A MIXED-EFFECTS MODEL TO ESTIMATE STAND VOLUME BY MEANS OF SMALL FOOTPRINT AIRBORNE LIDAR DATA FOR AN AMERICAN AND A GERMAN STUDY SITE

Johannes Breidenbach [a,*], Robert J. McGaughey [b], Hans-Erik Andersen [c], Gerald Kändler [a], Stephen E. Reutebuch [b]

[a] Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg, Abteilung Biometrie und Informatik
Wonnhaldestr. 4, 79100 Freiburg, Germany - johannes.breidenbach@forst.bwl.de
[b] USDA Forest Service, Pacific Northwest Research Station
University of Washington, PO Box 352100, Seattle, WA 98195-2100, USA
[c] USDA Forest Service, Pacific Northwest Research Station
Anchorage Forestry Sciences Laboratory, 3301 C Street, Suite 200, Anchorage, AK 99503, USA

KEY WORDS: Mixed-effects models, random effects, lidar, forest inventory, stand volume

ABSTRACT:

Similar datasets (inventory plots, stand maps and lidar data) were available for study sites in the USA and Germany. These datasets are grouped or hierarchical in that several sample plots are located within a stand and the stands are located within two study sites. Fixed-effects models and mixed-effects models with a random intercept on the stand level were fit to each dataset. Mean lidar raw data return height and its interaction term with canopy cover as well as its interaction term with the coniferous proportion were found to be the most influential predictor variables. The mixed-effects models significantly improved the estimates and especially reduced the bias which was present for numerous stands in the estimates of the fixed-effects models. This resulted in a slight increase of the variance within the stands. The RMSE for the German study site was higher (34.7% and 29.7% for fixed- and the mixed-effects model respectively) than on the US study site (19.2% and 16.8% for fixed- and the mixed-effects model respectively). A mixed-effects model with random effects on the study site and stand level was fit to the combined dataset. It showed almost the same errors as the local mixed-effects models (17.6% and 29.8% for the US and the German study site respectively). Hence a single model is sufficient to make estimates for both datasets. The study shows the potential of mixed-effects models in this context. It illustrates that the common practise of fitting different models for different strata may be unnecessary.

1 INTRODUCTION

Height and density metrics, derived from lidar (light detection and ranging) point clouds can be used as predictor variables in statistical models to estimate forest parameters at the stand or plot level (Næsset, 2004; Andersen et al., 2005, among others). Such models are usually fit using sample plots where both lidar (covariates) and ground-truth information (response) are available. To map the variable of interest, the entire lidar dataset is gridded into tiles having the same size as a sample plot. Then the predictor variables are computed and the regression models are applied to every tile. Compared to plot-based inventories, estimation errors can be significantly decreased for the area of interest (e.g. a single forest stand), since the number of observations (i.e. the tiles) is usually much higher than the number of sample plots within a stand.

The predictor variables derived from lidar data are mainly related to the vegetation height and structure (e.g., height- and density metrics, crown cover). The vegetation cover can, under certain circumstances, also be classified into broadleaf and coniferous trees. However, information about the site quality or tree species cannot be derived without additional data. Therefore, predictions for stands with rare site index classes or tree species compositions might deviate from the mean model, resulting in a bias.

If the grouping structure (i.e., the stand boundaries) is known, the deviation from the mean model of plot estimates within a stand can be utilized to reduce the bias using mixed-effects models (mixed models). From the statistical point of view, the grouping structure has to be considered since the observations are not independent. In a mixed model, the effects of the variable that indicates the level of grouping (i.e. the stand-ID) are assumed to be a random sample of a larger population that vary randomly around a population mean. This is referred to as random effects. Mixed models with forestry application were discussed by Lappi and Bailey (1988). An in-depth description of mixed models is given for example by Pinheiro and Bates (2002).

In a mixed model, the variance is split into within and between group variance. The coefficients and standard errors for predictor variables that vary less within than between the groups are therefore more accurate. Another advantage of a mixed model, compared to a fixed-effects model with the grouping level as a dummy variable, is that predictions can also be made for individuals with grouping levels that did not exist in the dataset used to fit the model (e.g., in our case those stands without sample plots). In a forest inventory context, a mixed model provides an additional advantage. A model can be fit to a large dataset (e.g., to a well inventoried public forest) and subsequently be calibrated with just a few sample plots for a new forest area (e.g., a small private forest). (A new model would need to be fit, if a fixed-effects model were used.)

Publications regarding the estimation of volume and biomass on the plot level include these of Næsset (2002) who created separate models for different ages and site qualities and achieved R^2 between 0.80 and 0.93 in a boreal forest and Means et al. (2000) for Douglas fir stands in the Cascade Mountains (Oregon, USA) who reported R^2 between 0.93 and 0.95. In a study by Packalén and Maltamo (2006) in a Finnish boreal forest, plot volume was assigned to tree species by using the k-MSN method. They report a RMSE of roughly 24% for estimates of total volume. Aardt et al. (2006) segmented homogeneous forest units first and used the lidar vegetation height distribution and the field data for the units to calibrate prediction models. They report R^2 between 0.58 and 0.79 for their study which was located in Virginia (temperate mixed forests).

The objective of this research was to develop a single statistical model for estimates on study sites located in the USA and Germany. Furthermore, we wanted to find out if information about the stand-level grouping of sample plots can be used to further improve the regression models. The datasets contain several levels of grouping or random effects: (i) The study sites are two random samples of all potentially existing study sites, (ii) the stands are grouped within the study site and are a random sample of the stands within each study site. Due to this structure, the resulting model is referred to as a multi-level or hierarchical model.

2 MATERIAL AND METHODS

2.1 American Dataset

The study site in the USA is part of Capitol State Forest and is managed by the Washington State Department of Natural Resources. The terrain is moderate with elevations varying from 300 to 425 meters and ground slopes up to 30°. The forest is composed primarily of Douglas-fir (*Pseudotsuga menziesii* (Mirb.) Franco; 81%) and western hemlock (*Tsuga heterophylla* (Raf.) Sarg.; 13%). Additional species present include western red cedar (*Thuja plicata* Donn ex D. Don; 2%) and few deciduous hardwoods such as red alder (*Alnus rubra* Bong.; 3%) and maple (*Acer spp.*; <1%). The height of dominant trees in the study area was approximately 50 meters (table 1). As part of a forest management study (Curtis et al., 2004), the canopy of the 70-year-old forest stand was partially harvested in 1998, resulting in four different residual canopy density classes.

A total of 98 fixed area field inventory plots were established over a range of stand conditions in 1999. Plot sizes ranged from 0.02 to 0.2 ha. Measurements acquired at each plot included species and diameter at breast height (DBH) for all trees greater than 14.2 cm in diameter. In addition, total height was measured on a representative selection of trees using a hand-held laser rangefinder. A detailed description of the plot measurement protocol can be found in Curtis et al. (2004). Inventory plot locations were surveyed with a Topcon ITS-1 total station and are accurate to within 1 m.

The Saab TopEye lidar system mounted on a helicopter was used to map approximately 5.25 km^2 of the study area in the spring of 1999 (before foliation). Table 2 summarizes the flight parameters and instrument settings for the data acquisition. Data for each return included the pulse number, return number for the pulse (up to four returns were recorded per pulse), X, Y, elevation, off-nadir angle and intensity.

2.2 German Dataset

The 49 km^2 study area is located approximately 60 km north of Freiburg. Elevations range between 400 and 1050 m above sea level. The average gradient across the site is approximately 12° with some slopes of up to 35°. The average forest stand is approximately 1.2 ha in size. Tree heights within the study area range from 5 to 47 m, with an average height of 23 m. Norway spruce (*Picea abies* (L.) Karst.; 65%), silver fir (*Abies alba*; 17%), beech (*Fagus sylvatica* L.; 9%) and Scotts pine (*Pinus sylvestris*; 6%) are the most common tree species. The forest is managed using a group selection system, where the regeneration phase may take several decades (clearcuts are not common in Germany).

A regular forest enterprize inventory was conducted in the second half of 2003 in the state forest of the study area, using plots positioned on the intersections of a 100 x 200 m sample grid. The horizontal accuracy of the inventory plot locations is estimated to be better than 10 m. Forest characteristics were recorded within sample plots consisting of four concentric circle plots (i.e. they have the same centre) with radii of 2 m, 3 m, 6 m and 12 m. Trees with a diameter at breast height (DBH) greater than 7 cm, 10 cm, 15 cm and 30 cm, respectively, were recorded within the four circle plots. The heights of the two tallest trees per species were measured in each plot using a *Vertex* angle measurement instrument. The height of the remaining trees within a plot were estimated using forest stand height curves and the DBH. Single tree volumes were calculated using DBH and height as parameters for taper and volume functions of the Baden-Württemberg state forest service (Kon-Allan et al., 2004). Plots intersecting stand or forest borders were excluded for this study. A total of 1061 inventory plots, with an overall area of 48 ha, were used as terrestrial reference data for the remotely sensed data. Stand boundaries were digitized from orthophotos in 2003 and were adjusted to meet operational purposes during the field work. Additional information describing the stands that could be used as covariates were not available for this study.

Lidar data were acquired in spring 2003 (before foliation) using the Optech ALTM 1225 airborne laser scanner. Adjacent swaths overlapped about 50% (table 2). First and last return laser data were automatically classified into ground and vegetation hits by the data provider (TopScan).

2.3 Computation of predictor variables

A digital terrain model (DTM) and a digital surface model (DSM) was computed for both test sites using the software *TreesVis* (Weinacker et al., 2004) for the German and *Fusion 2.0* (McGaughey et al., 2004) for the American study site. An evaluation of the American DTM, presented in Reutebuch et al. (2003), found an average lidar elevation error of 22 cm. For the German study site, a DSM was derived from the first ($\mathrm{DSM_F}$) and the last return ($\mathrm{DSM_L}$) vegetation returns. Canopy height models (CHM, $\mathrm{CHM_F}$, $\mathrm{CHM_L}$) were computed by subtracting the DTM from the according DSMs. The lidar vegetation height was determined by calculating the difference between the elevation of the lidar vegetation data (raw data) and the corresponding DTM raster bin elevation.

Circular subsets of the same radius as the corresponding sample plot were created from the lidar raw data. The 0$^{\text{th}}$, 25$^{\text{th}}$, 50$^{\text{th}}$, 75$^{\text{th}}$ and 100$^{\text{th}}$ percentiles and the mean of the lidar vegetation heights were calculated for each subset to characterize the vegetation height distribution. Vegetation density metrics were derived by dividing the range between the highest and lowest measurement into 10 classes and determining the proportion of measurements within each class. *Fusion 2.0* was used for the raw data manipulation.

Since broadleaf trees in leaf-off condition had only a few vegetation returns in the last return data, they do not show up in the $\mathrm{CHM_L}$. Therefore, a classification of the pixels into those belonging either to coniferous or broadleaf trees was possible by subtracting the $\mathrm{CHM_L}$ from the $\mathrm{CHM_F}$. The result was normalized with the $\mathrm{CHM_F}$. By comparison with orthophotos, a threshold of 0.3 was found to separate coniferous and broadleaf pixels well (equation 1). It should be noted that Larches (*Larix spp.*) are a potential problem for this classification approach, since they are deciduous conifers. However, few Larches were present in the study area so we felt the classification approach was applicable.

$$P_i \left\{ \begin{array}{ll} 1 & (\mathrm{CHM}_{\mathrm{F},i} - \mathrm{CHM}_{\mathrm{L},i})/\mathrm{CHM}_{\mathrm{F},i} \leq 0.3 \\ 0 & (\mathrm{CHM}_{\mathrm{F},i} - \mathrm{CHM}_{\mathrm{L},i})/\mathrm{CHM}_{\mathrm{F},i} > 0.3 \end{array} \right. \quad (1)$$

Parameter	German study site		American study site	
	Mean	Max	Mean	Max
Trees per ha [ha^{-1}]	411.10	2255.00	309.10	1093.00
Mean heights [m]	23.07	42.56	36.42	51.97
Volume [m^3 ha^{-1}]	347.10	1265.00	567.90	1167.00

Table 1. Summary of forest attributes derived from sample plot data for the study sites.

Parameter	Characteristic	
	German study site	American study site
Laser pulse frequency	25,000 Hz	7,000 Hz
Scan angle	$\pm$ 20°	$\pm$ 10°
Swath width	500-600 m	70 m
Laser pulse density	0.51 m^{-2}	4 m^{-2}
Flying height	900 m AGL	200 m AGL
Flying speed	80 m $sec.^{-1}$	25 m $sec.^{-1}$
Beam divergence	0.25 mrad	2 mrad
Vertical accuracy	0.15 m	n.a.
Horizontal accuracy	0.45 m	n.a.

Table 2. Operating parameters of the Lidar sensors (n.a. = not available).

The ith pixel P will have a value of 1 if it is classified as coniferous and 0 if classified as coming from a broadleaf tree.

The proportion of coniferous trees (coniferous proportion = *CP*) at a plot, which is assumed to be equal to the proportion of pixels classified as coniferous, was calculated as the sum of coniferous pixels divided by the overall sum of pixels within a sample plot. For numerical reasons, this 0...1 distributed variable was then stretched between -infinity and infinity using the logit-transformation ($\log(\frac{p}{1-p})$). Since the lidar dataset of the US study site was not separated into first- and last return data, a computation of *CP* was not possible.

The percentage of canopy cover (*CC*) on a sample plot was computed as the number of pixels in the CHM (CHM_F for the German data set) greater than 1 m divided by the total number of pixels within the plot. As for *CP*, a logit transformation was applied to this variable.

2.4 Modeling

Modeling consisted of two steps: (i) Select adequate predictor variables, (ii) fit mixed models by adding random effects. To select predictor variables, scatter plots and correlations of the response variables over the height metrics were analyzed for the German study site. The mean vegetation height measured by lidar data (*mean.l*) was found to be the most influential predictor variable. Since the variance increases as the response variable increases (heteroscedasticity), *mean.l* was also used as a predictor variable for the variance function. More precisely, a generalized-least-squares (GLS) regression was used with weights based on $mean.l^{2\delta}$ where δ was estimated during the fitting of the model (Pinheiro and Bates, 2002, p. 208).

Since the height metrics vary depending on the canopy structure, we wanted to know if the model improves as interaction terms between *mean.l* and the canopy cover and between *mean.l* and the crown shape (expressed as conifer proportion) are considered (equation 2). We also explored whether or not density metrics further improved the model. The selected model (fixed-effects model for the German study site) was then re-fit using data for the American study site and the coefficients were compared. As it was not possible to compute the coniferous proportion for the US study site, this variable was not included in the fixed effect model for the US data.

The fixed-effects model can be written as

$$\begin{aligned} y_k = \beta_0 + \beta_1 mean.l_k + \beta_2 CC_k + \beta_3 CP_k + \\ \beta_4 CC_k \cdot mean.l_k + \beta_5 CP_k \cdot mean.l_k + \epsilon_k, \\ k = 1, ..., n, \quad \epsilon_k \sim N(0, \sigma^2 mean.l_k^{2\delta}), \end{aligned} \tag{2}$$

where y_k is the response variable for the *k*th sample plot, $\beta_0..\beta_5$ are the coefficients, ϵ_k is an independent error term with a variance model depending on *mean.l* and δ and *n* is the number of sample plots.

In the second step, random effects for the intercept on the stand level were introduced for the local models (equation 3). Their results were compared to a global model with random effects for the intercept on the study site as well as on the stand level. For the global model, we checked if it was necessary to have a random effect for the coefficients. To do this, models with a random intercept on the study site and the stand level as well as a random effect for either one of the coefficients (equation 4) were compared with the global model with the random effect only for the intercept using a F-test.

The following equation is the general form of a local mixed model

$$\begin{aligned} y_{jk} = \beta_0 + b_{0,j} + \beta_1 x_{1,jk} + .. + \beta_m x_{m,jk} + \epsilon_{jk} \\ k = 1, ..., n_j, \quad b_j \sim N(0, \sigma_1^2), \\ \epsilon_{jk} \sim N(0, \sigma^2 mean.l_k^{2\delta})). \end{aligned} \tag{3}$$

Here y_{jk} is the response variable for the *k*th sample plot in the *j*th stand, $x_{1,jk}..x_{m,jk}$ are the *m* fixed effects, $\beta_0..\beta_m$ are the coefficients thereof and n_j is the number of sample plots within a stand. The stand random effects $b_{0,j}$ are assumed to be independent for different *j* and the within-group errors, ϵ_{jk} are assumed to be independent of different *j* and *k* and to be independent of the random effects.

If the response variable of the *k*th sample plot in the *j*th stand within the *i*th study site is denoted as y_{ijk} *i*=1,2; *j*=1,...,l_i; *k*=1,...,n_j, with l_i as the number of stands in the *i*th study site and $b_{0,i}; b_{0,ij}$ are random effects for the intercept on the study site and stand level respectively, an example for the global model including a random effect on the study site level for the coefficient of the first fixed effect ($b_{1,i}$) can be expressed as

$$y_{ijk} = \beta_0 + b_{0,i} + b_{0,ij} + (\beta_1 + b_{1,i})x_{1,ijk} + ..+ \beta_m x_{m,ijk} + \epsilon_{ijk}. \quad (4)$$

The random effects are, technically speaking, not parameters of the statistical model. Nevertheless, their values (Best Linear Unbiased Predictors, BLUPs) can be estimated. Details on the estimation of BLUPs can be found in Pinheiro and Bates (2002).

A leave-one-out cross-validation procedure was used to check for potential overfitting of the data. A close similarity of the RMSE and the RMSE of the cross-validation (RMSE.CV) indicates that the model is not overfitting the data (Andersen et al., 2005). All statistical analysis were carried out with the software package *R* (R-Development-Core-Team, 2006) including the library *nlme* (Pinheiro and Bates, 2002) for the fitting of mixed-effects models.

3 RESULTS

3.1 Selected models

Canopy closure and coniferous proportion as well as their interaction with *mean.l* significantly improved the linear model for the German study site. The addition of density metrics seemed to enhance the model fit significantly but improved the R^2 less than 1%. They were therefore not included in the model in order to keep the amount of predictor variables to a minimum. The selected model of the German study site explains about 70% of the variance and leads to an RMSE of ca. 35%.

The model including the same predictor variables as the model for the German study site showed better goodness-of-fit measures (R^2 of 0.86 and RMSE of ca. 17%) for the US study site. Nevertheless, it was also tested, if the model improves as other height metrics (e.g. the median or the 75^{th} percentile) serve as predictor variables instead of *mean.l*. But none of the models including those variables was significantly different from the model including *mean.l*. We concluded that the same predictor variables can be used for the German and for the US study site. Additional model attributes and RMSE can be found in tables 3 and 4.

3.2 Mixed effect models

Random effects on the stand level improved the models for both study sites significantly. In general, it can be observed that the median residual per stand is closer to zero, while the variance slightly increases. This also means that the prediction for some observations gets better, while the opposite is true for others. In other words, the mixed models lead to a decreased bias with a trade-off of higher variance. This of course, is most present in stands where the bias of the fixed effect model was large. However, the variance within the stands is relatively high, especially for the German study site. Therefore, the bias will not be eliminated completely (figure 1 and table 5).

For the global model, besides the random effects on the stand level, only a random effect for the coefficient of the interaction between the canopy cover and *mean.l* significantly improved the model. This suggests, that the other coefficients do not differ significantly between the study sites. Interestingly, the RMSE does not increase very much, meaning that this model can be used for predictions at both study sites.

The global model can be expressed as

$$\begin{aligned} y_{ijk} = \beta_0 + b_{0,i} + b_{0,ij} + \beta_1 mean.l_{ijk} + \beta_2 CC_{ijk} + \\ \beta_3 CP_{ijk} + (\beta_4 + b_{1,i})CC_{ijk} \cdot mean.l_{ijk} + \\ \beta_5 CP_{ijk} \cdot mean.l_{ijk} + \epsilon_{ijk}. \end{aligned} \quad (5)$$

The other models can be written the same way without $b_{0,i}$ and $b_{1,i}$ for the local models with random effects, without $b_{0,i}$, $b_{1,i}$ and $b_{0,ij}$ for the local models with fixed effects only and without CP for the US models.

3.3 Characteristics of the regression models

The slope of *mean.l* is slightly higher for the US study site than for the German site. This is also true for the coefficient of canopy cover. The models will predict higher volumes with an increase of canopy cover or an increased number of coniferous trees per plot. The global model produces almost the same predictions for given *mean.l* but differs slightly more from the local models given canopy cover (figures 2 and 3).

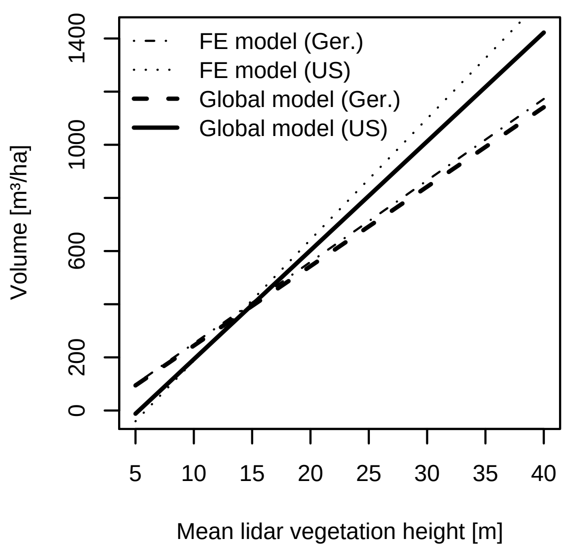

Figure 2. Comparison of predictions of local fixed-effects models (FE) with the global mixed model (Ger.=Germany). *mean.l* alters, *CC* and *CP* are fixed at 100%.

4 DISCUSSION

In this study, we compared fixed-effects models with mixed-effects models containing random effects on the stand level and on the study site level. The grouping information was used to calibrate the mixed models on the stand level using the variance information of sample plots located within a stand. A drawback of this method is that this information can only be used reliably for stands that contain several sample plots if the within stand variance is high as it was the case in this study.

Reasons for bias in some stands, besides rare tree species and site indices, might be uncommon taper shapes, varying density of small trees in the understory (i.e. two layers of trees, which probably does not change *mean.l* explicitly) or other incidents that change the canopy structure but are not reflected in the selected covariates.

Coefficients	**FE model (Ger.)**		**FE model (US)**		**RE model (Ger.)**		**RE model (US)**		**Global model**	
	Est.	p-val.	Est.	p-val.	Est.	p-val.	Est.	p-val.	Est.	p-val.
Intercept	-10.12	0.42	77.24	0.24	-3.93	0.74	57.47	0.39	-82.04	0.33
$mean.l$	15.44	<0.01	9.15	<0.01	14.98	<0.01	10.48	<0.01	14.48	<0.01
CC	-11.86	<0.01	-50.00	<0.01	-13.02	<0.01	-35.56	0.01	-13.69	<0.01
CP	5.45	0.14			3.52	0.24			5.86	0.06
$mean.l \cdot CC$	1.64	<0.01	5.27	<0.01	1.74	<0.01	4.31	<0.01	2.68	<0.01
$mean.l \cdot CP$	0.57	<0.01			0.59	<0.01			0.36	0.04
δ	0.46		0.34		0.58		0.44		0.46	

Table 3. Attributes of the fitted models (Est. = Estimate, p-val. = p-value).

	German models		**American models**		**Both**
	FE model	RE model	FE model	RE model	Global model
RMSE [m^3 ha^{-1}]	120.33	103.12	108.96	95.51	103.00
German study site					
RMSE [%]	34.67	29.71			29.76
RMSE.CV [%]	34.90	34.40			34.43
American study site					
RMSE [%]			19.19	16.82	17.58
RMSE.CV [%]			19.95	18.49	18.69
Both					
RMSE [%]					28.16
RMSE.CV [%]					32.35
R^2	0.70	0.78	0.86	0.89	0.81

Table 4. RMSE, RMSE of the cross-validation (RMSE.CV) and R^2 for the fitted models (FE = fixed effect, RE = random effect).

Stand-ID	**FE model**		**RE model**		**Global model**	
	SD	Bias	SD	Bias	SD	Bias
9	117.59	96.24	126.85	95.13	110.91	95.61
8	46.82	49.75	48.97	31.25	50.17	32.24
6	145.39	115.49	128.37	98.74	127.78	96.67
5	90.34	82.60	94.41	75.50	96.20	77.60
4	105.32	86.24	119.01	96.46	120.77	99.31
3	117.69	102.09	123.99	100.16	123.83	98.40
2	50.51	94.66	50.45	40.64	51.27	58.44

Table 5. Standard deviations (SD) of the residuals and bias for the stands on the American study site.

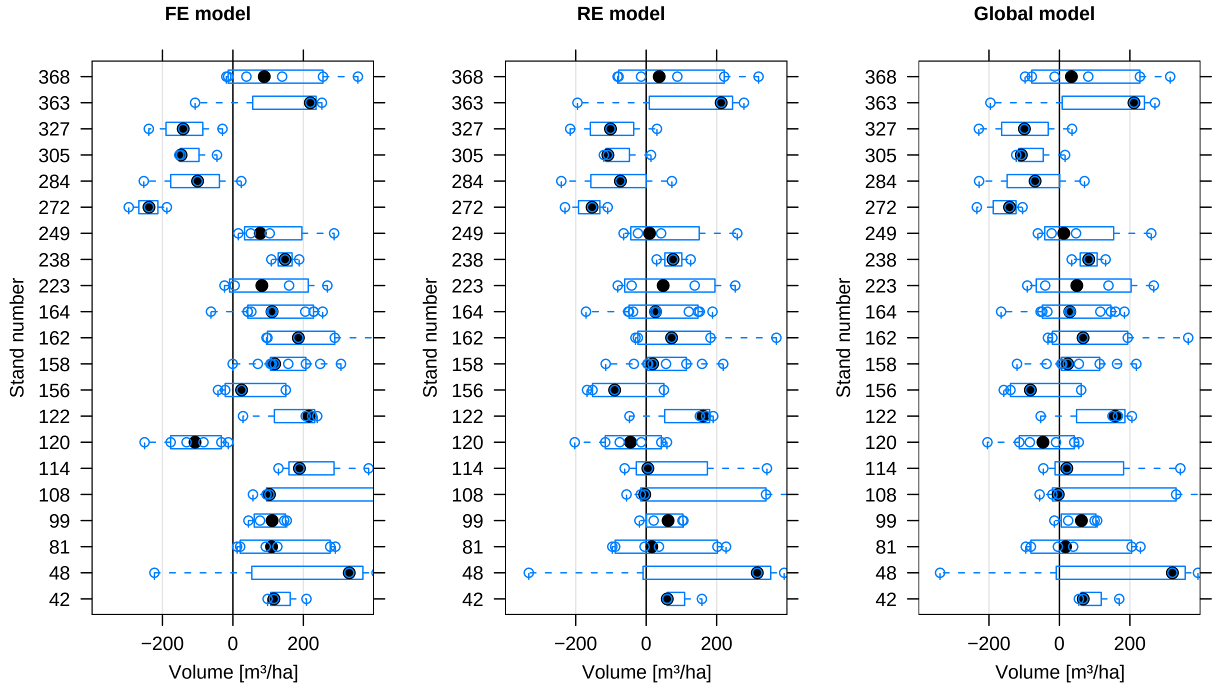

Figure 1. Residuals of a leave-one-out cross-validation for selected stands at the German study site. Stands with a mean residual >100 and at least 3 observations were selected for this graph.

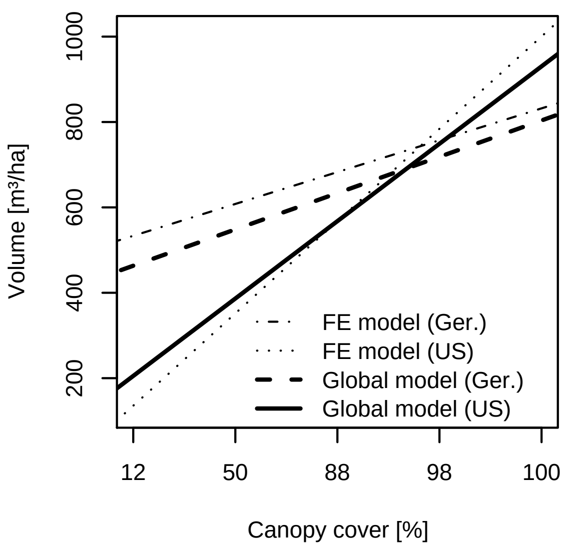

Figure 3. Comparison of predictions of local fixed-effects models (FE) with the global mixed model (Ger.=Germany). *CC* varies, *mean.l* is fixed at 30 m and *CP* at 100%.

We assume that differences in the model coefficients for the study sites can be attributed to variation in the vegetation cover and the lidar parameters:

1. The US study site is highly productive (high site index) and is stocked mainly by Douglas-fir, which is one of the fastest growing tree species in temperate forests. In comparison, the German study site encompasses a range of productivity classes, a broader range of elevations and a more diverse mix of tree species. In addition, the main coniferous tree species (Norway spruce) does not accumulate as much volume as Douglas-fir.

2. Although both lidar systems produced small footprint data, return density, footprint size and flying platforms were significantly different. This could influence the penetration rates through the canopy, amount of shadowing, etc.

However, whether vegetation or lidar parameters have a larger influence in these study results could not be determined. The same is true for possible interactions between lidar parameters and vegetation.

Interestingly, the predictor variable canopy cover improved the model more on the American study site than on the German study site. This improvement is likely related to the extensive changes in the canopy structure resulting from the silvicultural treatments carried out on the American study site. These treatments resulted in a wider range of canopy densities than was present on the German dataset.

The coefficients of the coniferous proportion indicate that the volume increases with an increasing amount of coniferous trees on a plot. This is consistent, since *mean.l* tends to be smaller for conifer dominated plots compared to plots dominated by deciduous species but having the same mean tree height due to the conifer crown shape (Breidenbach et al., 2007). Another reason for this effect is probably that the amount of usable timber is higher for most coniferous species, since the ratio of stem to branch volume is higher for coniferous trees. Therefore, similar heights correspond to more volume for conifer dominated sample plots.

The observed errors for the US study site ($\sim 17\%$) are comparable to those reported by Næsset (2002), but somewhat higher than those reported by Means et al. (2000) (73 m^3 ha^{-1} opposed to ~ 95 m^3 ha^{-1}). The errors for the German study site are much higher which is probably due to the wider range of tree species and stand types. Another reason could be that the horizontal accuracy of the field plot positions for the German study site is worse than for the US study site. Aardt et al. (2006), whose study site is probably more similar to the German site, report smaller absolute RMSE ($\sim$ 40-68 m^3 ha^{-1}) than we observed for the German site. However, since the range of stand volumes in their data is significantly smaller than in this study, the relative errors seem to be larger.

5 CONCLUSIONS AND FUTURE WORK

A mixed-effects model was fit to data from the USA and Germany. The goodness-of-fit metrics indicate, that the model fit to the combined data is almost as good as models fit to data for each site, although the stand conditions and lidar properties varied greatly between the study sites. It should be emphasized that the random effects at the stand level were able to significantly reduce the bias that was found at the stand level. The relatively expensive field data were consequently used twofold: (i) To fit prediction models, (ii) to reduce the bias by calibrating random effects and utilizing the information that they provide at the stand level. Therefore the effectiveness of the money spent to collect field data was increased using mixed models.

The results of this study indicate that other researchers that stratified their data and used different models for each stratum could potentially enhance their models with random effects on the level of these strata. An additional benefit would be that the amount of data for modeling is then larger.

Future work will strive to better understand the bias observed at the stand level. The stands, represented as polygons, on the German site were delineated based on operational considerations. Hence, small groups of trees were included with adjacent but different (in terms of species composition and age) stands to avoid creating small stands. We speculate that stand delineations that result in more homogeneous conditions within each stand will lead to lower within-stand variance and larger between-stand variance which could further improve the models. Furthermore it seems to be interesting to determine the contribution of the different lidar acquisition parameters (e.g. return density, foot print size) to the coefficients of the regression models. The use of mixed models can also reduce the number of sample plots needed for new study areas once a basic model exists. This issue will be discussed in an other study.

ACKNOWLEDGEMENTS

This paper is dedicated to Jürgen Bayer who died during the synthesis of this paper. He was a GIS-specialist at the Forstliche Versuchs- und Forschungsanstalt Baden-Württemberg and thus was involved in this research by consulting with questions regarding spatial data. We cannot express how much we miss him and his expertise.

We appreciate the comments of two unknown reviewers. We also would like to thank Dr. Edgar Kublin for his support regarding statistical analysis as well as Nick Vaughn and Nain Martinez for their general help. Special thanks go to Daniele Stöhr for the many years of good collaboration. Funds for this research were provided by the German Federal Ministry of Education and Research (0330616B). This document was typeset using LaTeX.

References

Aardt, J. A. N. v., Wynne, R. H. and Oderwald, R. G., 2006. Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *Forest Science* 52(6), pp. 636–649.

Andersen, H. E., McGaughey, R. J. and Reutebuch, S. E., 2005. Estimating forest canopy fuel parameters using lidar data. *Remote Sensing of Environment* 94(4), pp. 441–449.

Breidenbach, J., Koch, B., Kändler, G. and Kleusberg, A., 2007. Quantifying the influence of slope, aspect, crown shape and stem density on the estimation of tree height at plot level using lidar and insar data. *International Journal of Remote Sensing* (in review).

Curtis, R. O., Marshall, D. D. and DeBell, D. S. (eds), 2004. Silvicultural options for young-growth Douglas-fir forests: the Capitol Forest study-establishment and first results. *Gen. Tech. Rep. PNW-GTR-598*. U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, OR.

Kon-Allan, E., Goltz, v. H., Blust, M. and Nothdurft, A., 2004. *Manual for enterprise inventories*, version 1.1. Landesforstverwaltung Baden-Württemberg (German, title translated).

Lappi, J. and Bailey, R. L., 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. *Forest Science* 34(4), pp. 907–927.

McGaughey, R., Carson, W., Reutebuch, S. and Andersen, H.-E., 2004. Direct measurement of individual tree characteristics from lidar data. In: *Annual ASPRS Conference*, American Society of Photogrammetry and Remote Sensing, Denver, USA.

Means, J. E., Acker, S. A., Fitt, B. J., Renslow, M., Emerson, L. and Hendrix, C. J., 2000. Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering & Remote Sensing* 66(11), pp. 1367–1371.

Næsset, E., 2002. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment* 80(1), pp. 88– 99.

Næsset, E., 2004. Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research* 19(2), pp. 164 –179.

Packalén, P. and Maltamo, M., 2006. Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. *Forest Science* 52(6), pp. 611–622.

Pinheiro, J. and Bates, D., 2002. *Mixed Effects Models in S and S-Plus*. Springer.

R-Development-Core-Team, 2006. *R: A language and environment for statistical computing*. Technical report, R Foundation for Statistical Computing.

Reutebuch, S. E., McGaughey, R. J., Andersen, H.-E. and Carson, Ward, W., 2003. Accuracy of a high resolution lidar terrain model under a conifer forest canopy. *Canadian Journal of Remote Sensing* 29(5), pp. 527–535.

Weinacker, H., Koch, B., Heyder, U. and Weinacker, R., 2004. Development of filtering, segmentation and modelling modules for lidar and multispectral data as a fundament of an automatic forest inventory system. In: *ISPRS Working Group VIII/2 'Laser-scanners for forest and landscape assessment'*, University of Freiburg, Freiburg, Germany.