Application of mixed-effect model predictions in forestry

Lauri Mehtätalo[1]

[1]UEF School of Forest Sciences

Nordstat conference, Umeå, Sweden, 11.05.2012

Outline

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.

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.

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, because estimation method takes into account the correlation resulting from the grouping.
  2. Possibilitý to compute the predictions at different levels of the dataset. In many forest applications, this means plot and population level predictions (i.e. predictions for an average plot).
- 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 arises from the possibility to make predictions at different levels of hierarchy.**

Why random effects?

- Using mixed-effects models with hierarchical datasets result in
  1. More reliable inference on the model parameters, because estimation method takes into account the correlation resulting from the grouping.
  2. Possibilitý to compute the predictions at different levels of the dataset. In many forest applications, this means plot and population level predictions (i.e. predictions for an average plot).
- 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 arises from the possibility to make predictions at different levels of hierarchy.

Why random effects?

- Using mixed-effects models with hierarchical datasets result in
  1. More reliable inference on the model parameters, because estimation method takes into account the correlation resulting from the grouping.
  2. Possibilitý to compute the predictions at different levels of the dataset. In many forest applications, this means plot and population level predictions (i.e. predictions for an average plot).
- 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 arises from the possibility to make predictions at different levels of hierarchy.**

Topic of this presentation

I will demonstrate and discuss the use of mixed-effects model predictions in two forestry situations

- In localizing a previously fitted mixed effects model **for a new stand from outside the modeling data but from the same population of stands** using measured response of one or more individuals of the new group.
- In prediction of a treatment-free response in a dataset of crossed crouping structure to extract a pure treatment effect.

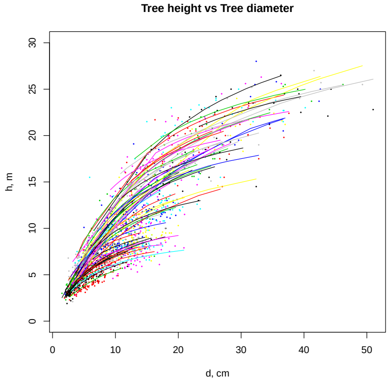
Topic of this presentation

I will demonstrate and discuss the use of mixed-effects model predictions in two forestry situations

- In localizing a previously fitted mixed effects model **for a new stand from outside the modeling data but from the same population of stands** using measured response of one or more individuals of the new group.
- In prediction of a treatment-free response in a dataset of crossed crouping structure to extract a pure treatment effect.

Why an H-D model?

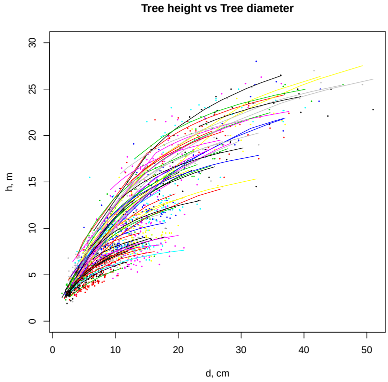
- H-D relationship varies much among sample plots, but height measurement is time-consuming.
- In a forest inventory, diameter is usully tallied for all trees of a sample plot, whereas height is measured only for 0 – 5 trees per plot.



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.

Why an H-D model?

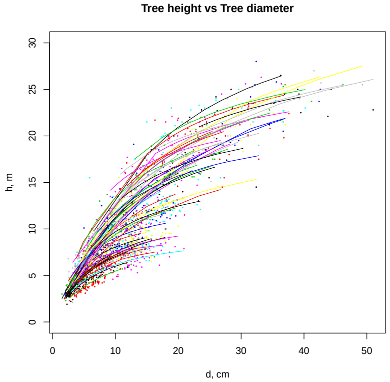
- H-D relationship varies much among sample plots, but height measurement is time-consuming.
- In a forest inventory, diameter is usully tallied for all trees of a sample plot, whereas height is measured only for 0 – 5 trees per plot.



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.

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)

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

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,
$\boldsymbol{b} = \begin{pmatrix} \alpha_k & \beta_k & \alpha_{k1} & \beta_{k1} & \alpha_{k2} & \beta_{k2} & \ldots \end{pmatrix}'$ 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

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

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

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

and the variance of prediction errors is

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

Example

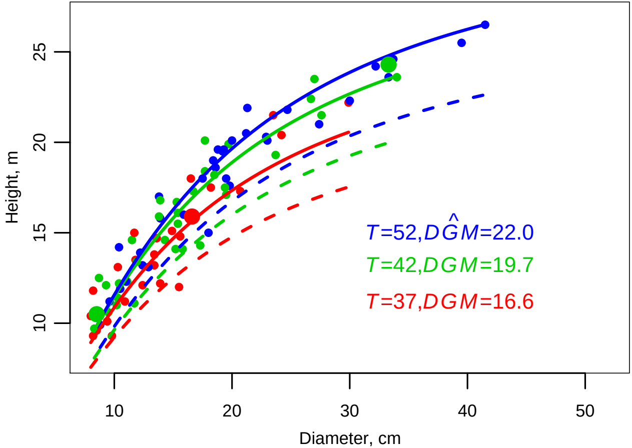
Height of one tree was measured 5 years ago and 2 trees at the current year. The matrices and vectors are

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

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

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

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.
- 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.

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.

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.

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.
- 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 $Volume \sim Diameter^2 Height$)

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 $Volume \sim Diameter^2 Height$)

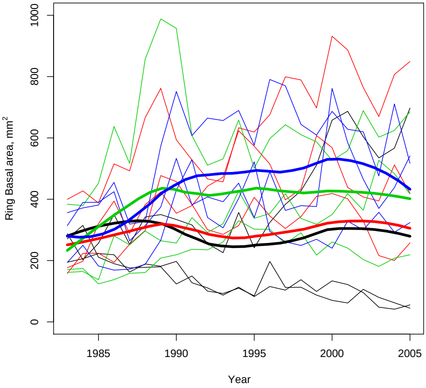
Study material

- 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 $Volume \sim Diameter^2 Height$)

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



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

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

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 restricted cubic spline with 3 knots),
u_k is a NID tree effect,
v_t is a NID year effect and
e_{kt} is a NID residual.

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 restricted cubic spline with 3 knots),
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.
- The pure thinning effects were estimated by subtracting the prediction from the observed growth

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

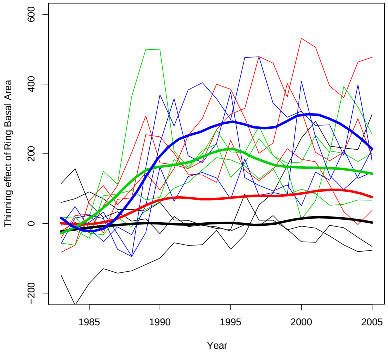
Extracting the thinning effects

- 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

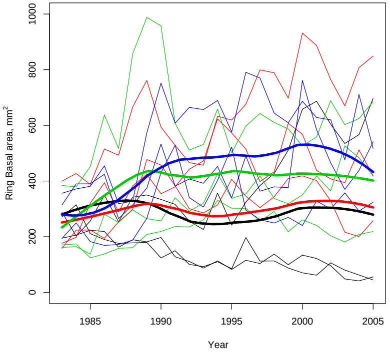
$$d_{kt} = y_{kt} - \tilde{y}_{kt} \tag{2}$$

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.

Discussion and conclusions

- The prediction of random effects for a **new** group is a powerful tool to localize models afterwards using very limited datasets.
- Since the proposal of this approach for taper curves (Lappi 1986), numerous forstry applications have been published, including
  - Site index curves (starting from Lappi and Bailey 1988)
  - Height-Diameter curves (Lappi 1997, Mehtätalo 2004, 2005a)
  - Diameter distributions (Mehtätalo 2005b, Mehtätalo et al. 2011)
  - Cross-calibration of seemingly unrealted mixed-effects models (Lappi 1991, Lappi et al 2006)
- I am wondering if other fields than forestry have or could have similar applications.

Discussion and conclusions

- The prediction of random effects for a **new** group is a powerful tool to localize models afterwards using very limited datasets.
- Since the proposal of this approach for taper curves (Lappi 1986), numerous forstry applications have been published, including
  - Site index curves (starting from Lappi and Bailey 1988)
  - Height-Diameter curves (Lappi 1997, Mehtätalo 2004, 2005a)
  - Diameter distributions (Mehtätalo 2005b, Mehtätalo et al. 2011)
  - Cross-calibration of seemingly unrealted mixed-effects models (Lappi 1991, Lappi et al 2006)
- I am wondering if other fields than forestry have or could have similar applications.

Discussion and conclusions

- The prediction of random effects for a **new** group is a powerful tool to localize models afterwards using very limited datasets.
- Since the proposal of this approach for taper curves (Lappi 1986), numerous forstry applications have been published, including
  - Site index curves (starting from Lappi and Bailey 1988)
  - Height-Diameter curves (Lappi 1997, Mehtätalo 2004, 2005a)
  - Diameter distributions (Mehtätalo 2005b, Mehtätalo et al. 2011)
  - Cross-calibration of seemingly unrealted mixed-effects models (Lappi 1991, Lappi et al 2006)
- I am wondering if other fields than forestry have or could have similar applications.

Discussion and conclusions

- In the other study, mixed-effects models provided an useful tool to extract the thinning effects by removing the nuisance effects caused by age trend, and random age and year effects.
- A somewhat similar analysis has previously been used for extracting year effects and exploring their correlation with climatic records (Gort et al 2011, Zubizarreta et al 2012).
- The extracted thinning effects can be further modeled using nonlinear mixed-effects models (Mehtätalo et al 2013).
- Forestry datasets are often hierarchical, and analyses of such datasets can significantly benefit from the use of mixed-model predictions. As an example, this presetation showed two examples, but the applicationa are not restricted to these.

Discussion and conclusions

- In the other study, mixed-effects models provided an useful tool to extract the thinning effects by removing the nuisance effects caused by age trend, and random age and year effects.
- A somewhat similar analysis has previously been used for extracting year effects and exploring their correlation with climatic records (Gort et al 2011, Zubizarreta et al 2012).
- The extracted thinning effects can be further modeled using nonlinear mixed-effects models (Mehtätalo et al 2013).
- Forestry datasets are often hierarchical, and analyses of such datasets can significantly benefit from the use of mixed-model predictions. As an example, this presetation showed two examples, but the applicationa are not restricted to these.

Discussion and conclusions

- In the other study, mixed-effects models provided an useful tool to extract the thinning effects by removing the nuisance effects caused by age trend, and random age and year effects.
- A somewhat similar analysis has previously been used for extracting year effects and exploring their correlation with climatic records (Gort et al 2011, Zubizarreta et al 2012).
- The extracted thinning effects can be further modeled using nonlinear mixed-effects models (Mehtätalo et al 2013).
- Forestry datasets are often hierarchical, and analyses of such datasets can significantly benefit from the use of mixed-model predictions. As an example, this presetation showed two examples, but the applicationa are not restricted to these.

Discussion and conclusions

- In the other study, mixed-effects models provided an useful tool to extract the thinning effects by removing the nuisance effects caused by age trend, and random age and year effects.
- A somewhat similar analysis has previously been used for extracting year effects and exploring their correlation with climatic records (Gort et al 2011, Zubizarreta et al 2012).
- The extracted thinning effects can be further modeled using nonlinear mixed-effects models (Mehtätalo et al 2013).
- Forestry datasets are often hierarchical, and analyses of such datasets can significantly benefit from the use of mixed-model predictions. As an example, this presetation showed two examples, but the applicationa are not restricted to these.

References

- Lappi, J. 1986. Mixed linear models for analyzing and predicting stem form variation of Scots pine. Communicationes Instituti Forestalis Fenniae 134. 69 p.
- Lappi, J. and Bailey, R. L. 1988. A height prediction model with random stand and tree parameters: an alternative to traditional site index methods. For. Sci 34: 907–927.
- Lappi, J. 1991. Calibration of height and volume equations with random parameters. For. Sci. 37(3): 781-801.
- Lappi, J. 1997. A longitudinal analysis of height/diameter curves. For. Sci. 43. 555–570.
- Lappi, J., Mehtätalo, L. and Korhonen, K.T. 2006. Generalizing sample tree information. in Kangas, A. and Maltamo, M (editors) Forest inventory - methodology & applications. Springer.
- Mehtätalo, L. 2004. A longitudinal height-diameter model for Norway spruce in Finland. Can. J. For. Res. 34(1): 131-140.
- Mehtätalo, L. 2005a. Height-diameter models for Scots pine and birch in Finland. Silva Fennica 39(1): 55-66.
- Mehtätalo, L. 2005b. Localizing a Predicted Diameter Distribution Using Sample Information For. Sci. 51(4):292Ű303.
- Mehtätalo, L., Comas, C., Pukkala, T., and Palahi, M. 2011. Combining a predicted diameter distribution with an estimate based on a small sample of diameters. Can. J. For. Res. 41:750-762.
- Mehtätalo, L., Peltola, H., Antti 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.
- Zubizarreta-Gerendiain, A., Gort, J., Mehtätalo, L., Peltola, H., Venäläinen, A., and Pulkkinen, P. 2012. Effects of cambial age, clone and climatic factors on ring width and ring density in Norway spruce (Picea abies) in southeastern Finland. Forest Ecology and Management Volume 263:9-16.
- Gort, J., Mehtätalo, L., Peltola, H., Zubizarreta-Gerendian, A., Pulkkinen, P. and Venäläinen, A. 2011. Effects of spacing and genetic entry on radial growth and ring density development in Scots pine (Pinus sylvestris L.). Annals of Forest Science 68(7): 1233-1243.