FINAL SCRIPT: APPROVED FOR FILMING

Submission ID #: 60827
Scriptwriter Name: Anastasia Gomez
Project Page Link: https://www.jove.com/account/file-uploader?src=18552283

Title: Development of an Individual-Tree Basal Area Increment Model using a Linear Mixed-Effects Approach

Authors and Affiliations:

Wenwen Wang[1], Yanfeng Bai[2], Chunqian Jiang[2], Jinghui Meng[1]

1. Research Center of Forest Management Engineering of National Forestry and Grassland Administration, Beijing Forestry University, Beijing, China
2. Research Institute of Forestry, Chinese Academy of Forestry, Beijing, China

Corresponding Authors:

Jinghui Meng
jmeng@bjfu.edu.cn

Email Addresses for All Authors:

wangwenwen_0321@163.com
baiyf@caf.ac.cn
jiangchq@caf.ac.cn
jmeng@bjfu.edu.cn

Author Questionnaire

1. Microscopy: Does your protocol involve video microscopy, such as filming a complex dissection or microinjection technique? **No**

2. Software: Does the part of your protocol being filmed include step-by-step descriptions of software usage? **Yes, all set**

3. Filming location: Will the filming need to take place in multiple locations? **No**

Current Protocol Length

Number of Steps:	8
Number of Shots:	9, 8 SC and 1 establishing shot for videographer.

Introduction

1. Introductory Interview Statements

REQUIRED:

1.1. **Wenwen Wang:** This protocol provides the key procedures of developing an individual-tree basal area increment model using a linear mixed-effects approach.

1.1.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera.

1.2. **Wenwen Wang:** The main advantage of this technique is that it can powerfully analyzes data with complex structures in forestry and significantly improve the performance of forest growth models.

1.2.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera.

Protocol

2. Linear mixed-effects model development with the package "nlme" in R software

2.1. Begin by reading the model development dataset **[1]** and loading the Package nlme *(spell out 'N-L-M-E')* in R software. Select sample plots as random effects to develop the mixed-effects model. **[2]**.

2.1.1. WIDE: Establishing shot of talent at the computer performing the protocol.

2.1.2. SCREEN: 60827_screenshot_1.mp4. 0:00 – 0:18. *Video Editor: Emphasize the following text in the command window:*

```
>model.development.dataset=read.csv("E:/DATA/JoVE/modelingdata.csv",
header=TRUE)
>library(nlme)
```

2.2. Fit all possible combinations of random effects with the maximum likelihood method and output the results. Set the intercept to random parameters, then change the random statements until all combinations are fitted. In the process of fitting, the codes may report errors due to the nonconvergence of the fitted model **[1]**.

2.2.1. SCREEN: 60827_screenshot_1.mp4. 0:19 – 1:27. *Video Editor: Speed this up after 0:30 and emphasize the following text in the command window:*

```
>Model<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML", random =~1|PLOT)

>summary(Model)
```

2.3. Select the best model by Akaike's information criterion, the Bayesian information criterion, the logarithm likelihood, and the likelihood ratio test **[1]**.

2.3.1. SCREEN: 60827_screenshot_1.mp4. 1:28 – 1:42. *Video Editor: Emphasize the following text in the command window:*

```
>anova(Model.1, Model.6)
>anova(Model.6, Model.23)

>anova(Model.23, Model.30)
```

2.4. Observe whether the residuals have heteroscedasticity from the residual plot. If there is heteroscedasticity, introduce the constant plus power function, the power function, and the exponential function to model the errors variance structure **[1]**.

2.4.1. SCREEN: 60827_screenshot_1.mp4. 1:43 – 2:12. *Video Editor: Emphasize the following text in the command window:*

```
>Model.30.1<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML",random=~1/DBH1+BAL+NT|PLOT,weights=varConstPower(form=~
fitted(.)))
>summary(Model.30.1)
>Model.30.2<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML",random=~1/DBH1+BAL+NT|PLOT,weights=varPower(form=~ fitted(.)))
>summary(Model.30.2)
>Model.30.3<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML",random=~1/DBH1+BAL+NT|PLOT,weights=varExp(form=~ fitted(.)))

>summary(Model.30.3)
```

2.5. Determine the best variance function for the model according to Akaike's information criterion, the Bayesian information criterion, the logarithm likelihood, and the likelihood ratio test **[1]**.

2.5.1. SCREEN: 60827_screenshot_1.mp4. 2:13 – 2:26. *Video Editor: Emphasize the following text in the command window:*

```
>anova(Model.30, Model.30.1)
>anova(Model.30, Model.30.2)

>anova(Model.30, Model.30.3)
```

2.6. Next, introduce the compound symmetry structure, first-order autoregressive structure, and a combination of first-order autoregressive and moving average structures to account for autocorrelation **[1]**.

2.6.1. SCREEN: 60827_screenshot_1.mp4. 2:27 – 2:57. *Video Editor: Emphasize the following text in the command window:*

```
>Model.30.3.1<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML",random=~1/DBH1+BAL+NT|PLOT,weights=varExp(form=~fitted(.)),
corr= corCompSymm())
>summary(Model.30.3.1)
>Model.30.3.2<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML", random=~1/DBH1+BAL+NT|PLOT,weights=varExp(form=~ fitted(.)),
corr=corAR1())
>summary(Model.30.3.2)
>Model.30.3.3<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="ML",random=~1/DBH1+BAL+NT|PLOT,weights=varExp(form=~ fitted(.)),
corr=corARMA(q=1,p=1))

>summary(Model.30.3.3)
```

2.7. Determine the best autocorrelation structure according to Akaike's information criterion, the Bayesian information criterion, the logarithm likelihood, and the likelihood ratio test **[1]**.

2.7.1. SCREEN: 60827_screenshot_1.mp4. 2:58 – 3:06. *Video Editor: Emphasize the following text in the command window:*

```
>anova(Model.30.3, Model.30.3.2)
```

2.8. Output the final results of the mixed-effects model using the restricted maximum likelihood method **[1]**.

2.8.1. SCREEN: 60827_screenshot_1.mp4. 3:06 – 3:21. *Video Editor: Emphasize the following text in the command window:*

```
>Mixed.model<-lme(Y~1/DBH1+BAL+NT+EL,data=model.development.dataset,
method="REML",random=~1/DBH1+BAL+NT|PLOT,weights=varExp(form=~ fitted(.)),
corr=corAR1())

>summary(Mixed.model)
```

Results

3. Results: Basal Area Increment Model for *P. asperata*

3.1. The basic basal area increment model for *P. asperata* is expressed with this equation **[1]**. The parameter estimates, their corresponding standard errors, and the lack-of-fit statistics are shown here **[2]**. Pronounced heteroscedasticity of the residuals was observed **[3]**.

3.1.1. Equation 7 from manuscript (below). *Video Editor: Just show the equation.*

$$\log(\mathrm{DBH}_2^2 - \mathrm{DBH}_1^2 + 1) = \beta_1 + \beta_2 1/\mathrm{DBH}_1 + \beta_3 \mathrm{BAL} + \beta_4 \mathrm{NT} + \beta_5 \mathrm{EL} + e_i$$

3.1.2. LAB MEDIA: Table 2.

3.1.3. LAB MEDIA: Figure 1.

3.2. There were 31 possible combinations of random-effects parameters for the basic basal area increment model. After fitting, 30 combinations reached convergence **[1]**.

3.2.1. LAB MEDIA: Table 3. *Video Editor: Scroll down the table as VO talks.*

3.3. Among these 30 combinations, Model 30 was selected **[1]** because it yielded the lowest AIC **[2]**, the lowest BIC **[3]**, and the largest LogLik *(pronounce 'log-like')* **[4]**. Furthermore, the LRT was significantly different when compared with the other models **[5]**.

3.3.1. LAB MEDIA: Table 3. *Video Editor: Just show the table headers and row 30. Also, show equation 8 from the manuscript below the table.*

$$\log(\mathrm{DBH}_2^2 - \mathrm{DBH}_1^2 + 1) = (\beta_1 + b_1) + (\beta_2 + b_2)1/\mathrm{DBH}_1 + (\beta_3 + b_3)\mathrm{BAL} + (\beta_4 + b_4)\mathrm{NT} + \beta_5 \mathrm{EL} + e_i$$

3.3.2. LAB MEDIA: Table 3, just headers and row 30. *Video Editor: Emphasize the AIC.*

3.3.3. LAB MEDIA: Table 3, just headers and row 30. *Video Editor: Emphasize the BIC.*

3.3.4. LAB MEDIA: Table 3, just headers and row 30. *Video Editor: Emphasize the LogLic.*

3.3.5. LAB MEDIA: Table 3, just headers and row 30. *Video Editor: Emphasize the LRT.*

3.4. The linear mixed-effects models with variance functions and correlation structures are shown here **[1]**. According to the AIC, BIC, Loglik, and LRT, the exponential function and AR-1 were selected as the best variance function and autocorrelation structure, respectively **[2]**.

3.4.1. LAB MEDIA: Table 4.

3.4.2. LAB MEDIA: Table 4. *Video Editor: Emphasize Model 30.3.2 row.*

3.5. The final linear mixed-effects individual-tree basal area increment model was proposed using the REML method **[1]**. The estimated fixed parameters, their corresponding standard errors, and the lack-of-fit statistics are shown here **[2]**. A significant improvement was observed in the residuals **[3]**.

3.5.1. Equation 7 from manuscript (below). *Video Editor: Just show the equation.*

$$\log(\mathrm{DBH}_2^2 - \mathrm{DBH}_1^2 + 1) = (2.8086 + \mathrm{b}_1) + (-6.24021 + \mathrm{b}_2)1/\mathrm{DBH}_1 + (-0.1324 + \mathrm{b}_3)\mathrm{BAL} + (-0.0001 + \mathrm{b}_4)\mathrm{NT} - 0.0003\mathrm{EL} + e_i$$

$$b_i = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \\ b_4 \end{bmatrix} \sim \mathrm{N}\left\{ \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, \psi_i = \begin{pmatrix} 0.1302 & -0.2088 & -0.0271 & -1.59\mathrm{E}-05 \\ -0.2088 & 6.1589 & -0.0552 & -2.28\mathrm{E}-04 \\ -0.0271 & -0.0552 & 0.0143 & 1.19\mathrm{E}-05 \\ -1.59\mathrm{E}-05 & -2.28\mathrm{E}-04 & 1.19\mathrm{E}-05 & 1.73\mathrm{E}-08 \end{pmatrix} \right\}$$

$$e_i \sim \mathrm{N}(0,\ R_i = 0.0886 G_i^{0.5} \Gamma_i G_i^{0.5})$$

$$G_i = \exp(-0.0104 y_i);\ \Gamma_i = \mathrm{AR}(1),\ \rho = 0.04245$$

3.5.2. LAB MEDIA: Table 5.

3.5.3. LAB MEDIA: Figure 2.

3.6. Prediction statistics of the two models show that the performance of the linear mixed-effects model was significantly improved compared to the basic model **[1]**.

3.6.1. LAB MEDIA: Table 6.

Conclusion

4. Conclusion Interview Statements

4.1. **Wenwen Wang:** When the model comparisons are completed, remember to use REML method to output the final results.

4.1.1. INTERVIEW: Named talent says the statement above in an interview-style shot, looking slightly off-camera. *Suggested B-roll: 2.8.1.*