Predicting stand structure using limited measurements

Lauri Mehtätalo
lauri.mehtatalo@yale.edu

June 12, 2006,
University of Minnesota
Department of Forest Resources

Outline of the presentation

Introduction

Methods
- Linear Prediction
- Examples of BLP

Predicting the distribution of diameters
- Percentile-based diameter distribution
- Distributions of order statistics
- BLP in this case
- The effect of order statistics on prediction accuracy

Discussion

The inventory of a single stand

The inventory of a single stand

- For forestry decision making, the most important information from a stand is the amount and structure of the growing stock

The inventory of a single stand

- For forestry decision making, the most important information from a stand is the amount and structure of the growing stock
- In this study, these are described by the total basal area, diameter distribution and height-diameter curve, each of which is expressed by tree species.

The inventory of a single stand

- For forestry decision making, the most important information from a stand is the amount and structure of the growing stock
- In this study, these are described by the total basal area, diameter distribution and height-diameter curve, each of which is expressed by tree species.
- The main aim of this study was to develop tools for producing a stand description from ground-measured data

The inventory of a single stand

- For forestry decision making, the most important information from a stand is the amount and structure of the growing stock
- In this study, these are described by the total basal area, diameter distribution and height-diameter curve, each of which is expressed by tree species.
- The main aim of this study was to develop tools for producing a stand description from ground-measured data
- Because of limited measurement resources for a single stand, the available information is very limited and the use of measurement information should be as effective as possible.

We have two kinds of measurements from a stand

We have two kinds of measurements from a stand

Group 1: Stand variables

- Some basic information (stand age, site fertility class ...)
- Basal area and basal area weighted median diameter (*DGM*) from angle count plot(s)

Can be used as predictors in the regression models of stand structure.

We have two kinds of measurements from a stand

Group 1: Stand variables

- Some basic information (stand age, site fertility class ...)
- Basal area and basal area weighted median diameter (*DGM*) from angle count plot(s)

Can be used as predictors in the regression models of stand structure.

Group2: Sample measurements (of response)

- Sample tree height(s)
- Sample order statistic(s) of the angle count sample plot(s), called quantile trees.

The use is based on **linear prediction**.

Linear prediction of random variables

Suppose that a random vector $\mathbf{x}$ is partitioned into unobserved and observed parts, $\mathbf{x}_1$ and $\mathbf{x}_2$.

Linear prediction of random variables

Suppose that a random vector $\mathbf{x}$ is partitioned into unobserved and observed parts, $\mathbf{x}_1$ and $\mathbf{x}_2$.
Suppose that that the first and second order properties of $\mathbf{x}$ are

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} \\ \mathbf{V}_{12}' & \mathbf{V}_2 \end{bmatrix} \right),$$

where everything except for $\mathbf{x}_1$ are known.

Linear prediction of random variables

Suppose that a random vector $\mathbf{x}$ is partitioned into unobserved and observed parts, $\mathbf{x}_1$ and $\mathbf{x}_2$.

Suppose that that the first and second order properties of $\mathbf{x}$ are

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} \\ \mathbf{V}_{12}' & \mathbf{V}_2 \end{bmatrix} \right),$$

where everything except for $\mathbf{x}_1$ are known.

The Best Linear Predictor (BLP) of $\mathbf{x}_1$ is

$$\widehat{\mathbf{x}_1} = \boldsymbol{\mu}_1 + \mathbf{V}_{12}\mathbf{V}_2^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$$

and the variance of prediction errors is

$$\mathrm{var}(\widehat{\mathbf{x}_1} - \mathbf{x}_1) = \mathbf{V}_1 - \mathbf{V}_{12}\mathbf{V}_2^{-1}\mathbf{V}_{12}'.$$

Linear prediction of random variables

Suppose that a random vector $\mathbf{x}$ is partitioned into unobserved and observed parts, $\mathbf{x}_1$ and $\mathbf{x}_2$.
Suppose that that the first and second order properties of $\mathbf{x}$ are

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} \\ \mathbf{V}_{12}' & \mathbf{V}_2 \end{bmatrix} \right) ,$$

where everything except for $\mathbf{x}_1$ are known.
The Best Linear Predictor (BLP) of $\mathbf{x}_1$ is

$$\widehat{\mathbf{x}_1} = \boldsymbol{\mu}_1 + \mathbf{V}_{12}\mathbf{V}_2^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$$

and the variance of prediction errors is

$$\mathrm{var}(\widehat{\mathbf{x}_1} - \mathbf{x}_1) = \mathbf{V}_1 - \mathbf{V}_{12}\mathbf{V}_2^{-1}\mathbf{V}_{12}' .$$

Under normality, BLP is the Best Predictor (BP)

Linear prediction of random variables

Suppose that a random vector $\mathbf{x}$ is partitioned into unobserved and observed parts, $\mathbf{x}_1$ and $\mathbf{x}_2$.

Suppose that that the first and second order properties of $\mathbf{x}$ are

$$\begin{bmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{bmatrix} \sim \left(\begin{bmatrix} \boldsymbol{\mu}_1 \\ \boldsymbol{\mu}_2 \end{bmatrix}, \begin{bmatrix} \mathbf{V}_1 & \mathbf{V}_{12} \\ \mathbf{V}_{12}' & \mathbf{V}_2 \end{bmatrix} \right),$$

where everything except for $\mathbf{x}_1$ are known.

The Best Linear Predictor (BLP) of $\mathbf{x}_1$ is

$$\widehat{\mathbf{x}_1} = \boldsymbol{\mu}_1 + \mathbf{V}_{12}\mathbf{V}_2^{-1}(\mathbf{x}_2 - \boldsymbol{\mu}_2)$$

and the variance of prediction errors is

$$\mathrm{var}(\widehat{\mathbf{x}_1} - \mathbf{x}_1) = \mathbf{V}_1 - \mathbf{V}_{12}\mathbf{V}_2^{-1}\mathbf{V}_{12}'.$$

Under normality, BLP is the Best Predictor (BP)

In applications, we usually use estimated $\boldsymbol{\mu}$'s and $\mathbf{V}$'s to obtain an Estimated BLP (EBLP).

Examples of BLP

Examples of BLP

- In a mixed model, we predict a vector of random effects, $\mathbf{x}_1$, using a vector of observed responses, $\mathbf{x}_2$.

Examples of BLP

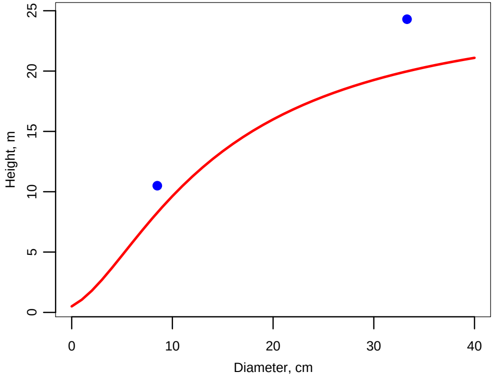
- In a mixed model, we predict a vector of random effects, $\mathbf{x}_1$, using a vector of observed responses, $\mathbf{x}_2$.
- If we have two models with correlated residuals, and the response for the first model is measured ($\mathbf{x}_2$), the response of the other model can be predicted as $\mathbf{x}_1$. For example, we can predict volumes of sample trees with known heights.

Examples of BLP

- In a mixed model, we predict a vector of random effects, $\mathbf{x}_1$, using a vector of observed responses, $\mathbf{x}_2$.
- If we have two models with correlated residuals, and the response for the first model is measured ($\mathbf{x}_2$), the response of the other model can be predicted as $\mathbf{x}_1$. For example, we can predict volumes of sample trees with known heights.
- In spatial statistics, kriging is a method where unobserved responses, $\mathbf{x}_1$ at spatial locatiobns are predicted using observed responses, $\mathbf{x}_2$ at spatial locations. Matrices $\mathbf{V}$ are obtained from an estimated variogram.

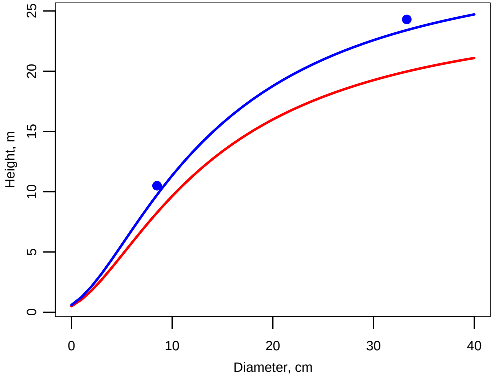
Predicting Height-Diameter curve

The fixed part of a mixed model gives the expected height for a tree with a given diameter. If we have sample tree heights measured, we can predict the random effects of the model ($\mathbf{x}_1$) using sample tree heights ($\mathbf{x}_2$)



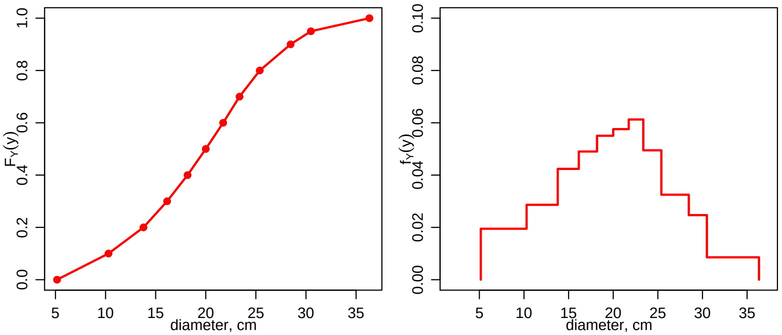
Predicting Height-Diameter curve

The fixed part of a mixed model gives the expected height for a tree with a given diameter. If we have sample tree heights measured, we can predict the random effects of the model ($\mathbf{x}_1$) using sample tree heights ($\mathbf{x}_2$)



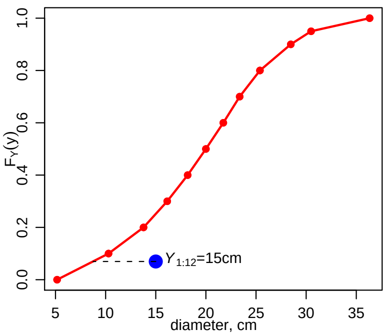
Predicting the diameter distribution

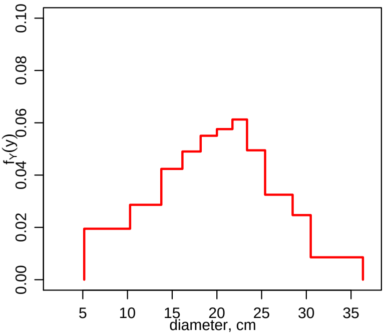
- I predict diameter percentiles using models that are estimated from some data *a priori*.



Predicting the diameter distribution

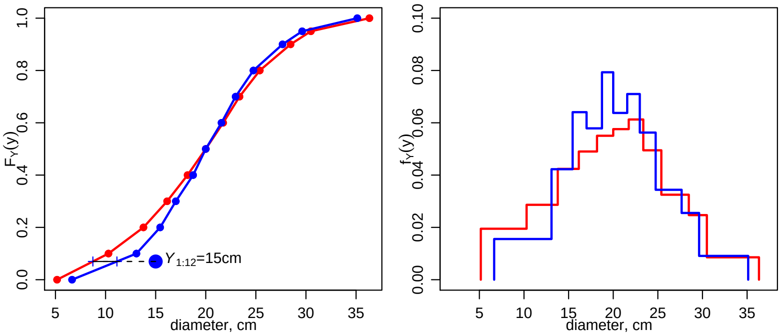
- I predict diameter percentiles using models that are estimated from some data *a priori*.
- The quantile tree is interpreted as a measured percentile,





Predicting the diameter distribution

- I predict diameter percentiles using models that are estimated from some data *a priori*.
- The quantile tree is interpreted as a measured percentile,
- so we can use it ($\mathbf{x}_2$) to predict the residuals of our percentile models ($\mathbf{x}_1$).



Percentile-based diameter distribution

Certain percentiles in stand k are modeled using Group 1 variables

$$
\begin{aligned}
d_{0\%,k} &= E(d0\%|DGM_k, G_k, T_k, \ldots) + e_{0\%,k} \\
d_{10\%,k} &= E(d10\%|DGM_k, G_k, T_k, \ldots) + e_{10\%,k} \\
&\vdots \\
d_{90\%,k} &= E(d90\%|DGM_k, G_k, T_k, \ldots) + e_{90\%,k} \\
d_{100\%,k} &= E(d100\%|DGM_k, G_k, T_k, \ldots) + e_{100\%,k}
\end{aligned}
$$

The continuous distribution function is obtained by linear interpolation.

Percentile-based diameter distribution

Certain percentiles in stand k are modeled using Group 1 variables

$$\begin{aligned}
d_{0\%,k} &= E(d0\%|DGM_k, G_k, T_k, \ldots) + e_{0\%,k} \\
d_{10\%,k} &= E(d10\%|DGM_k, G_k, T_k, \ldots) + e_{10\%,k} \\
&\vdots \\
d_{90\%,k} &= E(d90\%|DGM_k, G_k, T_k, \ldots) + e_{90\%,k} \\
d_{100\%,k} &= E(d100\%|DGM_k, G_k, T_k, \ldots) + e_{100\%,k}
\end{aligned}$$

The continuous distribution function is obtained by linear interpolation.

The same in matrix form:

$$\mathbf{d}_k = E(\mathbf{d}|DGM_k, G_k, T_k, \ldots) + \mathbf{e}_k \,,$$

where residuals $\mathbf{e}_k$ are actually stand effects with nondiagonal $\mathrm{var}(\mathbf{e}_k) = \mathbf{D}$. We will drop index k hereafter.

Sample order statistics

Let $Y_{r:n}$ be r^{th} smallest observation in a sample of size n from population with distribution $F_Y(y)$ and density $f_Y(y)$.

Sample order statistics

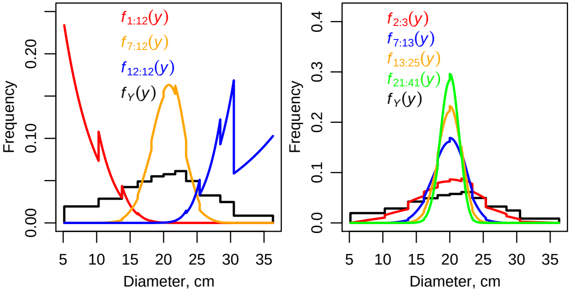
Let $Y_{r:n}$ be r^{th} smallest observation in a sample of size n from population with distribution $F_Y(y)$ and density $f_Y(y)$.
$Y_{r:n}$ is a random variable, which has density

$$f_{r:n}(y) = \frac{n!}{(r-1)!(n-r)!} f_Y(y)[F_Y(y)]^{r-1}[1 - F_Y(y)]^{n-r}$$

Sample order statistics

Let $Y_{r:n}$ be r^{th} smallest observation in a sample of size n from population with distribution $F_Y(y)$ and density $f_Y(y)$.
$Y_{r:n}$ is a random variable, which has density

$$f_{r:n}(y) = \frac{n!}{(r-1)!(n-r)!} f_Y(y)[F_Y(y)]^{r-1}[1 - F_Y(y)]^{n-r}$$

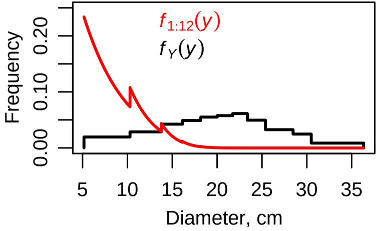


Examples of distributions, black line is the underlying population distribution

Expectation and variance of $Y_{r:n}$

Recall the example, where $Y_{1:12}$ was observed

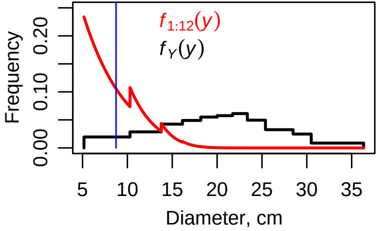
Assuming that the predicted $f_Y(y)$ is the diameter distribution, we know the distribution of $Y_{1:12}$.



Expectation and variance of $Y_{r:n}$

Recall the example, where $Y_{1:12}$ was observed

Assuming that the predicted $f_Y(y)$ is the diameter distribution, we know the distribution of $Y_{1:12}$.
Further we calculate
$E(Y_{1:12}) = 8.73\text{cm}$ and
$\text{var}(Y_{1:12}) = 8.09$

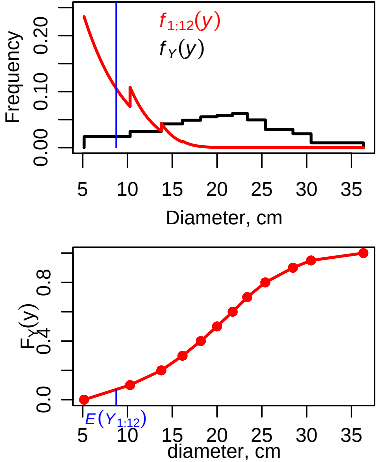


Expectation and variance of $Y_{r:n}$

Recall the example, where $Y_{1:12}$ was observed

Assuming that the predicted $f_Y(y)$ is the diameter distribution, we know the distribution of $Y_{1:12}$.
Further we calculate
$E(Y_{1:12}) = 8.73\mathrm{cm}$ and
$\mathrm{var}(Y_{1:12}) = 8.09$

Based on $p = F_Y(E(Y_{1:12})) = 0.070$, $Y_{1:12}$ is interpreted as a measurement of 7th percentile of distribution $F_Y(y)$

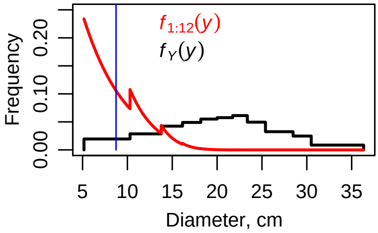


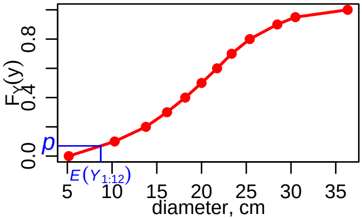
Expectation and variance of $Y_{r:n}$

Recall the example, where $Y_{1:12}$ was observed

Assuming that the predicted $f_Y(y)$ is the diameter distribution, we know the distribution of $Y_{1:12}$.
Further we calculate
$E(Y_{1:12}) = 8.73\text{cm}$ and
$\text{var}(Y_{1:12}) = 8.09$

Based on $p = F_Y(E(Y_{1:12})) = 0.070$, $Y_{1:12}$ is interpreted as a measurement of 7th percentile of distribution $F_Y(y)$





The model for the measured percentile

Recall the model for the predefined percentiles

$$\mathbf{d}_k = E(\mathbf{d}|DGM, G, T, \ldots) + \mathbf{e}_k \,,$$

where $\mathrm{var}(\mathbf{e}_k) = \mathbf{D}$.
The model for this observation is

$$d_{7\%} + \epsilon = E(d_{7\%}|DGM, G, T, \ldots) + e_{7\%} + \epsilon \,,$$

where ϵ is the measurement error of of the percentile, with $\mathrm{var}(\epsilon) = \mathrm{var}(Y_{1:12}) = 8.09$.
We approximate $E(d_{7\%}|DGM, G, T, \ldots)$, $\mathrm{var}(e_{7\%})$ and $\mathrm{cov}((e), e_{7\%})$ by interpolating linearly $E(\mathbf{d}|DGM, G, T, \ldots)$ and $\mathbf{D}$ for the 7th percentile.

The model for the measured percentile

Recall the model for the predefined percentiles

$$\mathbf{d}_k = E(\mathbf{d}|DGM, G, T, \ldots) + \mathbf{e}_k\,,$$

where $\mathrm{var}(\mathbf{e}_k) = \mathbf{D}$.
The model for this observation is

$$d_{7\%} + \epsilon = E(d_{7\%}|DGM, G, T, \ldots) + e_{7\%} + \epsilon\,,$$

where ϵ is the measurement error of of the percentile, with $\mathrm{var}(\epsilon) = \mathrm{var}(Y_{1:12}) = 8.09$.
We approximate $E(d_{7\%}|DGM, G, T, \ldots)$, $\mathrm{var}(e_{7\%})$ and $\mathrm{cov}((e), e_{7\%})$ by interpolating linearly $E(\mathbf{d}|DGM, G, T, \ldots)$ and $\mathbf{D}$ for the 7th percentile.
We can now write

$$\begin{bmatrix} \mathbf{d} \\ d_{7\%} + \epsilon \end{bmatrix} \sim \left(\begin{bmatrix} E(\mathbf{d}|DGM, G, T, \ldots) \\ E(d_{7\%}|DGM, G, T, \ldots) \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathrm{cov}(\mathbf{e}, e_{7\%}) \\ \mathrm{cov}(e_{7\%}, \mathbf{e}) & \mathrm{var}(e_{7\%}) + \mathrm{var}(\epsilon) \end{bmatrix} \right)\,,$$

where all except for $\mathbf{d}$ is known.

BLP of stand effects

$$\begin{bmatrix} \mathbf{d} \\ d_{7\%} + \epsilon \end{bmatrix} \sim \left(\begin{bmatrix} E(\mathbf{d}|DGM, G, T, \ldots) \\ E(d_{7\%}|DGM, G, T, \ldots) \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathrm{cov}(\mathbf{e}, e_{7\%}) \\ \mathrm{cov}(e_{7\%}, \mathbf{e}) & \mathrm{var}(e_{7\%}) + \mathrm{var}(\epsilon) \end{bmatrix} \right),$$

The BLP is calculated as

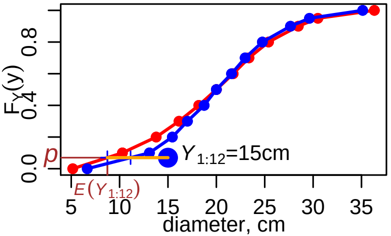
$$\begin{aligned} \widehat{\mathbf{d}} &= E(\mathbf{d}|DGM, G, T, \ldots) + \\ & \mathrm{cov}(\mathbf{e}, e_{7\%})(\mathrm{var}(e_{7\%}) + \mathrm{var}(\epsilon))^{-1}(d_{7\%} - E(d_{7\%}|DGM, G, T, \ldots)) \end{aligned}$$

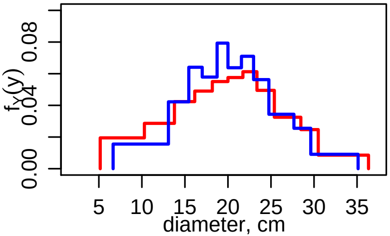
BLP of stand effects

$$\begin{bmatrix} \mathbf{d} \\ d_{7\%} + \epsilon \end{bmatrix} \sim \left(\begin{bmatrix} E(\mathbf{d}|DGM, G, T, \ldots) \\ E(d_{7\%}|DGM, G, T, \ldots) \end{bmatrix}, \begin{bmatrix} \mathbf{D} & \mathrm{cov}(\mathbf{e}, e_{7\%}) \\ \mathrm{cov}(e_{7\%}, \mathbf{e}) & \mathrm{var}(e_{7\%}) + \mathrm{var}(\epsilon) \end{bmatrix} \right),$$

The BLP is calculated as

$$\begin{aligned} \hat{\mathbf{d}} &= E(\mathbf{d}|DGM, G, T, \ldots) + \\ &\quad \mathrm{cov}(\mathbf{e}, e_{7\%})(\mathrm{var}(e_{7\%}) + \mathrm{var}(\epsilon))^{-1}(d_{7\%} - E(d_{7\%}|DGM, G, T, \ldots)) \end{aligned}$$





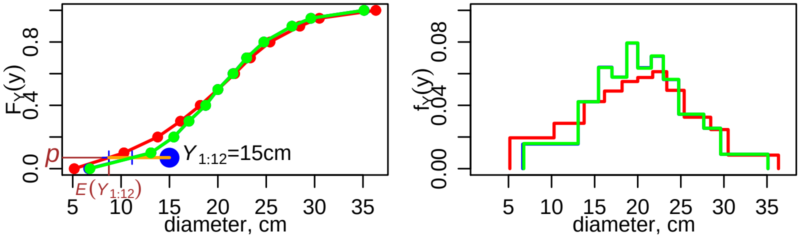
Oops

- Our $f_{r:n}(y)$ was based on assumption that $E(\mathbf{d}|DGM, G, T, \ldots)$ gives the *true* distribution of the stand.
- Now we have a better estimate for the true distribution, namely $\widehat{\mathbf{d}}$.
- That is why we use $\widehat{\mathbf{d}}$ to calculate p and $\mathrm{var}(Y_{r:n})$ again, and predict new $\widehat{\mathbf{d}}$.

Oops

- Our $f_{r:n}(y)$ was based on assumption that $E(\mathbf{d}|DGM, G, T, \ldots)$ gives the *true* distribution of the stand.
- Now we have a better estimate for the true distribution, namely $\widehat{\mathbf{d}}$.
- That is why we use $\widehat{\mathbf{d}}$ to calculate p and $\mathrm{var}(Y_{r:n})$ again, and predict new $\widehat{\mathbf{d}}$.

After 7 iterations the final values were $p = 0.064$, $E(Y_{1:12}) = 10.82\mathrm{cm}$ and $\mathrm{var}(Y_{r:n}) = 8.28$.



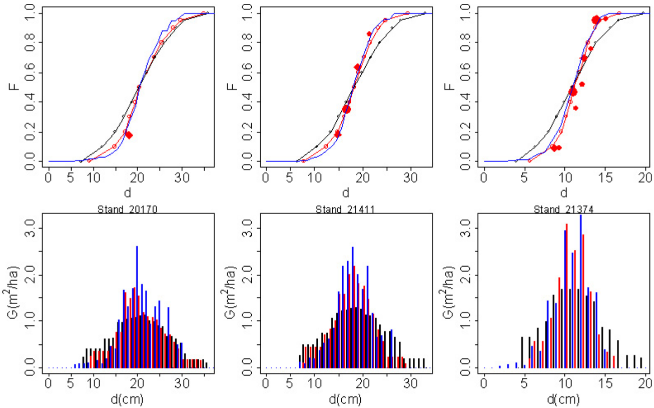
Several observations per stand

- If all measurements are from different plots, their "measurement errors" (ϵ's) are uncorrelated.

Several observations per stand

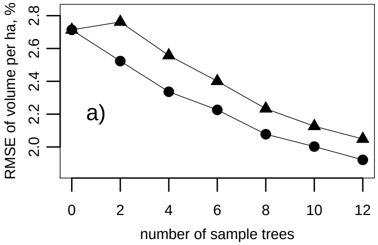
- If all measurements are from different plots, their "measurement errors" (ϵ's) are uncorrelated.
- In the case of several observation per plot, ϵ's are correlated. Joint distribution of two order statistics is needed to calculate the correlation:

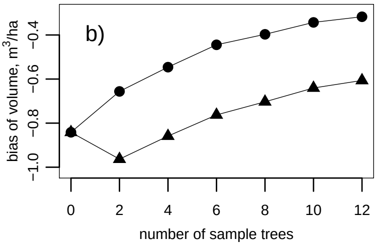
Examples of predictions (True trees from true plots)



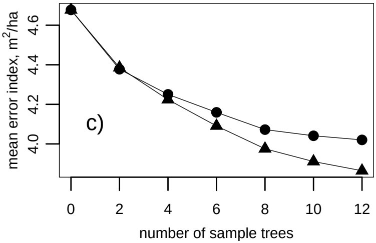
True distribution, prediction based on *DGM*, *G*, *T*, *soil* and the prediction based on *DGM*, *G*, *T*, *soil* and sample order statistics (the marks). Same mark is used for measurements from same plot.

True trees from simulated plots



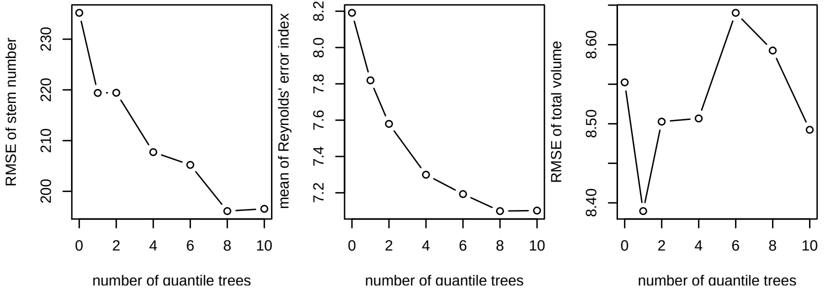


- Order statistics are taken from *random samples* of Norway spruce stands.
- Triangles show the prediction after first step, the circle after converged iteration.

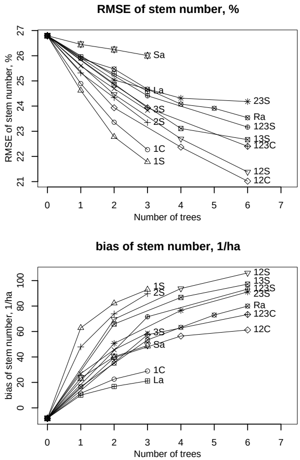


True trees from true plots

- Trees were randomly selected from 3 angle count sample plots
- Stand variables (DGM, G, N) included measurement errors
- Observations in a sample seem not to be independent observations from the population



Which quantile to measure?



Quantile trees were selected according to different strategies from 1-3 angle count plots

- 1S, 2S, 3S: 1st, 2nd and 3rd smallest tree of each plot
- 12S, 13S, 23S: 1st and 2nd, 1st and 3rd, and 2nd and 3rd smallest trees of each plot
- 1C, 2C, 3C: 1st, 2nd and 3rd closest tree of each plot
- 12C, 13C, 23C: 1st and 2nd, 1st and 3rd, and 2nd and 3rd closest trees of each plot
- Sa: Tree closest to sawtimber limit (17 cm in diameter)
- La: Largest tree of each plots
- Ra: Randomly selected trees

Discussion

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs
- Which quantiles are the easiest to measure?

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs
- Which quantiles are the easiest to measure?
- Using extremes seems to cause a bias problem

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs
- Which quantiles are the easiest to measure?
- Using extremes seems to cause a bias problem
- We used BA diameter distribution and angle count samples. With unweighted distributions we would need to use fixed area samples.

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs
- Which quantiles are the easiest to measure?
- Using extremes seems to cause a bias problem
- We used BA diameter distribution and angle count samples. With unweighted distributions we would need to use fixed area samples.
- Here we worked with percentile-based distribution. What about other families, such as Weibull

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs
- Which quantiles are the easiest to measure?
- Using extremes seems to cause a bias problem
- We used BA diameter distribution and angle count samples. With unweighted distributions we would need to use fixed area samples.
- Here we worked with percentile-based distribution. What about other families, such as Weibull
  - Modeling simultaneously Weibull-parameters and percentiles

Discussion

- Using sample information combined with models provides possibilities to control the accuracy of a forest inventory according to the information needs
- Which quantiles are the easiest to measure?
- Using extremes seems to cause a bias problem
- We used BA diameter distribution and angle count samples. With unweighted distributions we would need to use fixed area samples.
- Here we worked with percentile-based distribution. What about other families, such as Weibull
    - Modeling simultaneously Weibull-parameters and percentiles
    - Combining PPM-weibull and a Weibull fitted to a small sample by ML

Publications

Mehtätalo, L. 2005. Localizing a predicted diameter distribution using sample information. Forest Science 51(4): 292-302.

Mehtätalo, L. and Kangas, A. 2005. An approach to optimizing data collection in an inventory by compartments. Canadian Journal of Forest Research 35(1): 100-112.

Mehtätalo, L., Maltamo, M., and Kangas, A. The use of quantile trees in predicting the diameter distribution of a stand. To appear in Silva Fennica