., Hansen, M. H., Bechtold, W. A., Roesch, F. A., & Moisen, G. G. (2005). The forest inventory and analysis sampling frame. In W. A. Bechtold & P. L. Patterson (Eds.). *The enhanced forest inventory and analysis program – national sampling design and estimation procedures* (pp. 11–26). General Technical Report SRS-80. Asheville, NC: U.S. Department of Agriculture, Southern Research Station.

Russell, M. B. (2015). Influence of prior distributions and random effects on count regression models: implications for estimating standing dead tree abundance. *Environmental and Ecological Statistics, 22*, 145–160.

Schmidt, M., Kiviste, A., & von Gadow, K. (2011). A spatially explicit height–diameter model for Scots pine in Estonia. *European Journal of Forest Research, 130*, 303–315.

Temesgen, H., Monleon, V. J., & Hann, D. W. (2008). Analysis and comparison of nonlinear tree height prediction strategies for Douglas-fir forests. *Canadian Journal of Forest Research, 38*, 553–565.

Trincado, G., & Burkhart, H. E. (2006). A generalized approach for modeling and localizing stem profile curves. *Forest Science, 52*, 670–682.

Trincado, G., VanderSchaaf, C. L., & Burkhart, H. E. (2007). Regional mixed-effects height–diameter models for loblolly pine (*Pinus taeda* L.) plantations. *European Journal of Forest Research, 126*, 253–262.

U.S. Forest Service. (2007). Forest inventory and analysis national core field guide, volume 1: field data collection procedures for phase 2 plots, version 4.0. http://socrates.lv-hrc.nevada.edu/fia/dab/databandindex.html. Accessed 14 May 2015.

Vargas-Larreta, B., Castedo-Dorado, F., Álvarez-González, J. G., Barrio-Anta, M., & Cruz-Cobos, F. (2009). A generalized height-diameter model with random coefficients for uneven-aged stands in El Salto, Durango (Mexico). *Forestry, 82*, 445–462.

Vonesh, E. F., & Chinchilli, V. M. (1997). *Linear and nonlinear models for the analysis of repeated measurements*. New York: Marcel Dekker.

Vonesh, E. F., Chinchilli, V. M., & Pu, K. (1996). Goodness-of-fit in generalized nonlinear mixed-effects models. *Biometrics, 52*, 572–587.

Westfall, J. A. (2006). Predicting past and future diameter growth for trees in northeastern U.S. *Canadian Journal of Forest Research, 36*(6), 1551–1562.

Westfall, J. A. (2015). Spatial-scale considerations for a large area forest inventory regression model. *Forestry, 88*, 267–274.

Westfall, J. A., & Laustsen, K. M. (2006). A merchantable and total height model for tree species in Maine. *Northern Journal of Applied Forestry, 23*, 241–249.

Woodall, C. W., Heath, L. S., Domke, G. M., & Nichols, M. C. (2011). Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the U.S. forest inventory, 2010. General Technical Report NRS-88. Newtown Square, PA: U.S. Department of Agriculture, Forest Service, Northern Research Station.

Zhang, L., & Gove, J. H. (2005). Spatial assessment of model errors from four regression techniques. *Forest Science, 51*, 334–346.

Zhang, L., & Shi, H. (2004). Local modeling of tree growth by geographically weighted regression. *Forest Science, 50*, 225–244.