

MLM, exerc 12: repeated measurements Sitka spruce data

Gerrit Gort

July 2, 2021

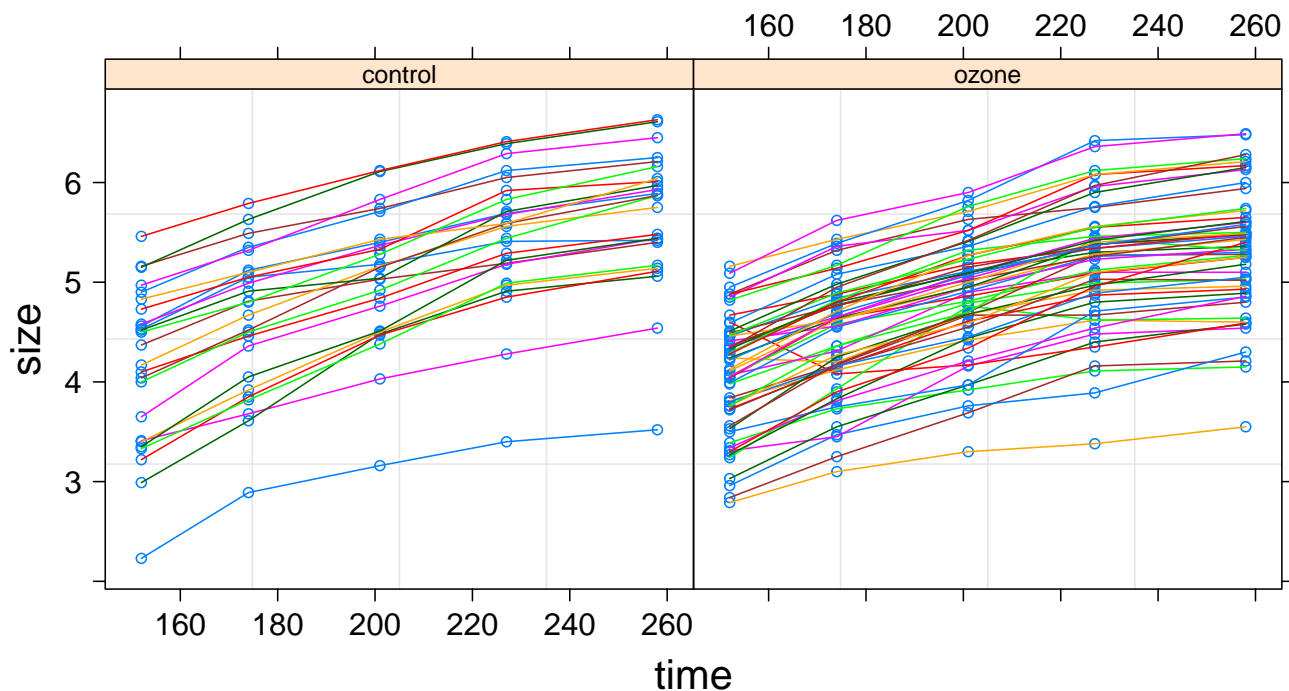
Read data

```
Sitka <- read.csv("exercise12.csv")
Sitka$tree <- factor(Sitka$tree)
Sitka$treat <- factor(Sitka$treat)
head(Sitka)
```

```
##   nr size time tree treat
## 1  1 4.51  152    1 ozone
## 2  2 4.98  174    1 ozone
## 3  3 5.41  201    1 ozone
## 4  4 5.90  227    1 ozone
## 5  5 6.15  258    1 ozone
## 6  6 4.24  152    2 ozone
```

Plot of size versus time per tree split by treatment

```
library(nlme)
Sitka.grp <- groupedData(size ~ time | tree, outer = ~treat,
  data = Sitka)
# nice plot of size vs time
plot(Sitka.grp, outer = ~treat, key = FALSE, aspect = 0.85, layout = c(2,
  1), xlab = list(fontsize = 20), ylab = list(fontsize = 20),
  scales = list(cex = 1.25))
```



Fitting different covariance structures

1: Compound symmetry: constant correlation among observations from same tree

```
(lme.tree1 <- gls(size ~ treat + time + treat:time, correlation = corCompSymm(form = ~1 | tree), data = Sitka))
```

```
## Generalized least squares fit by REML
## Model: size ~ treat + time + treat:time
## Data: Sitka
## Log-restricted-likelihood: -81.75852
##
## Coefficients:
## (Intercept)      treatozone          time treatozone:time
## 2.12171791      0.22167749      0.01414724     -0.00213851
##
## Correlation Structure: Compound symmetry
## Formula: ~1 | tree
## Parameter estimate(s):
```

```
##      Rho
## 0.9077829
## Degrees of freedom: 395 total; 391 residual
## Residual standard error: 0.6383461
```

```
summary(lme.tree1)
```

```
## Generalized least squares fit by REML
##  Model: size ~ treat + time + treat:time
##  Data: Sitka
##      AIC      BIC    logLik
## 175.517 199.3293 -81.75852
##
## Correlation Structure: Compound symmetry
## Formula: ~1 | tree
## Parameter estimate(s):
##      Rho
## 0.9077829
##
## Coefficients:
##              Value Std.Error  t-value p-value
## (Intercept)   2.1217179 0.15439225 13.742386  0.0000
## treatozone    0.2216775 0.18674206  1.187078  0.2359
## time          0.0141472 0.00046190 30.628557  0.0000
## treatozone:time -0.0021385 0.00055868 -3.827802  0.0002
##
## Correlation:
##              (Intr) tretzn time
## treatozone   -0.827
## time         -0.606  0.501
## treatozone:time 0.501 -0.606 -0.827
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -3.5274072 -0.5979164  0.1204015  0.6136729  2.1158148
##
## Residual standard error: 0.6383461
## Degrees of freedom: 395 total; 391 residual
```

```
logLik(lme.tree1)
```

```
## 'log Lik.' -81.75852 (df=6)
```

```
AIC(lme.tree1)
```

```
## [1] 175.517
```

```
# Different specification of same model
```

```
(lme.tree1.1 <- lme(size ~ treat + time + treat:time, random = ~1 |  
  tree, data = Sitka))
```

```
## Linear mixed-effects model fit by REML  
## Data: Sitka  
## Log-restricted-likelihood: -81.75852  
## Fixed: size ~ treat + time + treat:time  
## (Intercept) treatozone time treatozone:time  
## 2.12171791 0.22167749 0.01414724 -0.00213851  
##  
## Random effects:  
## Formula: ~1 | tree  
## (Intercept) Residual  
## StdDev: 0.6082011 0.1938483  
##  
## Number of Observations: 395  
## Number of Groups: 79
```

```
summary(lme.tree1.1)
```

```
## Linear mixed-effects model fit by REML  
## Data: Sitka  
## AIC BIC logLik  
## 175.517 199.3293 -81.75852  
##  
## Random effects:  
## Formula: ~1 | tree  
## (Intercept) Residual  
## StdDev: 0.6082011 0.1938483  
##  
## Fixed effects: size ~ treat + time + treat:time  
## Value Std.Error DF t-value p-value  
## (Intercept) 2.1217179 0.15439225 314 13.742386 0.0000  
## treatozone 0.2216775 0.18674206 77 1.187078 0.2388  
## time 0.0141472 0.00046190 314 30.628557 0.0000  
## treatozone:time -0.0021385 0.00055868 314 -3.827802 0.0002  
## Correlation:  
## (Intr) tretzn time  
## treatozone -0.827  
## time -0.606 0.501  
## treatozone:time 0.501 -0.606 -0.827  
##  
## Standardized Within-Group Residuals:  
## Min Q1 Med Q3 Max  
## -2.6949604 -0.4923652 0.1176021 0.5733781 4.3279068
```

```
##
## Number of Observations: 395
## Number of Groups: 79
```

```
logLik(lme.tree1.1)
```

```
## 'log Lik.' -81.75852 (df=6)
```

```
AIC(lme.tree1.1)
```

```
## [1] 175.517
```

2: AR(1): autocorrelation order 1

```
(lme.tree2 <- gls(size ~ treat + time + treat:time, correlation = corAR1(form = ~1 |
  tree), data = Sitka))
```

```
## Generalized least squares fit by REML
## Model: size ~ treat + time + treat:time
## Data: Sitka
## Log-restricted-likelihood: -34.93111
##
## Coefficients:
## (Intercept)      treatozone      time treatozone:time
## 2.154358716    0.231900511    0.013505267   -0.002219096
##
## Correlation Structure: AR(1)
## Formula: ~1 | tree
## Parameter estimate(s):
## Phi
## 0.9502363
## Degrees of freedom: 395 total; 391 residual
## Residual standard error: 0.6491015
```

```
summary(lme.tree2)
```

```
## Generalized least squares fit by REML
## Model: size ~ treat + time + treat:time
## Data: Sitka
## AIC      BIC      logLik
## 81.86221 105.6745 -34.93111
##
## Correlation Structure: AR(1)
```

```
## Formula: ~1 | tree
## Parameter estimate(s):
##      Phi
## 0.9502363
##
## Coefficients:
##              Value Std.Error   t-value p-value
## (Intercept)  2.1543587 0.19545769 11.022123  0.0000
## treatozone   0.2319005 0.23641195  0.980917  0.3272
## time         0.0135053 0.00073945 18.263939  0.0000
## treatozone:time -0.0022191 0.00089439 -2.481137  0.0135
##
## Correlation:
##              (Intr) tretzn time
## treatozone   -0.827
## time         -0.774  0.640
## treatozone:time 0.640 -0.774 -0.827
##
## Standardized residuals:
##      Min      Q1      Med      Q3      Max
## -3.2640774 -0.4222331  0.2900052  0.7576319  2.2674113
##
## Residual standard error: 0.6491015
## Degrees of freedom: 395 total; 391 residual
```

```
logLik(lme.tree2)
```

```
## 'log Lik.' -34.93111 (df=6)
```

```
AIC(lme.tree2)
```

```
## [1] 81.86221
```

3: Unstructured covariance matrix

- unstructured covariance matrix with heterogeneous variances
- unstructured covariance matrix with constant variance

```
(lme.tree3a <- gls(size ~ treat + time + treat:time, correlation = corSymm(form = ~1 |
  tree), weights = varIdent(form = ~1 | time), data = Sitka))
```

```
## Generalized least squares fit by REML
## Model: size ~ treat + time + treat:time
```

```
## Data: Sitka
## Log-restricted-likelihood: 21.56485
##
## Coefficients:
##      (Intercept)      treatozone      time treatozone:time
##      2.94552816      0.27498300      0.01034710      -0.00228976
##
## Correlation Structure: General
## Formula: ~1 | tree
## Parameter estimate(s):
## Correlation:
##  1      2      3      4
## 2 0.926
## 3 0.792 0.944
## 4 0.641 0.847 0.948
## 5 0.737 0.898 0.953 0.972
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | time
## Parameter estimates:
##      152      174      201      227      258
## 1.0000000 0.8426792 0.7946996 0.8491992 0.8401260
## Degrees of freedom: 395 total; 391 residual
## Residual standard error: 0.7585072
```

```
summary(lme.tree3a)
```

```
## Generalized least squares fit by REML
## Model: size ~ treat + time + treat:time
## Data: Sitka
##      AIC      BIC    logLik
## -5.12969 70.27575 21.56485
##
## Correlation Structure: General
## Formula: ~1 | tree
## Parameter estimate(s):
## Correlation:
##  1      2      3      4
## 2 0.926
## 3 0.792 0.944
## 4 0.641 0.847 0.948
## 5 0.737 0.898 0.953 0.972
## Variance function:
## Structure: Different standard deviations per stratum
## Formula: ~1 | time
## Parameter estimates:
##      152      174      201      227      258
```

```
## 1.0000000 0.8426792 0.7946996 0.8491992 0.8401260
##
## Coefficients:
##              Value Std.Error   t-value p-value
## (Intercept)  2.9455282 0.15558099 18.932443  0.0000
## treatozone   0.2749830 0.18817988  1.461277  0.1447
## time         0.0103471 0.00048447 21.357690  0.0000
## treatozone:time -0.0022898 0.00058598 -3.907589  0.0001
##
## Correlation:
##              (Intr) tretzn time
## treatozone   -0.827
## time         -0.633  0.523
## treatozone:time 0.523 -0.633 -0.827
##
## Standardized residuals:
##              Min          Q1          Med          Q3          Max
## -3.28773260 -0.68959530  0.01947652  0.52393143  2.12765425
##
## Residual standard error: 0.7585072
## Degrees of freedom: 395 total; 391 residual
```

```
logLik(lme.tree3a)
```

```
## 'log Lik.' 21.56485 (df=19)
```

```
AIC(lme.tree3a)
```

```
## [1] -5.12969
```

```
(lme.tree3b <- lme(size ~ treat + time + treat:time, random = list(tree = pdSymm(~1)),
  data = Sitka))
```

```
## Linear mixed-effects model fit by REML
## Data: Sitka
## Log-restricted-likelihood: -81.75852
## Fixed: size ~ treat + time + treat:time
##      (Intercept)      treatozone          time treatozone:time
##      2.12171791      0.22167749      0.01414724      -0.00213851
##
## Random effects:
## Formula: ~1 | tree
##      (Intercept) Residual
## StdDev:  0.6082011 0.1938483
##
## Number of Observations: 395
## Number of Groups: 79
```



```
summary(lme.tree3b)
```

```
## Linear mixed-effects model fit by REML
##   Data: Sitka
##       AIC      BIC    logLik
##  175.517 199.3293 -81.75852
##
## Random effects:
## Formula: ~1 | tree
##      (Intercept)  Residual
## StdDev:   0.6082011 0.1938483
##
## Fixed effects:  size ~ treat + time + treat:time
##               Value Std.Error DF   t-value p-value
## (Intercept)   2.1217179 0.15439225 314 13.742386 0.0000
## treatozone    0.2216775 0.18674206  77  1.187078 0.2388
## time          0.0141472 0.00046190 314 30.628557 0.0000
## treatozone:time -0.0021385 0.00055868 314 -3.827802 0.0002
## Correlation:
##              (Intr) tretzn time
## treatozone   -0.827
## time         -0.606  0.501
## treatozone:time 0.501 -0.606 -0.827
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.6949604 -0.4923652  0.1176021  0.5733781  4.3279068
##
## Number of Observations: 395
## Number of Groups: 79
```

```
logLik(lme.tree3b)
```

```
## 'log Lik.' -81.75852 (df=6)
```

```
AIC(lme.tree3b)
```

```
## [1] 175.517
```

4: random coefficient model

```
lme.tree4 <- lme(size ~ treat + time + treat:time, random = ~1 +
  time | tree, data = Sitka)
summary(lme.tree4)
```

```
## Linear mixed-effects model fit by REML
## Data: Sitka
##      AIC      BIC    logLik
## 146.1218 177.8714 -65.06088
##
## Random effects:
## Formula: ~1 + time | tree
## Structure: General positive-definite, Log-Cholesky parametrization
##      StdDev      Corr
## (Intercept) 0.803748734 (Intr)
## time        0.002538797 -0.651
## Residual    0.162608842
##
## Fixed effects: size ~ treat + time + treat:time
##      Value Std.Error DF t-value p-value
## (Intercept) 2.1217179 0.17944923 314 11.823500 0.0000
## treatozone 0.2216775 0.21704924 77 1.021323 0.3103
## time 0.0141472 0.00063871 314 22.149852 0.0000
## treatozone:time -0.0021385 0.00077253 314 -2.768176 0.0060
## Correlation:
##      (Intr) tretzn time
## treatozone -0.827
## time -0.729 0.603
## treatozone:time 0.603 -0.729 -0.827
##
## Standardized Within-Group Residuals:
##      Min      Q1      Med      Q3      Max
## -2.3056637 -0.6747055 0.1052744 0.5581490 3.2718433
##
## Number of Observations: 395
## Number of Groups: 79
```

```
logLik(lme.tree4)
```

```
## 'log Lik.' -65.06088 (df=8)
```

```
AIC(lme.tree4)
```

```
## [1] 146.1218
```

Comparison of AIC's:

- model 1 (compound symmetry): 175.5
- model 2 (autoregressive order 1): 81.9

- model 3a (unstructured, heterogeneous variance): -5.1
- model 3b (unstructured, common variance): 175.5
- model 4 (random coefficients): 146.1

Notice that the unstructured covariance matrix with heterogeneous variances gives by far the lowest AIC, meaning best fitting model. This is a quite costly structure with a lot parameters, though.

Compare F-values of fixed effects

```
anova(lme.tree1)
```

```
## Denom. DF: 391
##           numDF  F-value p-value
## (Intercept)      1 4904.888 <.0001
## treat          1    2.019 0.1562
## time           1 2383.472 <.0001
## treat:time      1   14.652 0.0002
```

```
anova(lme.tree2)
```

```
## Denom. DF: 391
##           numDF  F-value p-value
## (Intercept)      1 4695.482 <.0001
## treat          1    2.210 0.1379
## time           1  830.603 <.0001
## treat:time      1    6.156 0.0135
```

```
anova(lme.tree3a)
```

```
## Denom. DF: 391
##           numDF  F-value p-value
## (Intercept)      1 5272.340 <.0001
## treat          1    1.712 0.1915
## time           1 1038.341 <.0001
## treat:time      1   15.269 0.0001
```

```
anova(lme.tree3b)
```

```
##           numDF denDF  F-value p-value
## (Intercept)      1   314 4904.888 <.0001
## treat          1    77    2.019 0.1594
## time           1   314 2383.472 <.0001
## treat:time      1   314   14.652 0.0002
```

```
anova(lme.tree4)
```

```
##              numDF denDF  F-value p-value
## (Intercept)      1   314 4967.060 <.0001
## treat           1    77   2.118  0.1497
## time           1   314 1246.518 <.0001
## treat:time      1   314   7.663  0.0060
```

Notice that hypothesis test results for fixed effects do depend on the chosen covariance structure.

Extract variance covariance matrices

```
round(getVarCov(lme.tree1), 2)
```

```
## Marginal variance covariance matrix
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 0.41 0.37 0.37 0.37 0.37
## [2,] 0.37 0.41 0.37 0.37 0.37
## [3,] 0.37 0.37 0.41 0.37 0.37
## [4,] 0.37 0.37 0.37 0.41 0.37
## [5,] 0.37 0.37 0.37 0.37 0.41
## Standard Deviations: 0.64031 0.64031 0.64031 0.64031 0.64031
```

```
round(getVarCov(lme.tree2), 2)
```

```
## Marginal variance covariance matrix
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 0.42 0.40 0.38 0.36 0.34
## [2,] 0.40 0.42 0.40 0.38 0.36
## [3,] 0.38 0.40 0.42 0.40 0.38
## [4,] 0.36 0.38 0.40 0.42 0.40
## [5,] 0.34 0.36 0.38 0.40 0.42
## Standard Deviations: 0.64807 0.64807 0.64807 0.64807 0.64807
```

```
round(getVarCov(lme.tree3a), 2)
```

```
## Marginal variance covariance matrix
##      [,1] [,2] [,3] [,4] [,5]
## [1,] 0.58 0.45 0.36 0.31 0.36
## [2,] 0.45 0.41 0.36 0.35 0.37
## [3,] 0.36 0.36 0.36 0.37 0.37
## [4,] 0.31 0.35 0.37 0.41 0.40
## [5,] 0.36 0.37 0.37 0.40 0.41
## Standard Deviations: 0.76158 0.64031 0.6 0.64031 0.64031
```

```
round(getVarCov(lme.tree3b), 2)
```

```
## Random effects variance covariance matrix
##           (Intercept)
## (Intercept)      0.37
## Standard Deviations: 0.60828
```

```
getVarCov(lme.tree4, type = "marginal")
```

```
## tree 1
## Marginal variance covariance matrix
##           1      2      3      4      5
## 1 0.41740 0.38328 0.37385 0.36477 0.35395
## 2 0.38328 0.40516 0.37312 0.36772 0.36130
## 3 0.37385 0.37312 0.39866 0.37135 0.37032
## 4 0.36477 0.36772 0.37135 0.40128 0.37900
## 5 0.35395 0.36130 0.37032 0.37900 0.41580
## Standard Deviations: 0.64606 0.63652 0.63139 0.63347 0.64483
```

```
getVarCov(lme.tree4, type = "random.effects")
```

```
## Random effects variance covariance matrix
##           (Intercept)      time
## (Intercept)  0.6460100 -1.3289e-03
## time        -0.0013289  6.4455e-06
## Standard Deviations: 0.80375 0.0025388
```

```
# This is the covariance matrix of intercepts - slopes
```

To get correlation matrix do e.g.

```
covmat3a <- getVarCov(lme.tree3a)
std3a <- sqrt(diag(covmat3a))
corrmat3a <- diag(1/std3a) %*% covmat3a %*% diag(1/std3a)
round(corrmat3a, 2)
```

```
##           [,1] [,2] [,3] [,4] [,5]
## [1,] 1.00 0.93 0.79 0.64 0.74
## [2,] 0.93 1.00 0.94 0.85 0.90
## [3,] 0.79 0.94 1.00 0.95 0.95
## [4,] 0.64 0.85 0.95 1.00 0.97
## [5,] 0.74 0.90 0.95 0.97 1.00
```

Do the same for the other 4 covariance matrices...

Notice the rather unexpected pattern for the unstructured correlation matrix: correlation decreases until the third timepoint, but then increases again.

This may be related to a failing systematic part of the model. In the starting plot in this report clearly curved relationships can be seen. You could try to add a quadratic effect of time and next refit the models with the different covariance structures.

Try yourself...