

Modelling Maize Project: Cob.weight

Modelling with random effects

Author: Nisia Trisconi | Zurich Data Scientists
Reviewer: Dr. Luisa Barbanti | Zurich Data Scientists

October 30, 2023

Contents

1	Freezing Package versions	2
2	Load packages	2
3	Settings	2
4	Getting data	3
5	Design	5
6	Response variable: <i>Cob_weight</i>	5
6.1	Aim	5
6.2	Model fitting	5
7	Session information	8

1 Freezing Package versions

The following code lines are commented out because the `{checkpoint}` package no longer works.

```
## (messages are omitted in this chunk)
##
# library(checkpoint)
# checkpoint(snapshot_date = "2022-11-15")
```

2 Load packages

```
## (messages are omitted from this chunk)
##
library(dplyr)
library(kableExtra)
library(ggplot2)
library(tibble) ## function rownames_to_column()
library(lme4)
```

3 Settings

Global settings:

```
Sys.setenv(lang = "en_US")
theme_set(theme_bw())

if (!dir.exists("Prepared_data_and_models")) {
  dir.create("Prepared_data_and_models")
}
```

4 Getting data

```
d.maize <- readRDS(file = paste0("Prepared_data_and_models/",
                                   "d.maize_PreparedData.RDS"))
```

Overview of the data:

```
dim(d.maize)
```

```
[1] 108 33
```

```
head(d.maize)[1:min(ncol(d.maize), 30)]
```

```
# A tibble: 6 x 30
  pot   soil      well depth seed.weight fungus date.germinated observations
  <chr> <chr>    <chr> <dbl>    <dbl> <chr>    <chr>          <chr>
1 A1    Bio garden a      3      30 <NA>    2022-05-11    <NA>
2 A1    Bio garden b      5      34 <NA>    2022-05-11    <NA>
3 A1    Bio garden c      2      35 <NA>    2022-05-09    <NA>
4 A1    Bio garden d      1      40 <NA>    2022-05-10    <NA>
5 A1    Bio garden e      4      46 <NA>    2022-05-11    <NA>
6 A1    Bio garden f      6      37 <NA>    2022-05-11    <NA>
# i 22 more variables: height_2022_07_05 <chr>, cob_weight <chr>, ...12 <dbl>,
# pot.fac <fct>, soil.fac <fct>, well.fac <fct>, seed.weight.grams <dbl>,
# fungus.fac <fct>, date.germinated.asDate <date>, obs.time <fct>,
# broken <lgl>, height_2022_07_05.num <dbl>, plant.found <lgl>,
# cob_weight.num <dbl>, germinated.in.lab <lgl>, germinated.in.field <lgl>,
# germinated.yes <lgl>, days.to.germination <dbl>,
# days.to.germination.censored <dbl>, seed_coord_y <dbl>, ...
```

```
str(d.maize)
```

```
tibble [108 x 33] (S3: tbl_df/tbl/data.frame)
 $ pot           : chr [1:108] "A1" "A1" "A1" "A1" ...
 $ soil          : chr [1:108] "Bio garden" "Bio garden" "Bio garden" "Bio garden" ...
 $ well          : chr [1:108] "a" "b" "c" "d" ...
 $ depth         : num [1:108] 3 5 2 1 4 6 6 4 5 1 ...
 $ seed.weight   : num [1:108] 30 34 35 40 46 37 27 16 23 22 ...
 $ fungus        : chr [1:108] NA NA NA NA ...
 $ date.germinated : chr [1:108] "2022-05-11" "2022-05-11" "2022-05-09" "2022-05-10" ...
 $ observations   : chr [1:108] NA NA NA NA ...
 $ height_2022_07_05 : chr [1:108] "217" "131" "143" "194" ...
 $ cob_weight     : chr [1:108] "117" "26" "61" "109" ...
 $ ...12         : num [1:108] NA NA NA NA NA NA NA NA NA NA ...
 $ pot.fac       : Factor w/ 18 levels "A1","A2","A3",...: 1 1 1 1 1 1 2 2 2 2 ...
 $ soil.fac      : Factor w/ 4 levels "Bio garden","Composana",...: 1 1 1 1 1 1 3 3 3 3 ..
 $ well.fac      : Factor w/ 6 levels "a","b","c","d",...: 1 2 3 4 5 6 1 2 3 4 ...
 $ seed.weight.grams : num [1:108] 0.3 0.34 0.35 0.4 0.46 0.37 0.27 0.16 0.23 0.22 ...
 $ fungus.fac    : Factor w/ 2 levels "no","yes": 1 1 1 1 1 1 1 1 1 1 ...
 $ date.germinated.asDate : Date[1:108], format: "2022-05-11" "2022-05-11" ...
 $ obs.time      : Factor w/ 2 levels "morning","night": 2 2 2 2 2 2 2 2 2 2 ...
 $ broken        : logi [1:108] FALSE FALSE FALSE FALSE FALSE FALSE ...
 $ height_2022_07_05.num : num [1:108] 217 131 143 194 206 233 158 282 241 232 ...
 $ plant.found    : logi [1:108] TRUE TRUE TRUE TRUE TRUE TRUE ...
 $ cob_weight.num : num [1:108] 117 26 61 109 106 156 57 286 51 120 ...
 $ germinated.in.lab : logi [1:108] TRUE TRUE TRUE TRUE TRUE TRUE ...
```

```

$ germinated.in.field      : logi [1:108] FALSE FALSE FALSE FALSE FALSE FALSE ...
$ germinated.yes          : logi [1:108] TRUE TRUE TRUE TRUE TRUE TRUE ...
$ days.to.germination      : num [1:108] 11 11 9 10 11 11 11 11 NA 9 ...
$ days.to.germination.censored: num [1:108] 11 11 9 10 11 11 11 11 14 9 ...
$ seed_coord_y            : num [1:108] 1 1 2 2 3 3 1 1 2 2 ...
$ seed_coord_x            : num [1:108] 1 2 1 2 1 2 3 4 3 4 ...
$ position_field_x        : num [1:108] 1 1 1 1 1 1 1 1 1 1 ...
$ position_field_x_cm     : num [1:108] 50 50 50 50 50 50 50 50 50 50 ...
$ position_field_y        : int [1:108] 1 2 3 4 5 6 7 8 9 10 ...
$ position_field_y_cm     : num [1:108] 25 50 75 100 125 150 175 200 225 250 ...

```

5 Design

108 maize seeds are planted in 18 different pots, each with 6 wells.

Inside one pot, the same soil is used. The soils that were used are: Bio garden (4 pots), Composana (4 pots), herbs (6 pots), mixture (4 pots).

In each well, one maize seed is planted at a pre-defined depth (in cm), which is allocated randomly to the well. The maximum value for depth is 6cm and this corresponds to planting the seed directly in the coconut fiber that makes up the pot.

Wells in the same pots are allocated as follows:

```
      [,1] [,2]
[1,] "e"  "f"
[2,] "c"  "d"
[3,] "a"  "b"
```

The pots are arranged as follows on a table in the lab:

```
      [,1] [,2] [,3] [,4] [,5] [,6]
[1,] "C1" "C2" "C3" "C4" "C5" "C6"
[2,] "B1" "B2" "B3" "B4" "B5" "B6"
[3,] "A1" "A2" "A3" "A4" "A5" "A6"
```

Seeds are watered for the first time on 04.30.2022 and are transferred to the field on 05.15.2022 according to the same scheme.

Some seeds are broken when planted, one seed develops a fungus.

Some seeds germinate in the lab, others in the field, while some seeds never germinate.

On 07.05.2022, the height of all maize plants is measured in cm. The plants that were not measured for time reasons receive a height value of `not measured`, while the plants that were not found and could hence not be measured present missing values.

On 09.16.2022, the weight of the cob is measured for all plants that have a cob. The variety of maize that was planted typically yields 1 cob per plant.

6 Response variable: *Cob_weight*

6.1 Aim

We are interested in testing whether *cob_weight.num* is influenced by the following variables:

- Soil (variable *soil.fac*)
- Seed weight (variable *seed.weight*)
- Depth (variable *depth*)

Introducing *well.fac* and *pot.fac* as random effects.

6.2 Model fitting

The variable *cob_weight.num* is a continuous variable (in particular it is an amount), whose density is already well-centered, so there is no need to log transform it.

Based on the graphical analysis, it is not necessary to transform any of the explanatory variables. Therefore, they can be used as they are in the model.

seed.weight and *depth* are numeric variables which are introduced in the model as continuous variables; all the others are introduced as factors.

pot.fac and *well.fac* are introduced in the model as random effects, to account for the fact that observations inside the same pot (or in the same well position) could have more similar weight.

The first step is to fit a model using all the available explanatory variables.

```
lm.cob_weight <- lmer(cob_weight.num ~
  soil.fac +
  depth +
  seed.weight +
  (1 | pot.fac) +
  (1 | well.fac),
  data = d.maize,
  na.action = na.omit)
```

boundary (singular) fit: see help('isSingular')

```
##
summary(lm.cob_weight)
```

Linear mixed model fit by REML ['lmerMod']

Formula: cob_weight.num ~ soil.fac + depth + seed.weight + (1 | pot.fac) +
(1 | well.fac)

Data: d.maize

REML criterion at convergence: 812.2

Scaled residuals:

Min	1Q	Median	3Q	Max
-2.12043	-0.45249	0.01522	0.47835	3.82456

Random effects:

Groups	Name	Variance	Std.Dev.
pot.fac	(Intercept)	0.0	0.000
well.fac	(Intercept)	0.0	0.000
Residual		2431.6	49.311

Number of obs: 80, groups: pot.fac, 18; well.fac, 6

Fixed effects:

	Estimate	Std. Error	t value
(Intercept)	123.55538	30.18738	4.0929
soil.facComposana	14.53272	16.07131	0.9043
soil.facherbs	24.43720	14.66386	1.6665
soil.facmixture	9.04197	16.92374	0.5343
depth	0.79582	3.37105	0.2361
seed.weight	-0.78343	0.92070	-0.8509

Correlation of Fixed Effects:

	(Intr)	sl.fcC	sl.fch	sl.fcm	depth
sol.fcCmpsn	-0.213				
soil.fchrbs	-0.320	0.522			
soil.fcmxtr	-0.246	0.440	0.479		
depth	-0.265	-0.077	-0.097	0.095	
seed.weight	-0.851	-0.009	0.091	-0.032	-0.127

optimizer (nloptwrap) convergence code: 0 (OK)

boundary (singular) fit: see `help('isSingular')`

We observe that the variance for the random effects is numerically equal to “0”, thus we can exclude them and come back to the linear model (in the document “2_Modelling_maize_lm_cob.weight.pdf”).

7 Session information

```
sessionInfo()
```

```
R version 4.3.1 (2023-06-16)
```

```
Platform: aarch64-apple-darwin20 (64-bit)
```

```
Running under: macOS Sonoma 14.0
```

```
Matrix products: default
```

```
BLAS: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRblas.0.dylib
```

```
LAPACK: /Library/Frameworks/R.framework/Versions/4.3-arm64/Resources/lib/libRlapack.dylib; LAPACK vers
```

```
locale:
```

```
[1] en_US.UTF-8/en_US.UTF-8/en_US.UTF-8/C/en_US.UTF-8/en_US.UTF-8
```

```
time zone: Europe/Zurich
```

```
tzcode source: internal
```

```
attached base packages:
```

```
[1] stats      graphics  grDevices  utils      datasets  methods    base
```

```
other attached packages:
```

```
[1] lme4_1.1-34      Matrix_1.6-1.1  tibble_3.2.1    ggplot2_3.4.4
```

```
[5] kableExtra_1.3.4 dplyr_1.1.3     knitr_1.44
```

```
loaded via a namespace (and not attached):
```

```
[1] gtable_0.3.4      compiler_4.3.1  Rcpp_1.0.11     webshot_0.5.5
[5] tidyselect_1.2.0  xml2_1.3.5      stringr_1.5.0   splines_4.3.1
[9] systemfonts_1.0.5 scales_1.2.1     boot_1.3-28.1   yaml_2.3.7
[13] fastmap_1.1.1     lattice_0.21-8  R6_2.5.1        generics_0.1.3
[17] MASS_7.3-60       nloptr_2.0.3    munsell_0.5.0   minqa_1.2.6
[21] svglite_2.1.2     pillar_1.9.0    rlang_1.1.1     utf8_1.2.4
[25] stringi_1.7.12    xfun_0.40       viridisLite_0.4.2 cli_3.6.1
[29] withr_2.5.1       magrittr_2.0.3  digest_0.6.33   rvest_1.0.3
[33] grid_4.3.1        rstudioapi_0.15.0 nlme_3.1-162    lifecycle_1.0.3
[37] vctrs_0.6.4       evaluate_0.22   glue_1.6.2      fansi_1.0.5
[41] colorspace_2.1-0  rmarkdown_2.25  httr_1.4.7      tools_4.3.1
[45] pkgconfig_2.0.3   htmltools_0.5.6.1
```