Application of mixed-effect model predictions in forestry

Lauri Mehtätalo[1]

[1] University of Eastern Finland, School of Computing & School of Forest Sciences

August 30, 2012

Types of forest datasets

- Forest datasets are usually hierarchical e.g.
  - needles within branches
  - branches within trees
  - **trees within sample plots**
  - sample plots within forest stands
  - forest stand within regions
  - repeated measurements of trees, branches etc.
  - ...

Types of forest datasets

- Forest datasets are usually hierarchical e.g.
  - needles within branches
  - branches within trees
  - **trees within sample plots**
  - sample plots within forest stands
  - forest stand within regions
  - repeated measurements of trees, branches etc.
  - ...
- Also crossed grouping structures are common
  - Tree increments for different calendar years
  - Trees or forest stands on aerial images

Types of forest datasets

- Forest datasets are usually hierarchical e.g.
  - needles within branches
  - branches within trees
  - **trees within sample plots**
  - sample plots within forest stands
  - forest stand within regions
  - repeated measurements of trees, branches etc.
  - ...
- Also crossed grouping structures are common
  - Tree increments for different calendar years
  - Trees or forest stands on aerial images

These datasets are naturally modeled using random effect models.

Why random effects?

- Using mixed-effects models with hierarchical datasets result in
  1. More reliable inference on the model parameters
  2. Possibilitý to compute the predictions at different levels of the dataset.

Why random effects?

- Using mixed-effects models with hierarchical datasets result in
  1. More reliable inference on the model parameters
  2. Possibilitý to compute the predictions at different levels of the dataset.
- If the main interest is the inference (e.g. the effects of certain medical treatments on individuals) the first property is more important.

Why random effects?

- Using mixed-effects models with hierarchical datasets result in
  1. More reliable inference on the model parameters
  2. Possibilitý to compute the predictions at different levels of the dataset.
- If the main interest is the inference (e.g. the effects of certain medical treatments on individuals) the first property is more important.
- If the main interest is prediction, then greatest benefit may arise from the possibility to make predictions at different levels of hierarchy. **This is possible also for observations from outside the modeling data**

Topic of this presentation

I will demonstrate and discuss the use of mixed-effects models in four forestry situations (The main benefit of mixed-effects models arsing either from prediction (P) or inference (I)).

- Using a previously fitted linear mixed-effects model for tree height prediction (P)

Topic of this presentation

I will demonstrate and discuss the use of mixed-effects models in four forestry situations (The main benefit of mixed-effects models arsing either from prediction (P) or inference (I)).

- Using a previously fitted linear mixed-effects model for tree height prediction (P)
- Using a linear mixed-effect model with crossed grouping structure to predict a treatment-free response in a dataset of a thinning experiment (P).

Topic of this presentation

I will demonstrate and discuss the use of mixed-effects models in four forestry situations (The main benefit of mixed-effects models arsing either from prediction (P) or inference (I)).

- Using a previously fitted linear mixed-effects model for tree height prediction (P)
- Using a linear mixed-effect model with crossed grouping structure to predict a treatment-free response in a dataset of a thinning experiment (P).
- Using nonlinear mixed-effect-models to analyse the previously extracted tree-level thinning effects (I)

Topic of this presentation

I will demonstrate and discuss the use of mixed-effects models in four forestry situations (The main benefit of mixed-effects models arsing either from prediction (P) or inference (I)).

- Using a previously fitted linear mixed-effects model for tree height prediction (P)
- Using a linear mixed-effect model with crossed grouping structure to predict a treatment-free response in a dataset of a thinning experiment (P).
- Using nonlinear mixed-effect-models to analyse the previously extracted tree-level thinning effects (I)
- Using a system of linear mixed-effects models to predict of tree level size and quality-related characteristics using remotely sensed information and field sample tree measurements (P)

Simple mixed-effects models

Let y_{ki} be the observed response for individual i in group k, and let x_{ki} be a fixed predictor, just as in the one-predictor regression. In a linear mixed-effects model, one may have both fixed (population level) parameters and random parameters, e.g.,

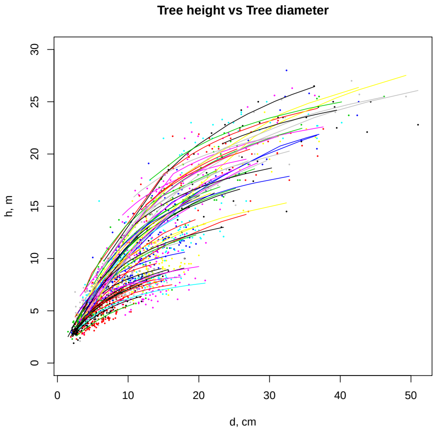
$$y_{ki} = a + bx_{ki} + \alpha_k + \beta_k x_{ki} + \epsilon_{ki} ,$$

where $(\alpha_k, \beta_k)' \sim MVN(0, \boldsymbol{D})$ and $\epsilon_{ki} \sim N(0, \sigma^2)$. a and b are the fixed parameters.

- The model allows population and group-level predictions: $\widehat{a} + \widehat{b}x_{ki}$ and $\widehat{a} + \widehat{\alpha_k} + (\widehat{b} + \widehat{\beta}_k)x_{ki}$ and corresponding residuals
- Generalizes to multiple groupings, with either nested or crossed grouring structures
- Generalizes to nonlinear models, where some parameters of a nonlinear function are allowed to vary between groups

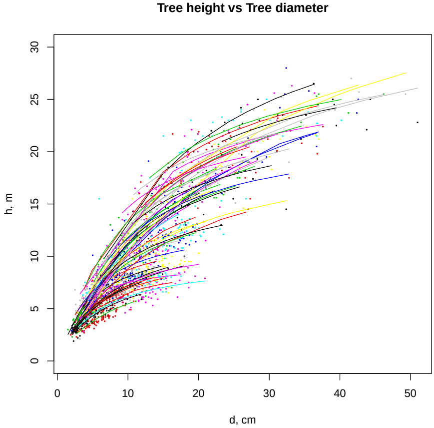
Why an H-D model?

- H-D relationship varies much among sample plots, but height measurement is time-consuming.



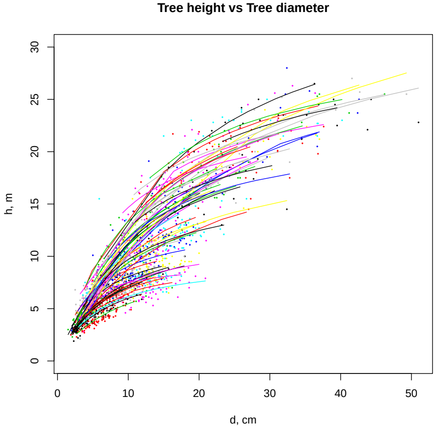
Why an H-D model?

- H-D relationship varies much among sample plots, but height measurement is time-consuming.
- In a forest inventory, diameter is usully tallied for all trees of a sample plot, whereas height is measured only for 0 – 5 trees per plot.



Why an H-D model?

- H-D relationship varies much among sample plots, but height measurement is time-consuming.
- In a forest inventory, diameter is usully tallied for all trees of a sample plot, whereas height is measured only for 0 – 5 trees per plot.



If a previously fitted H-D model is available, it can be localized, or calibrated, for the new plot by predicting the random effects using the sampled tree heights.

The Height-Diameter model

The logarithmic heigth H_{kti} for tree i in stand k at time t with diameter D_{kti} at the breast height is expressed by

$$\ln(H_{kti}) = a(DGM_{kt}) + \alpha_k + \alpha_{kt} + (b(DGM_{kt}) + \beta_k + \beta_{kt})D_{kti} + \epsilon_{kti} ,$$

where $a(DGM_{kt})$ and $b(DGM_{kt})$ are known fixed functions of plot-specific mean diameter DGM_{kt}, $(\alpha_k, \beta_k)'$ and $(\alpha_{kt}, \beta_{kt})'$ are the plot and measurement occasion -level random effects with varainces (correlations)

$$\mathrm{var}\begin{bmatrix}\alpha_k \\ \beta_k\end{bmatrix} = \begin{bmatrix} 0.108^2 & (0.269) \\ 0.0028 & 0.0958^2 \end{bmatrix} \qquad \mathrm{var}\begin{bmatrix}\alpha_{kt} \\ \beta_{kt}\end{bmatrix} = \begin{bmatrix} 0.0168^2 & (-0.681) \\ -0.0003 & 0.0223^2 \end{bmatrix}$$

and ϵ_{kti} are independent normal residuals with $\mathrm{var}(\epsilon_{kti}) = 0.401^2 \left(\max(D_{kti}, 7.5)\right)^{-1.068}$

The stand level mixed-effects model

The sample tree heights of a new stand can be described by

$$\boldsymbol{y} = \boldsymbol{\mu} + \boldsymbol{Z}\boldsymbol{b} + \boldsymbol{\epsilon}\,,$$

where
$\boldsymbol{y}$ includes the observed sample tree heights,
$\boldsymbol{\mu}$ is the fixed part,
$\boldsymbol{b} = \left(\alpha_k \quad \beta_k \quad \alpha_{k1} \quad \beta_{k1} \quad \alpha_{k2} \quad \beta_{k2} \quad \ldots \right)'$ includes the random effects,
$\boldsymbol{Z}$ is the corresponding design matrix, and
$\boldsymbol{\epsilon}$ includes the residuals.
We denote $\mathrm{var}(\boldsymbol{b}) = \boldsymbol{D}$ and $\mathrm{var}(\boldsymbol{\epsilon}) = \boldsymbol{R}$.

Prediction of random effects

The variances and covariances between random effects and observed heights can be written as

$$\begin{bmatrix} \boldsymbol{b} \\ \boldsymbol{y} \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{0} \\ \boldsymbol{\mu} \end{bmatrix}, \begin{bmatrix} \boldsymbol{D} & \boldsymbol{DZ'} \\ \boldsymbol{ZD} & \boldsymbol{ZDZ'} + \boldsymbol{R} \end{bmatrix} \right)$$

The Empirical Best Linear Unbiased Predictor (EBLUP) of random effects is

$$\widehat{\boldsymbol{b}} = \boldsymbol{DZ}'(\boldsymbol{ZDZ'} + \boldsymbol{R})^{-1}(\boldsymbol{y} - \boldsymbol{\mu}) .$$

and the variance of prediction errors is

$$\mathrm{var}(\widehat{\boldsymbol{b}} - \boldsymbol{b}) = \boldsymbol{D} - \boldsymbol{DZ}'(\boldsymbol{ZDZ}' + \boldsymbol{R})^{-1}\boldsymbol{ZD}$$

Example

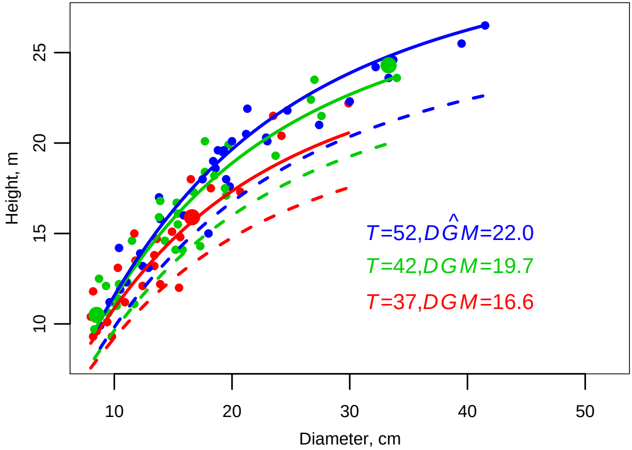
Height of one tree was measured 5 years ago and 2 trees at the current year. The matrices and vectors are

$$\boldsymbol{\mu} = \begin{bmatrix} 2.59 \\ 2.11 \\ 2.99 \end{bmatrix} \quad \boldsymbol{y} = \begin{bmatrix} 2.77 \\ 2.35 \\ 3.19 \end{bmatrix}$$

$$\boldsymbol{Z} = \begin{bmatrix} 1 & -0.36 & 1 & -0.36 & 0 & 0 \\ 1 & -1.22 & 0 & 0 & 1 & -1.22 \\ 1 & 0.058 & 0 & 0 & 1 & 0.058 \end{bmatrix} \quad \boldsymbol{R} = \begin{bmatrix} 0.008 & 0 & 0 \\ 0 & 0.016 & 0 \\ 0 & 0 & 0.004 \end{bmatrix}$$

$$\boldsymbol{b} = \begin{bmatrix} \alpha_k \\ \beta_k \\ \alpha_{k1} \\ \beta_{k1} \\ \alpha_{k2} \\ \beta_{k2} \end{bmatrix} \quad \boldsymbol{D} = \begin{bmatrix} 0.0118 & 0.0028 & 0 & 0 & 0 & 0 \\ 0.0028 & 0.0092 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0.0003 & 0.0004 & 0 & 0 \\ 0 & 0 & 0.0004 & 0.0005 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0.0003 & 0.0004 \\ 0 & 0 & 0 & 0 & 0.0004 & 0.0005 \end{bmatrix}$$

Uncalibrated and calibrated predictions



dashed=fixed part only; solid= calibrated (fixed+random)

Why thinning effects?

- Forest managers use silvicultural thinnings to decrease the competition of neighboring trees and, consequently, to increase the growth rate of the remaining trees for faster production of sawtimber.

Why thinning effects?

- Forest managers use silvicultural thinnings to decrease the competition of neighboring trees and, consequently, to increase the growth rate of the remaining trees for faster production of sawtimber.
- To understand the dynamics of thinning, one may wish to analyse the effect of thinnings on tree growth.

Why thinning effects?

- Forest managers use silvicultural thinnings to decrease the competition of neighboring trees and, consequently, to increase the growth rate of the remaining trees for faster production of sawtimber.
- To understand the dynamics of thinning, one may wish to analyse the effect of thinnings on tree growth.
- However, the growth is affected also by other factors, especially by the site productivity, tree age, and annual weather.

Why thinning effects?

- Forest managers use silvicultural thinnings to decrease the competition of neighboring trees and, consequently, to increase the growth rate of the remaining trees for faster production of sawtimber.
- To understand the dynamics of thinning, one may wish to analyse the effect of thinnings on tree growth.
- However, the growth is affected also by other factors, especially by the site productivity, tree age, and annual weather.
- Mixed-effects models can be used to model out these nuisance effects.

Study material

- Thinning experiment sample plots were established in naturally generated Scots pine stands at the age of $\sim$ 25 years in Mekrijärvi, Finland in 1986.

Study material

- Thinning experiment sample plots were established in naturally generated Scots pine stands at the age of $\sim$ 25 years in Mekrijärvi, Finland in 1986.
- One of the four following thinning treatments were applied to each plot: No thinning (I, Control), light (II), moderate (III), and heavy (IV) thinnings.

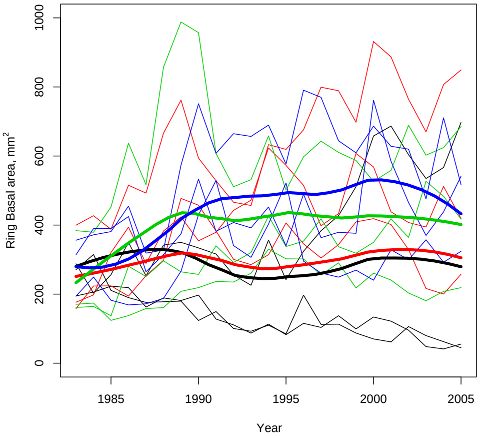
Study material

- Thinning experiment sample plots were established in naturally generated Scots pine stands at the age of $\sim$ 25 years in Mekrijärvi, Finland in 1986.
- One of the four following thinning treatments were applied to each plot: No thinning (I, Control), light (II), moderate (III), and heavy (IV) thinnings.
- 88 trees were felled in 2006, and the complete time series of diameter increments between 1983 and 2006 was measured for each tree using an X-ray densiometer.

Study material

- Thinning experiment sample plots were established in naturally generated Scots pine stands at the age of $\sim$ 25 years in Mekrijärvi, Finland in 1986.
- One of the four following thinning treatments were applied to each plot: No thinning (I, Control), light (II), moderate (III), and heavy (IV) thinnings.
- 88 trees were felled in 2006, and the complete time series of diameter increments between 1983 and 2006 was measured for each tree using an X-ray densiometer.
- The diameter growths were transformed to basal area growths, because $Volume \sim Diameter^2 Height$)

The raw data



- THICK: treatment-specific trends
- THIN: 12 randomly selected trees
- One can see
  - (Age trend)
  - climate-related year effects
  - tree effects

I (control) - black; II (light) - red
III (moderate) - green; IV (heavy) - blue

Modeling the non-thinned response

- A dataset without thinning treatments was produced by including from the original data
  - The control treatment for whole follow-up period
  - The thinned treatments until the year of thinning (1986)

Modeling the non-thinned response

- A dataset without thinning treatments was produced by including from the original data
  - The control treatment for whole follow-up period
  - The thinned treatments until the year of thinning (1986)
- A linear mixed effect model with random year and tree effects was fitted to the unthinned data

$$y_{kt} = f(T_{kt}) + u_k + v_t + e_{kt} \quad (1)$$

where y_{ckt} is the basal area growth of tree k at year t,
$f(T_{ckt})$ is the age trend (modeled using a spline),
u_k is a NID tree effect,
v_t is a NID year effect and
e_{kt} is a NID residual.

Extracting the thinning effects

- Using the estimated age trend and predicted year and tree effects, the growth without thinning, $\tilde{y}_{kt}$ was predicted for treatments II -IV after the thinning year.

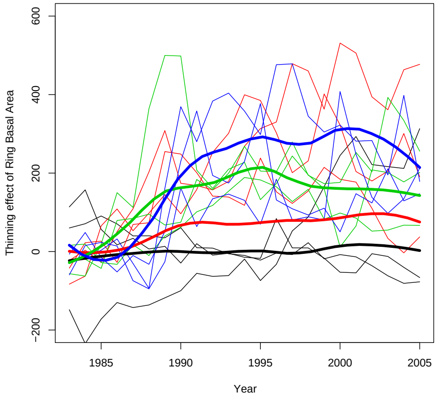
Extracting the thinning effects

- Using the estimated age trend and predicted year and tree effects, the growth without thinning, $\tilde{y}_{kt}$ was predicted for treatments II -IV after the thinning year.
- The pure thinning effects were estimated by subtracting the prediction from the observed growth

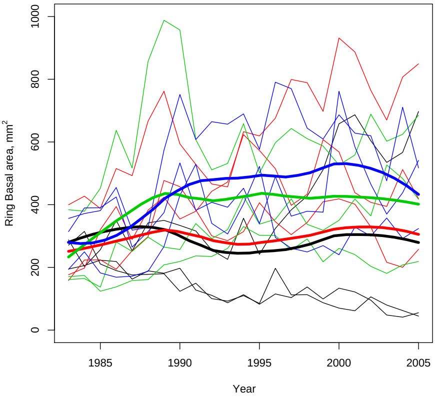
$$d_{kt} = y_{kt} - \tilde{y}_{kt} \quad (2)$$

The estimated thinning effects

Extracted thinning effects



Raw data



Line color specifies treatment (I:black, II: red, III: green IV: blue). Thick lines show the treatment-specific mean trends; thin lines show 12 randomly selected trees.

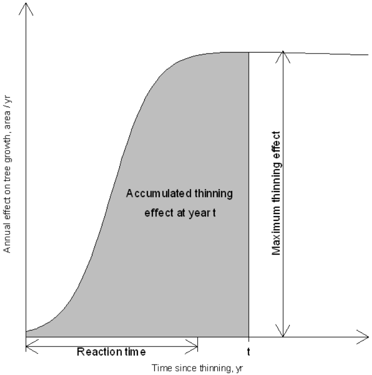
Modeling the thinning effects

- The thinning effects seem to switch on during a short time called **Reaction time** and stabilize thereafter at a level of **Maximum thinning effect**.
- To explore what predictors control these two parameters, the thinning effects of the thinnend treatments 2-4 were modeled using a nonlinear mixed-effects model.
- The random effects were used to take into account the data hierarchy for more reliable inference.

Nonlinear mixed-effects model for thinning effect

The thinning effect of tree *k* at time *t* was modeled using a logistic curve

$$d_{kt} = \frac{M_k}{1+\exp\left(4-8\frac{x_{kt}}{R_k}\right)} + e_{kt}$$



- d_{kt} - thinning effect
- x_{kt} - time since thinning
- $M_k = \mu_0 + \mu_1 T_2 + \mu_2 T_3 + \mu_4 x_{kt} + m_k$ - maximum thinning effect
- $T_2, \ldots, T_3$ - treatments
- $R_k = \rho_0 + \rho_1 z_k + r_k$ - reaction time
- z_k - standardized diameter
- $\begin{bmatrix} m_k \\ r_k \end{bmatrix} \sim MVN(\mathbf{0}, \boldsymbol{D}_{2x2})$
- e_{kt} - normal heteroscedastic residual with AR(1) structure within a tree.

The fitted model

- The reaction time was 6 years. It did not significantly vary among treatments but was shorter for large trees.

Fixed parameters	Estimate	s.e.	p-value
μ_0	112.8	23.29	0.0000
μ_1	91.91	30.45	0.0026
μ_2	169.2	32.14	0.0000
μ_3	-3.214	1.006	0.0014
ρ_0	5.749	0.4458	0.0000
ρ_1	-1.461	0.4568	0.0014
Random parameters			
var(r_k)	93.012		
var(m_k)	2.0852		
cor(r_k,m_k)	0.203		
Residual			
σ^2	8.157*10-4		
δ_1	8.746*104		
δ_2	1.886		
δ_3	0.5888		

The fitted model

- The reaction time was 6 years. It did not significantly vary among treatments but was shorter for large trees.
- The maximum thinning effect increased with thinning intensity, being 282 mm/yr for treatment IV, which indicates a 87% increase in the basal area growth compared to the control.

Fixed parameters	Estimate	s.e.	p-value
μ_0	112.8	23.29	0.0000
μ_1	91.91	30.45	0.0026
μ_2	169.2	32.14	0.0000
μ_3	-3.214	1.006	0.0014
ρ_0	5.749	0.4458	0.0000
ρ_1	-1.461	0.4568	0.0014
Random parameters			
var(r_k)	93.012		
var(m_k)	2.0852		
cor(r_k,m_k)	0.203		
Residual			
σ^2	8.157*10-4		
δ_1	8.746*104		
δ_2	1.886		
δ_3	0.5888		

Motivation

- Airborne Laser Scanners (ALS) provide information on the $3D$- structure of forest
- Majority of large individual trees can be detected from an ALS point cloud
- Point cloud characteristics can be assigned to field-measured tree characteristics to estimate a system of predictive models for tree characteristics, such as stem volume, height, diameter, crown base height, dead crown height.
- These tree-specific characteristics are correlated within a forest stand
- Also the stand effects are correlated across models
- These correlations can be utilized to predict the random effects of a mixed-effects model for a given stand for all 5 models using even one observation of one characteristics only
- Enables improved predictions of hard-to-measure characteristics by using easy-to-measure characteristics

The model

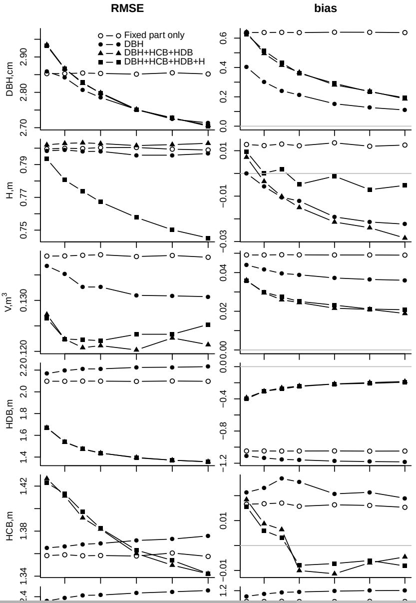
The model includes a system of 5 mixed-effects models of form for tree *i* in stand *k*:

$$
\begin{aligned}
y1_{ki} &= a1 + b1x1_{ki} + \ldots + \alpha1_k + \beta1_k x1_{ki} + \epsilon1_{ki} \\
y2_{ki} &= a2 + b1x2_{ki} + \ldots + \alpha2_k + \beta2_k x2_{ki} + \epsilon2_{ki} \\
&\vdots \\
y5_{ki} &= a5 + b5x5_{ki} + \ldots + \alpha5_k + \beta5_k x5_{ki} + \epsilon5_{ki}
\end{aligned}
$$

where the fixed parts are as with the previous mixed-effects models and include the ALS-based predictors.

- The assumptions on the random effects and residuals are $(\alpha1_k, \beta1_k, \alpha2_k, \beta2_k, \ldots, \alpha5_k, \beta5_k)' \sim MVN(0, \boldsymbol{D}_{10x10})$, and $(\epsilon1_{k1}, \epsilon2_{ki}, \ldots, \epsilon5_{ki}) \sim MVN(0, \boldsymbol{R}_{5x5})$
- The intended use of the model is prediction applying the random effects.
- The previously presented principles were used to predict the random effects of the model system by using 1-10 sample trees per stand and 3 different measurement strategies

Results


RMSE
bias
Fixed part only
DBH
DBH+HCB+HDB
DBH+HCB+HDB+H
DBH,cm
H,m
V,m3
HDB,m
HCB,m

Discussion, conclusions and references

- Mixed-effects models are useful tools for analyzing grouping datasets using models of various types.

References

- Mehtätalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. Can. J. For. Res. 34(1): 131-140.
- Mehtätalo, L., Peltola, H., Kilpeläinen, A. and Ikonen, V.-P. 2013. The effect of thinning on the basal area growth of Scots Pine: a longitudinal analysis using nonlinear mixed-effects model. Submitted manuscript.
- Maltamo, M., Mehtätalo, L., Vauhkonen, J. and Packalén P. 2012. Predicting and calibrating tree attributes by means of airborne laser scanning and field measurements. Can. J. For. Res. (In press)

Discussion, conclusions and references

- Mixed-effects models are useful tools for analyzing grouping datasets using models of various types.
- The prediction of random effects for a **new** group is a powerful tool to localize models afterwards using very limited datasets.

References

- Mehtätalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. Can. J. For. Res. 34(1): 131-140.
- Mehtätalo, L., Peltola, H., Kilpeläinen, A. and Ikonen, V.-P. 2013. The effect of thinning on the basal area growth of Scots Pine: a longitudinal analysis using nonlinear mixed-effects model. Submitted manuscript.
- Maltamo, M., Mehtätalo, L., Vauhkonen, J. and Packalén P. 2012. Predicting and calibrating tree attributes by means of airborne laser scanning and field measurements. Can. J. For. Res. (In press)

Discussion, conclusions and references

- Mixed-effects models are useful tools for analyzing grouping datasets using models of various types.
- The prediction of random effects for a **new** group is a powerful tool to localize models afterwards using very limited datasets.
- I am wondering if other fields than forestry have or could have similar applications.

References

- Mehtätalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. Can. J. For. Res. 34(1): 131-140.
- Mehtätalo, L., Peltola, H., Kilpeläinen, A. and Ikonen, V.-P. 2013. The effect of thinning on the basal area growth of Scots Pine: a longitudinal analysis using nonlinear mixed-effects model. Submitted manuscript.
- Maltamo, M., Mehtätalo, L., Vauhkonen, J. and Packalén P. 2012. Predicting and calibrating tree attributes by means of airborne laser scanning and field measurements. Can. J. For. Res. (In press)