Getting started with VetResearchLMM

Muhammad Yaseen

2026-11-10

Table of contents

1 Abstract **1**

2 Why a mixed model? **1**

3 The data **2**

4 Fit the model **2**

5 Read the results in plain language **4**

6 Means, estimates and contrasts **5**

7 Predicting random effects **7**

8 Where next **8**

9 References **8**

1 Abstract

VetResearchLMM contains the ten data sets and every analysis of Duchateau, Janssen and Rowlands (1998), *Linear Mixed Models. An Introduction with Applications in Veterinary Research*, and an engine for linear mixed models that gives the results the book prints from SAS PROC MIXED and PROC GLM. This vignette fits the split-plot experiment of Example 1.2.5 in a few lines, reads the results, and points to the rest of the package.

2 Why a mixed model?

An animal experiment rarely has a single source of random variation. In Example 1.2.5, two drugs (Berenil and Samorin) were given to herds, one herd per drug in each of six regions, and two doses to groups of animals within each herd.

Regions differ, herds within a region differ, and animals differ: three sources of variation. A drug is compared between herds, so its standard error must include the herd variation; a dose is compared within herds, so its standard error does not. A model with a single error term gets one of the two wrong. The linear mixed model

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \mathbf{e}, \qquad \mathbf{u} \sim N(\mathbf{0}, \mathbf{G}), \quad \mathbf{e} \sim N(\mathbf{0}, \mathbf{R}),$$

keeps the effects of interest (drug, dose) fixed in $\boldsymbol{\beta}$ and the sources of variation (region, herd) random in $\mathbf{u}$, so that $\mathrm{Var}(\mathbf{y}) = \mathbf{V} = \mathbf{Z}\mathbf{G}\mathbf{Z}^\top + \mathbf{R}$ has a part for each source.

3 The data

```
str(ex125)
```

```
'data.frame':   24 obs. of  4 variables:
 $ Region: int  1 1 1 1 2 2 2 2 3 3 ...
 $ Drug  : Factor w/ 2 levels "BERENIL","samorin": 1 1 2 2 1 1 2 2 1 1 ...
 $ dose  : Factor w/ 2 levels "h","l": 1 2 1 2 1 2 1 2 1 2 ...
 $ Pcv   : num  22.6 21.8 19.1 16.4 29 28.8 25.3 18.2 24 23.7 ...
```

4 Fit the model

The SAS program of the book (Appendix B.2.4) is

```
proc mixed data=ex125;
  class drug dose region;
  model pcv = drug dose drug*dose / solution ddfm=satterth;
  random region drug*region;
run;
```

and the same model with **VetResearchLMM** is

```
fm <-
  vlm_fit(
      fixed    = Pcv ~ Drug + dose + Drug:dose
    , data     = ex125
    , random   = ~ Region + Region:Drug
    , repeated = NULL
    , method   = "REML"
    , ddfm     = "satterth"
    , type     = 3
    , class    = NULL
    , bound    = TRUE
    , control  = list()
```

```
    )
fm
```

```
Linear mixed model: Pcv ~ Drug + dose + Drug:dose
Estimation: REML; denominator df: satterth; converged after 6 iterations
Random: Region; Region:Drug

Covariance parameter estimates
   parameter subject estimate     se      z      p
      Region            5.151  3.739  1.378 0.0842
 Region:Drug           0.3870  1.022 0.3788 0.3524
    Residual            2.096 0.9375  2.236 0.0127

Fit statistics
                statistic   value
    -2 Res Log Likelihood 92.3786
  AIC (smaller is better) 98.3786
 AICC (smaller is better) 99.8786
  BIC (smaller is better) 101.366

Solution for fixed effects
      effect     level estimate     se    df      t        p
 (Intercept)              17.13  1.128 8.208  15.19  < 1e-04
        Drug   BERENIL    7.150 0.9098 11.82  7.859  < 1e-04
        Drug   samorin        0                           .
        dose         h    4.350 0.8359 10.00  5.204 0.000399
        dose         l        0                           .
   Drug:dose BERENIL h   -3.183  1.182 10.00 -2.693 0.022594
   Drug:dose BERENIL l        0                           .
   Drug:dose samorin h        0                           .
   Drug:dose samorin l        0                           .

Type 3 tests of fixed effects
    effect num_df den_df     F        p
      Drug      1  5.000 64.58 0.000483
      dose      1  10.00 21.78 0.000886
 Drug:dose      1  10.00 7.251 0.022594

Sources: Duchateau, Janssen and Rowlands (1998); Patterson and Thompson (1971);
Gilmour, Thompson and Cullis (1995); Littell, Milliken, Stroup and Wolfinger
(1996); Satterthwaite (1946); Giesbrecht and Burns (1985); Fai and Cornelius
(1996). Full references: vlm_cite(x).
```

The printout has the tables of SAS: the covariance parameter estimates (5.151 for regions, 0.387 for herds within regions and 2.096 for animals, as in Table 2.4.2.1 of the book), the fit statistics, the solution for the fixed effects with the

last level of each effect set to zero (17.13, 7.15, 4.35 and −3.18, Example 2.5.1.1) and the Type 3 tests. Each component is a `data.table`:

```
fm$covariance
```

```
     parameter subject estimate       se         z          p
1:      Region         5.150833 3.739172 1.3775332 0.08417372
2: Region:Drug         0.387000 1.021605 0.3788156 0.35241241
3:    Residual         2.096417 0.937546 2.2360680 0.01267366
```

5 Read the results in plain language

```
vlm_interpret(
    x     = fm
  , alpha = 0.05
  , top   = 3
  )
```

```
Linear mixed model

Pcv ~ Drug + dose + Drug:dose, fitted by REML with 3 covariance parameters.

Findings:
- Of the total variance 7.63, 67.5% is due to Region, 5.1% is due to
  Region:Drug, 27.5% is due to the residual (within) variation.
- Significant at the 0.05 level: Drug (F = 64.6 on 1 and 5 df, p = 0.00048);
  dose (F = 21.8 on 1 and 10 df, p = 0.00089); Drug:dose (F = 7.25 on 1 and
  10 df, p = 0.023).

Recommendations:
- The interaction Drug:dose is significant: compare the levels of one factor
  within each level of the other (vlm_lsmeans() of the interaction, or
  vlm_estimate()) rather than main-effect means.
- Report means with standard errors in the broad inference space
  (vlm_lsmeans()); use vlm_estimate() with `random` for the narrow space of
  the observed levels.

Sources:
- Duchateau, L., Janssen, P. and Rowlands, G. J. (1998). Linear Mixed Models.
  An Introduction with Applications in Veterinary Research. International
  Livestock Research Institute, Nairobi. ISBN 92-9146-038-9.
- Patterson, H. D. and Thompson, R. (1971). Recovery of inter-block
  information when block sizes are unequal. Biometrika, 58(3), 545-554.
  doi:10.1093/biomet/58.3.545
- Gilmour, A. R., Thompson, R. and Cullis, B. R. (1995). Average information
```

REML: an efficient algorithm for variance parameter estimation in linear mixed models. Biometrics, 51(4), 1440-1450. doi:10.2307/2533274
- Littell, R. C., Milliken, G. A., Stroup, W. W. and Wolfinger, R. D. (1996). SAS System for Mixed Models. SAS Institute, Cary, NC.
- Satterthwaite, F. E. (1946). An approximate distribution of estimates of variance components. Biometrics Bulletin, 2(6), 110-114. doi:10.2307/3002019
- Giesbrecht, F. G. and Burns, J. C. (1985). Two-stage analysis based on a mixed model: large-sample asymptotic theory and small-sample simulation results. Biometrics, 41(2), 477-486. doi:10.2307/2530872
- Fai, A. H.-T. and Cornelius, P. L. (1996). Approximate F-tests of multiple degree of freedom hypotheses in generalized least squares analyses of unbalanced split-plot experiments. Journal of Statistical Computation and Simulation, 54(4), 363-378. doi:10.1080/00949659608811740

6 Means, estimates and contrasts

Least-squares means of the four drug-dose combinations, with Tukey-Kramer comparisons:

```
lm1 <-
  vlm_lsmeans(
      fit    = fm
    , specs  = ~ Drug:dose
    , at     = NULL
    , diff   = TRUE
    , adjust = "tukey"
    , level  = 0.95
    )
lm1
```

```
Least-squares means (denominator df: satterth; 95% limits)

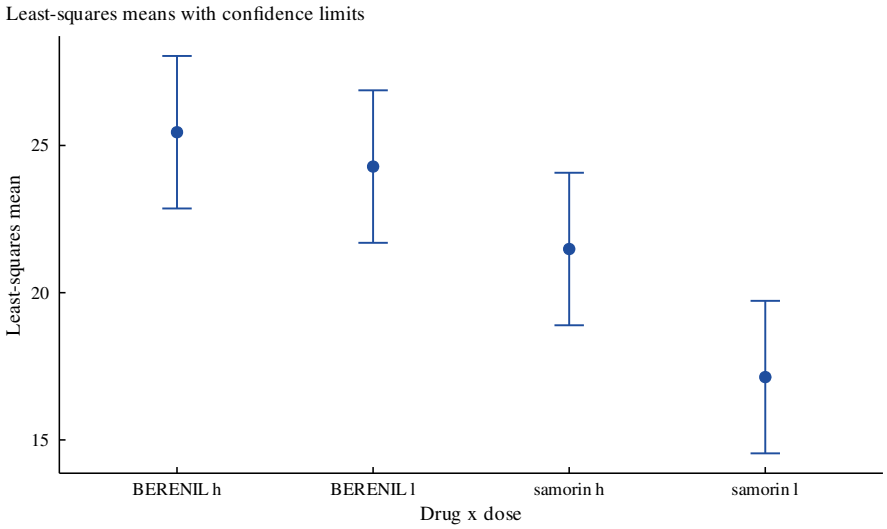
Least squares means
    Drug dose estimate    se    df     t      p lower upper
 BERENIL    h    25.45 1.128 8.208 22.56 <1e-04 22.86 28.04
 BERENIL    l    24.28 1.128 8.208 21.53 <1e-04 21.69 26.87
 samorin    h    21.48 1.128 8.208 19.05 <1e-04 18.89 24.07
 samorin    l    17.13 1.128 8.208 15.19 <1e-04 14.54 19.72

Differences of least squares means (adjustment: tukey)
   level_1   level_2 estimate     se    df     t        p p_adjusted   lower
 BERENIL h BERENIL l    1.167 0.8359 10.00 1.396 0.193041    0.52926  -1.391
 BERENIL h samorin h    3.967 0.9098 11.82 4.360 0.000962    0.00459   1.259
 BERENIL h samorin l    8.317 0.9098 11.82 9.141  < 1e-04    < 1e-04   5.609
 BERENIL l samorin h    2.800 0.9098 11.82 3.077 0.009738    0.04197 0.09258
```

```
 BERENIL l samorin l    7.150 0.9098 11.82 7.859  < 1e-04    < 1e-04   4.443
 samorin h samorin l    4.350 0.8359 10.00 5.204 0.000399    0.00188   1.793
 upper
 3.724
 6.674
 11.02
 5.507
 9.857
 6.907
```

Sources: Duchateau, Janssen and Rowlands (1998); Searle, Speed and Milliken (1980); Kramer (1956); Satterthwaite (1946); Giesbrecht and Burns (1985); Fai and Cornelius (1996). Full references: vlm_cite(x).

```
plot(lm1)
```



Any linear combination is an ESTIMATE statement, coefficients given per effect in the order of the levels of `fm$fixed`:

```
vlm_estimate(
    fit    = fm
  , L      = list(
       "drug dif"     = list(Drug = c(-1, 1), "Drug:dose" = c(-0.5,
        ↪  -0.5, 0.5, 0.5))
     , "Samorin mean" = list(Intercept = 1, Drug = c(0, 1), dose =
 ↪  c(0.5, 0.5), "Drug:dose" = c(0, 0, 0.5, 0.5))
     )
  , random = NULL
  , level  = 0.95
```

```
  )
```

```
Estimates (denominator df: satterth)

Estimates
        label estimate     se    df      t        p  lower  upper space
     drug dif   -5.558 0.6917 5.000 -8.036 0.000483 -7.336 -3.780 broad
 Samorin mean    19.31  1.048 6.205  18.43  < 1e-04  16.77  21.85 broad

Sources: Duchateau, Janssen and Rowlands (1998); McLean, Sanders and Stroup
(1991); Satterthwaite (1946); Giesbrecht and Burns (1985); Fai and Cornelius
(1996). Full references: vlm_cite(x).
```

7 Predicting random effects

The book's Example 2.4.3.1 predicts the merit of eight sires from the weaning weights of their lambs:

```
fs <-
  vlm_fit(
      fixed    = Ww ~ 1
    , data     = ex127
    , random   = ~ sire
    , repeated = NULL
    , method   = c("REML", "ML", "Type3")
    , ddfm     = c("default", "contain", "residual", "betwithin",
↪   "satterth", "kr")
    , type     = 3
    , class    = NULL
    , bound    = TRUE
    , control  = list()
    )
bl <-
  vlm_blup(
      fit   = fs
    , level = 0.95
    )
bl
```

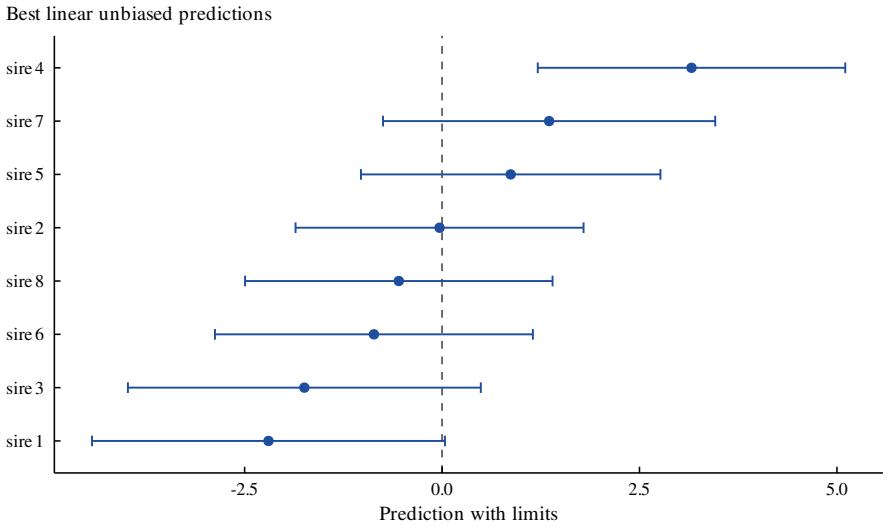
```
Best linear unbiased predictions of the random effects

Solution for random effects
 effect level estimate     se df        t       p   lower   upper
   sire     1   -2.198  1.101 35   -1.997 0.05364  -4.432 0.03641
   sire     2 -0.03189 0.8983 35 -0.03550 0.97188  -1.855   1.792
   sire     3   -1.743  1.101 35   -1.584 0.12216  -3.978  0.4908
   sire     4    3.159 0.9589 35    3.294 0.00227   1.212   5.106
```

```
sire     5   0.8692 0.9340 35   0.9306 0.35842  -1.027   2.765
sire     6  -0.8628 0.9915 35  -0.8702 0.39013  -2.876   1.150
sire     7   1.356  1.036 35   1.309 0.19922 -0.7475   3.459
sire     8  -0.5476 0.9589 35  -0.5710 0.57162  -2.494   1.399
```

Sources: Duchateau, Janssen and Rowlands (1998); Henderson (1975). Full references: vlm_cite(x).

```
plot(bl)
```



8 Where next

- `vignette("VetResearchLMM_02_theory")`: the methods, with formulas.
- `vignette("VetResearchLMM_03_book_chapters_1_2")` and `vignette("VetResearchLMM_04_book_cha` every example of the book, with the printed numbers next to the computed ones.
- `vignette("VetResearchLMM_05_validation")`: checks against the book and against **lme4**, **lmerTest**, **pbkrtest** and **nlme**.
- `vignette("VetResearchLMM_06_app")`: the point-and-click application, `vlm_app()`.
- The help pages `?Examp1.3.2` to `?Examp3.3` give the code of each example.

9 References

Duchateau, L., Janssen, P. and Rowlands, G. J. (1998). *Linear Mixed Models. An Introduction with Applications in Veterinary Research.* International Livestock Research Institute, Nairobi.