Lathyrus ms2: Selection on reaction norms - Extraction of RN parameters using BLUPs

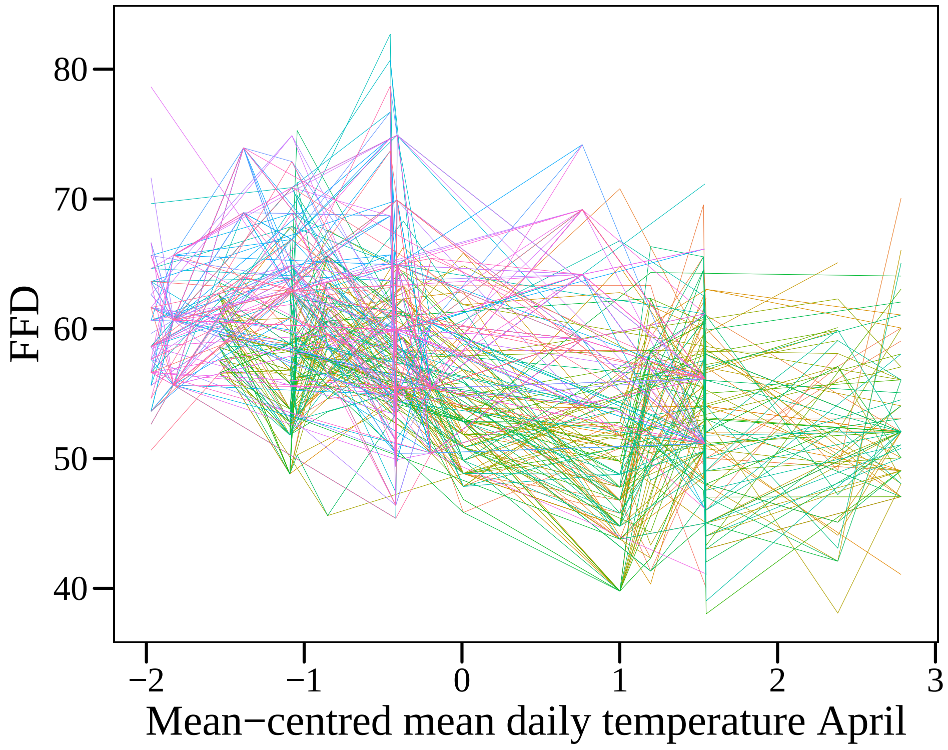
Contents

Code based on Arnold et al 2019 New Phyt.

The x variable (mean daily temperature April) is mean-centered (substracting the mean), so the intercepts reflect average values for the population and individuals. From here on, we use this mean-centred temperature (cmean_4).

```
data_3yrs$cmean_4<-scale(data_3yrs$mean_4,center=T,scale=F)
data_4yrs$cmean_4<-scale(data_4yrs$mean_4,center=T,scale=F)
data_5yrs$cmean_4<-scale(data_5yrs$mean_4,center=T,scale=F)
```

Plot the main effects (raw values of FFD against mean-centred temperatures for each plant id)



Basic linear model

Fit a linear model for the fixed effect of temperature on FFD and observe the average population-level reaction norm. Note that we also add a random effect for ‘year’ to take account of the repeated measures at each temperature (to account for differences among the years across which each id is represented). The mixed model is to be fitted using ML rather than REML so that models that contain different fixed effects can be compared directly.

Using blmer which does maximum a posteriori estimation for linear and generalized linear mixed-effects models in a Bayesian setting. Allows the user to do Bayesian inference or penalized maximum likelihood, with priors imposed on the different model components.

```
model1.1 <- blmer(FFD ~ cmean_4 + (1|year), REML = FALSE, data = data_4yrs,
                  lmerControl(optimizer = "Nelder_Mead"))
summary(model1.1)
```

```
## Cov prior  : year ~ wishart(df = 3.5, scale = Inf, posterior.scale = cov, common.scale = TRUE)
## Prior dev  : -0.1223
##
## Linear mixed model fit by maximum likelihood  ['blmerMod']
## Formula: FFD ~ cmean_4 + (1 | year)
##    Data: data_4yrs
## Control: lmerControl(optimizer = "Nelder_Mead")
##
##      AIC      BIC   logLik deviance df.resid
##   8780.5   8801.7  -4386.3   8772.5     1451
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0238 -0.6794 -0.0937  0.5932  4.7466
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  year     (Intercept) 24.76    4.976
##  Residual             22.82    4.777
## Number of obs: 1455, groups:  year, 22
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  57.9447     1.0695  54.179
## cmean_4      -2.4452     0.7941  -3.079
##
## Correlation of Fixed Effects:
##         (Intr)
## cmean_4 -0.004
```

```
r.squaredGLMM(model1.1)
```

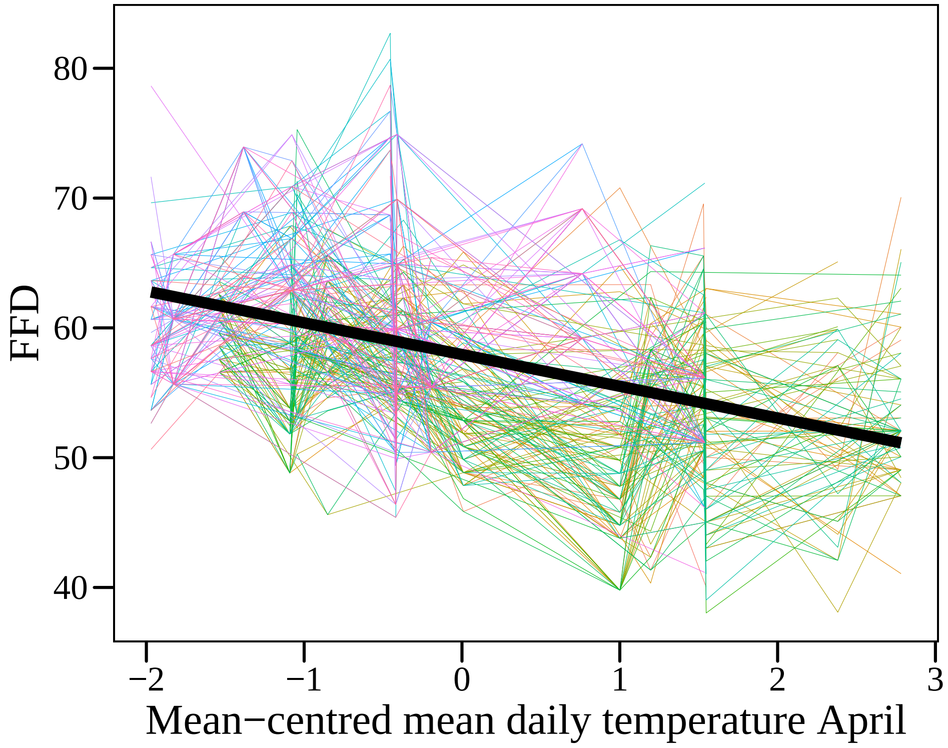
```
## Warning: 'r.squaredGLMM' now calculates a revised statistic. See the help
## page.
```

```
##            R2m       R2c
## [1,] 0.181672 0.6075111
```

Visually assess how well the linear model fits the raw data by overlaying the regression line from model1.1 as an average population-level reaction norm. Use the predict function to predict y-values accross the continuous x-axis and then plot the fixed effect of temperature from model1.1 over the raw id-specific reaction norms.

```r
temperature_pred <- data.frame(cmean_4 = seq(from = min(data_4yrs$cmean_4),
                                              to = max(data_4yrs$cmean_4),
                                              length.out = 50))
temperature_pred$fit1.1 <- predict(model1.1, newdata = temperature_pred, re.form = NA)
# re.form=NA includes no random effects

# Plot the raw data and overlay the fit of Model1.1
ggplot(temperature_pred, aes(x = cmean_4, y = fit1.1)) +
  geom_line(data = data_4yrs, aes(y = FFD, colour = id),size = 0.1) +
  geom_line(size = 2) +  ylab("FFD") +
  xlab("Mean-centred mean daily temperature April")+my_theme()
```



Quadratic fixed effects model

Fit a quadratic model for the fixed effect of temperature on FFD.

```r
model1.2 <- blmer(FFD ~ poly(cmean_4, 2, raw = T) + (1|year),
                  REML = FALSE, data = data_4yrs,lmerControl(optimizer = "Nelder_Mead"))
summary(model1.2)
```

```
## Cov prior  : year ~ wishart(df = 3.5, scale = Inf, posterior.scale = cov, common.scale = TRUE)
## Prior dev  : -0.122
##
## Linear mixed model fit by maximum likelihood  ['blmerMod']
## Formula: FFD ~ poly(cmean_4, 2, raw = T) + (1 | year)
##    Data: data_4yrs
## Control: lmerControl(optimizer = "Nelder_Mead")
##
```

```
##      AIC      BIC   logLik deviance df.resid
##   8782.5   8808.9  -4386.3   8772.5     1450
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.0244 -0.6793 -0.0936  0.5928  4.7467
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  year     (Intercept) 24.75    4.975
##  Residual             22.82    4.777
## Number of obs: 1455, groups:  year, 22
##
## Fixed effects:
##                            Estimate Std. Error t value
## (Intercept)                57.87195    1.55134  37.304
## poly(cmean_4, 2, raw = T)1 -2.46952    0.87858  -2.811
## poly(cmean_4, 2, raw = T)2  0.04015    0.62077   0.065
##
## Correlation of Fixed Effects:
##              (Intr) p(_4,2,r=T)1
## p(_4,2,r=T)1  0.307
## p(_4,2,r=T)2 -0.724 -0.428
```

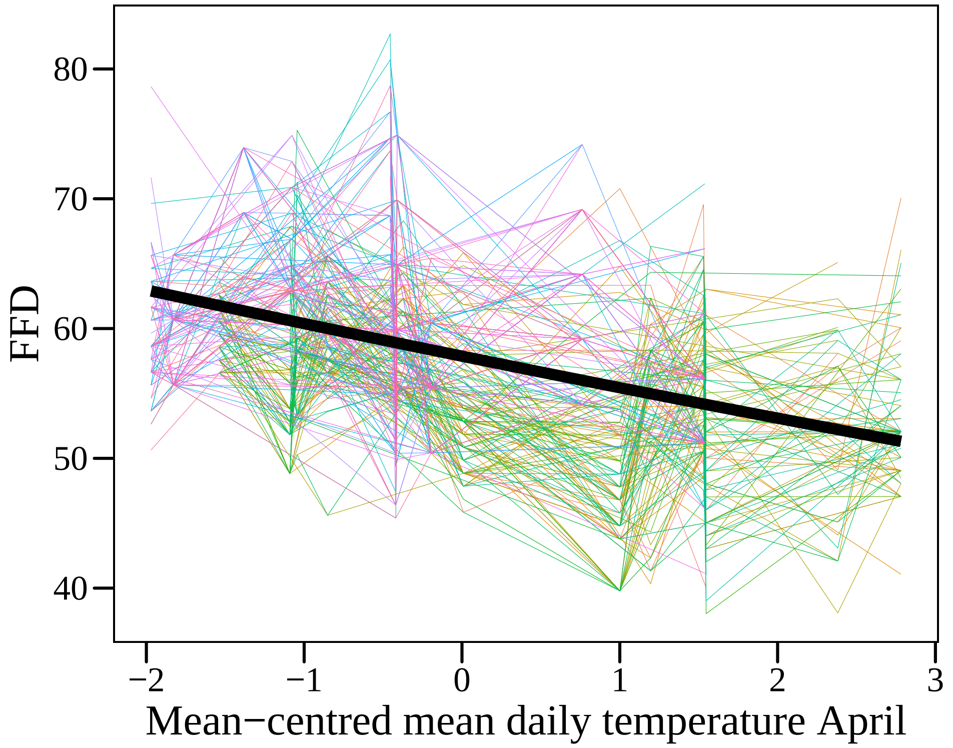
```r
r.squaredGLMM(model1.2)
```

```
##            R2m       R2c
## [1,] 0.1821737 0.6077058
```

Predict values based on the model fit and plot the overall model fit over the top of the raw data.

```r
temperature_pred$fit1.2 <- predict(model1.2, newdata = temperature_pred, re.form = NA)

ggplot(temperature_pred, aes(x = cmean_4, y = fit1.2)) +
  geom_line(data = data_4yrs, aes(y = FFD, colour = id),size = 0.1) +
  geom_line(size = 2) + ylab("FFD") +
  xlab("Mean-centred mean daily temperature April")+my_theme()
```



Compare with previous model using likelihood ratio test and AIC.

```
chi2 <- 2*(summary(model1.2)$logLik - summary(model1.1)$logLik)
1-pchisq(chi2,1)
```

```
## 'log Lik.' 0.9463939 (df=5)
```

```
AIC(model1.1, model1.2)
```

```
##          df      AIC
## model1.1  4 8780.527
## model1.2  5 8782.523
```

Very small increase in R2, non-significant LRT p-value and larger AIC. Model1.1 with only linear (non-quadratic effects) is better.

Linear fixed effects with random intercepts model

Fit a linear mixed effects model (random intercepts only) for the fixed effect of temperature on FFD and random effect of id intercepts. We are allowing the y-intercept value to vary among ids.

```
model1.3 <- blmer(FFD ~ cmean_4 + (1|year) + (1|id), REML = FALSE, data = data_4yrs,
                 lmerControl(optimizer = "Nelder_Mead"))
summary(model1.3)
```

```
## Cov prior  : id ~ wishart(df = 3.5, scale = Inf, posterior.scale = cov, common.scale = TRUE)
##            : year ~ wishart(df = 3.5, scale = Inf, posterior.scale = cov, common.scale = TRUE)
## Prior dev  : 2.3146
##
## Linear mixed model fit by maximum likelihood  ['blmerMod']
## Formula: FFD ~ cmean_4 + (1 | year) + (1 | id)
##    Data: data_4yrs
## Control: lmerControl(optimizer = "Nelder_Mead")
##
##      AIC      BIC   logLik deviance df.resid
##   8723.8   8750.2  -4356.9   8713.8     1450
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1404 -0.6183 -0.0938  0.5703  4.5899
##
## Random effects:
##  Groups   Name        Variance Std.Dev.
##  id       (Intercept)  3.315   1.821
##  year     (Intercept) 24.583   4.958
##  Residual             19.527   4.419
## Number of obs: 1455, groups:  id, 243; year, 22
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  58.0814     1.0716  54.202
## cmean_4      -2.4478     0.7913  -3.093
##
## Correlation of Fixed Effects:
##         (Intr)
## cmean_4 -0.003
```

```
r.squaredGLMM(model1.3)
```

```
##            R2m       R2c
## [1,] 0.1824562 0.6633812
```

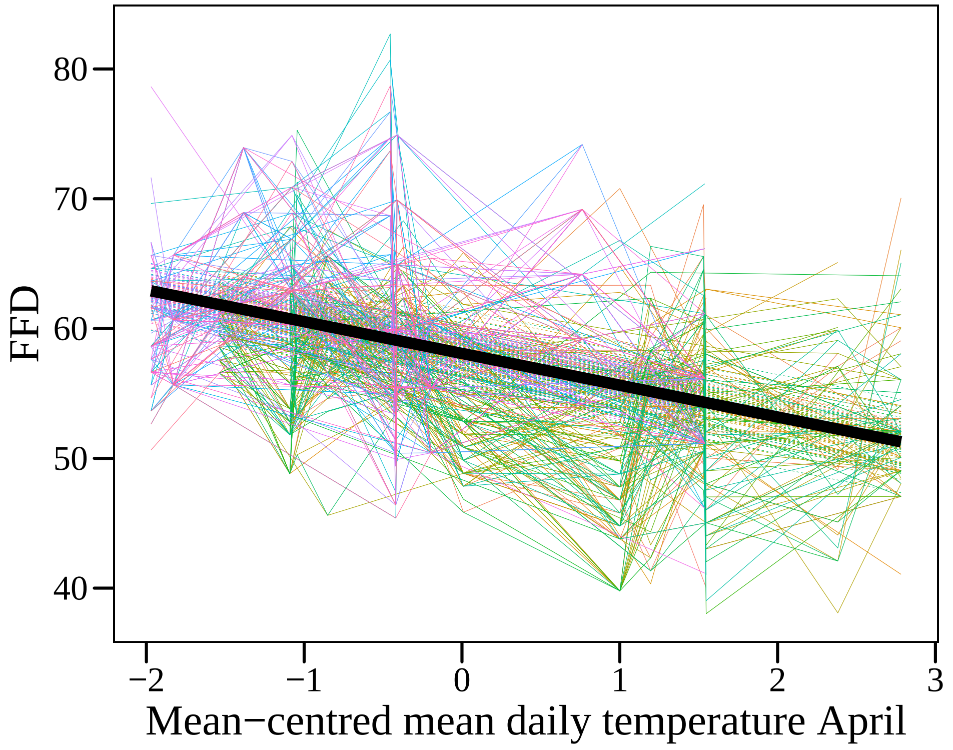
The outcome of the mixed-effects model is the linear effect of temperature on FFD, whilst allowing the intercepts of each id's FFD to account for some of the residual variance in the model.

Predict values based on the model fit and plot the overall model fit over the top of the raw data.

```
temperature_pred$fit1.3 <- predict(model1.3, newdata = temperature_pred, re.form = NA)

# Make a prediction for the average population-level mean reaction norm
# and append it to the dataset
data_4yrs$pred_pop1.3  <- predict(model1.3, re.form = NA)
# Make predictions for each id-level reaction norm
data_4yrs$pred_id1.3 <- predict(model1.3, re.form = ~(1|id))

# Plot predicted id reaction norms over the raw data, along with the overall mean
ggplot(temperature_pred, aes(x = cmean_4, y = fit1.3)) +
  geom_line(data = data_4yrs, aes(y = pred_id1.3, group = id, colour = id), lty = 2,size = 0.1)+
  geom_line(data = data_4yrs, aes(y = FFD, group = id, colour = id),size = 0.1) +
  geom_line(size = 2) +
  ylab("FFD") + xlab("Mean-centred mean daily temperature April") +  my_theme()
```



Compare with previous model using likelihood ratio test and AIC.

```
chi2 <- 2*(summary(model1.3)$logLik - summary(model1.1)$logLik)
1-pchisq(chi2, 1)
```

```
## 'log Lik.' 1.787459e-14 (df=5)
```

```
AIC(model1.1, model1.2, model1.3)
```

```
##          df      AIC
## model1.1  4 8780.527
## model1.2  5 8782.523
## model1.3  5 8723.776
```

R2 has substantially increased, significant LRT p-value and smaller AIC. Model1.3 with random intercepts explains more residual variance than model1.1 without trading-off against increased model complexity.

Linear fixed effects with linear random regression model

Fit a linear mixed effects model for the fixed effect of temperature on FFD and random effect of id intercepts and slopes. Allows the slopes of ids to vary in addition to the intercepts, so that the random regression slopes might be fit better to the observed patterns in the raw data. The addition of '+x-variable' (here: '1+cmean_4') to the left side of the random effect term ('|id') in model1.3 allows the slopes of the random id regressions to vary across mean-centred temperature.

```
model1.4 <- blmer(FFD ~ cmean_4 + (1|year) + (1+cmean_4|id), REML = FALSE,
                  data = data_4yrs,lmerControl(optimizer = "Nelder_Mead"))
summary(model1.4)
```

```
## Cov prior  : id ~ wishart(df = 4.5, scale = Inf, posterior.scale = cov, common.scale = TRUE)
##             : year ~ wishart(df = 3.5, scale = Inf, posterior.scale = cov, common.scale = TRUE)
## Prior dev  : 8.5314
##
## Linear mixed model fit by maximum likelihood  ['blmerMod']
## Formula: FFD ~ cmean_4 + (1 | year) + (1 + cmean_4 | id)
##    Data: data_4yrs
## Control: lmerControl(optimizer = "Nelder_Mead")
##
##      AIC      BIC   logLik deviance df.resid
##   8702.8   8739.7  -4344.4   8688.8     1448
##
## Scaled residuals:
##     Min      1Q  Median      3Q     Max
## -3.1728 -0.6228 -0.0900  0.5931  4.2738
##
## Random effects:
##  Groups   Name        Variance Std.Dev. Corr
##  id       (Intercept)  3.1686  1.7800
##           cmean_4      0.5995  0.7743   0.73
##  year     (Intercept) 24.6169  4.9615
##  Residual             18.6299  4.3162
## Number of obs: 1455, groups:  id, 243; year, 22
##
## Fixed effects:
##             Estimate Std. Error t value
## (Intercept)  58.0695     1.0723  54.154
## cmean_4      -2.4152     0.7941  -3.041
##
## Correlation of Fixed Effects:
##         (Intr)
## cmean_4 0.003
```

```
r.squaredGLMM(model1.4)
```

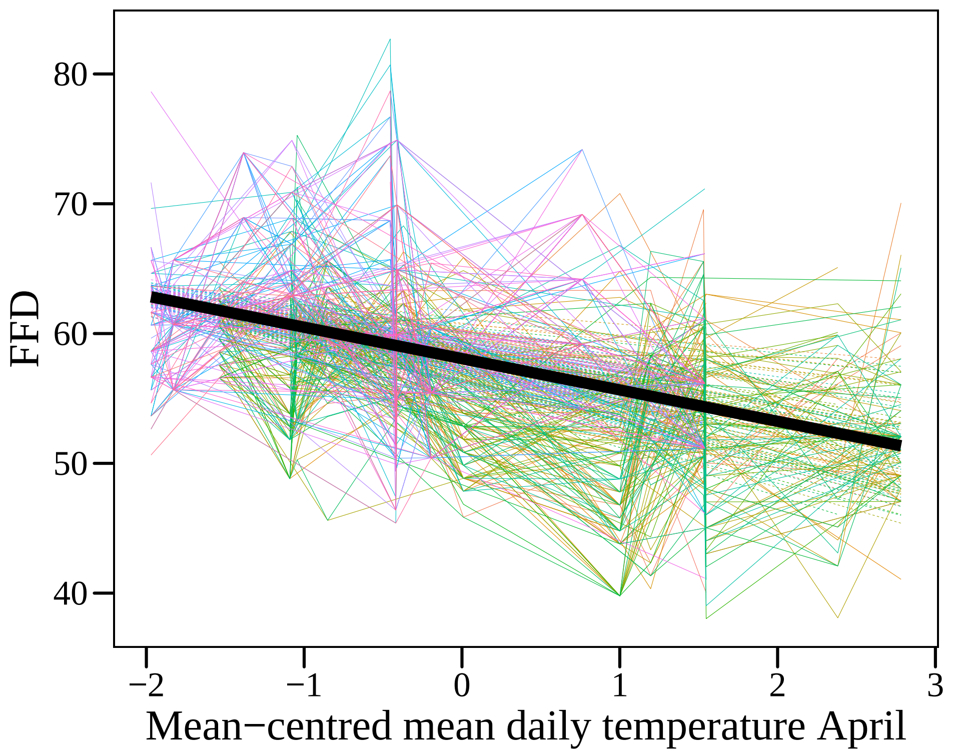
```
##            R2m       R2c
## [1,] 0.1783396 0.6775581
```

Predict values based on the model fit and plot the overall model fit over the top of the raw data.

```
temperature_pred$fit1.4 <- predict(model1.4, newdata = temperature_pred, re.form = NA)

# Make a prediction for the population-level mean reaction norm
# and append it to the dataset
data_4yrs$pred_pop1.4  <- predict(model1.4, re.form = NA)
# Make predictions for the id-level reaction norms
data_4yrs$pred_id1.4 <- predict(model1.4, re.form = ~(1+cmean_4|id))

# Plot predicted id reaction norms over the raw data, along with the overall mean
ggplot(temperature_pred, aes(x = cmean_4, y = fit1.4)) +
  geom_line(data = data_4yrs, aes(y = pred_id1.4, group = id, colour = id), lty = 2,size = 0.1) +
  geom_line(data = data_4yrs, aes(y = FFD, group = id, colour = id),size = 0.1) +
  geom_line(size = 2) +
  ylab("FFD") + xlab("Mean-centred mean daily temperature April") +  my_theme()
```



Compare with previous model using likelihood ratio test and AIC.

```
chi2 <- 2*(summary(model1.4)$logLik - summary(model1.3)$logLik)
# The df difference between models can be checked by
# looking at the df within the models being compared
summary(model1.3)$logLik
```

```
## 'log Lik.' -4356.888 (df=5)
```

```
summary(model1.4)$logLik
```

```
## 'log Lik.' -4344.385 (df=7)
```

```
# Note that between model1.3 and model1.4 there is a change of 2 df, so the
# pchisq change needs to be specified with 2 df rather than 1 as in previous comparisons.
1-pchisq(chi2, 2)
```

```
## 'log Lik.' 3.716043e-06 (df=7)
```

```
AIC(model1.1, model1.2, model1.3, model1.4)
```

```
##          df      AIC
## model1.1  4 8780.527
## model1.2  5 8782.523
## model1.3  5 8723.776
## model1.4  7 8702.770
```

R2 increases, significant LRT p-value and smaller AIC. The random regression mixed model (model1.4) has significantly improved the model fit to the data.Model1.4 is the best model for these data. We will proceed hereafter with model1.4 to extract the best linear unbiased predictors (BLUPs) for each id.

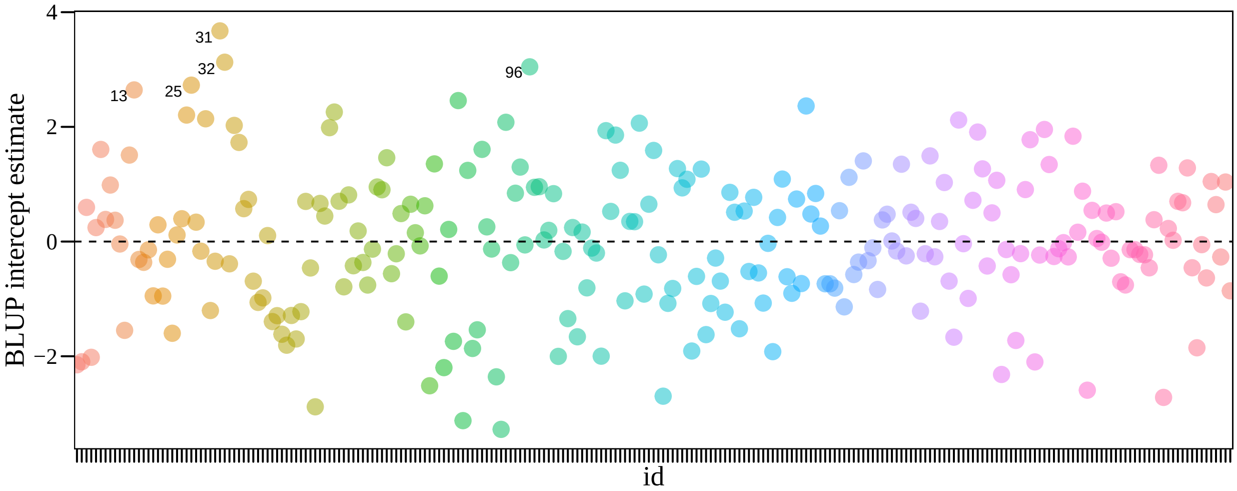
Extract BLUPs from model1.4 (linear random regression mixed model)

BLUPs represent the response of a given id to the fixed effect of temperature as the difference between that id's predicted response and the population-level average predicted response. Here, we calculate and plot BLUPs for ranking plasticity.

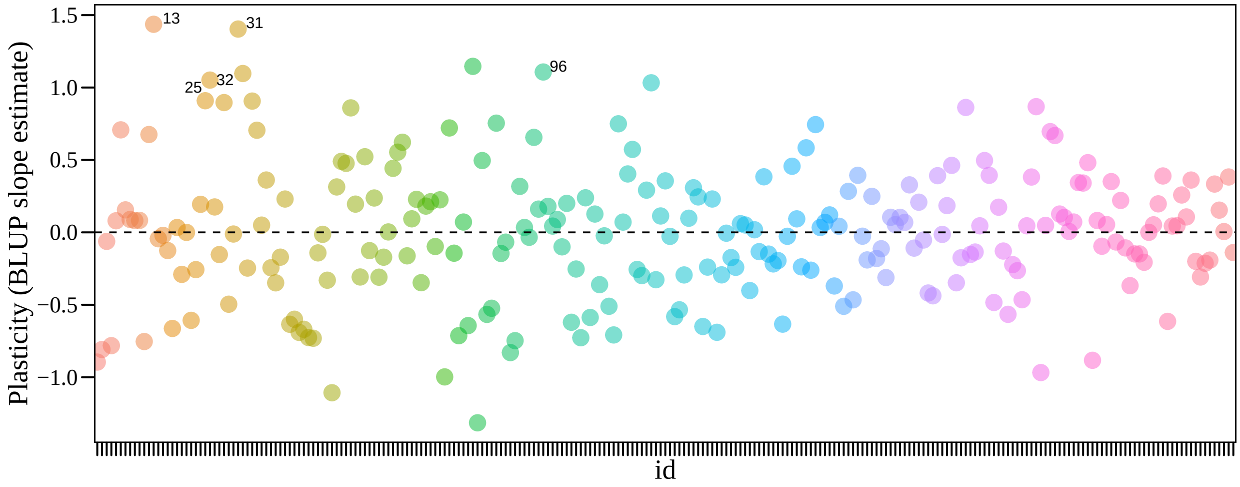
```
id_blups <- ranef(model1.4)$`id`
id_index <- as.factor(c(1:243))
id_data  <- cbind(id_index, id_blups)
colnames(id_data) <- c("id", "BLUP_int", "BLUP_slope")
with(id_data,cor(BLUP_int,BLUP_slope)) # highly correlated!
```

```
## [1] 0.9297162
```

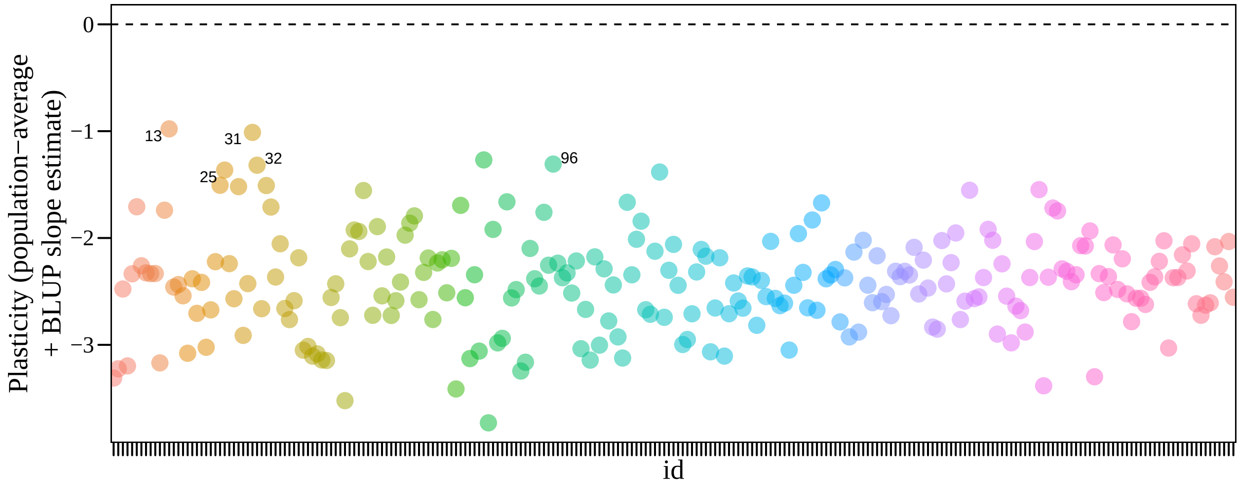
The BLUP intercept term indicates the difference in id elevation relative to the population-average, so more positive values of BLUP intercept indicate that the id's reaction norm occurs above the population-level average and negative values are below the population-level average. The BLUP intercept values are not a measure of plasticity, but these values may be correlated with BLUP slope values and otherwise may be a parameter of interest for comparing among ids.



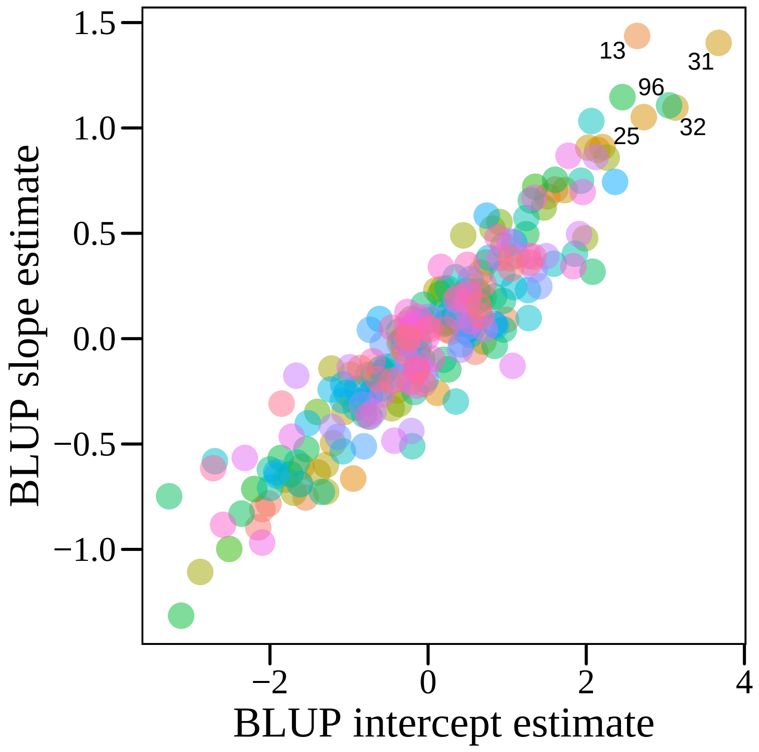
The BLUP slope estimate is the difference in slope (relative steepness of change) between the population-level average response and the response of the id. Here, that is the difference in slope of FFD for each value of temperature relative to the population-level average slope.



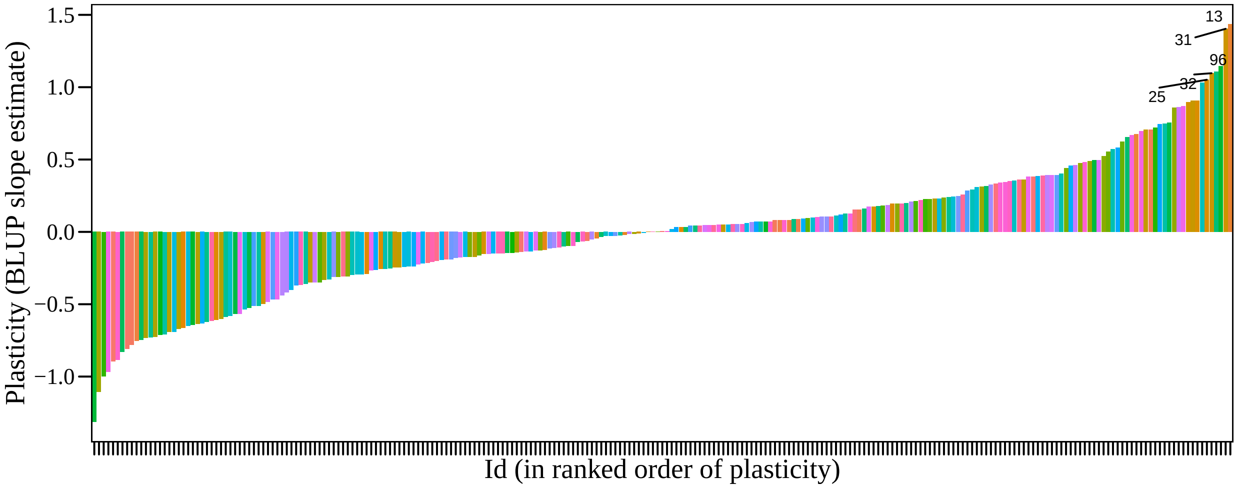
We now add the BLUP slopes for the ids to the population average. Because the population-level average response is negative overall, all ids have a negative slope when the BLUP slope estimates are added to the population-level average slope estimate from model1.4.



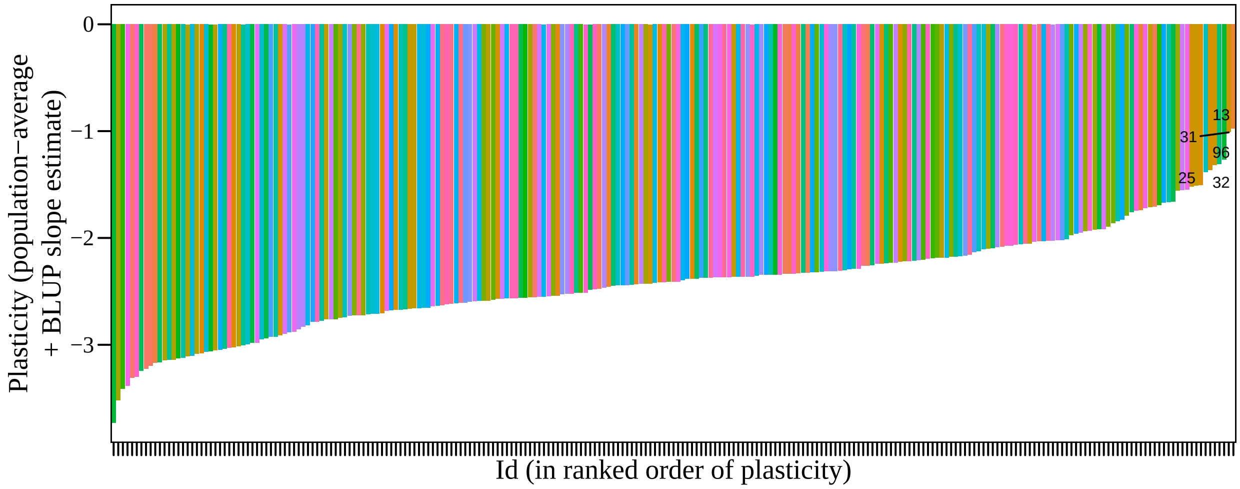
The BLUP intercept and slope estimates are sometimes correlated. The correlation coefficient is given in the random effects correlation from the model1.4 summary, which is 0.89. This positive relationship can clearly be seen when plotting the BLUP slope estimate against the BLUP intercept estimate. Ids with the most positive BLUP slope estimate (labelled) have the highest positive intercept, and have the least plasticity across growth temperatures (see previous figs with the same individuals labelled).



We can rank the BLUPs in order: sorting BLUPs by slope of most to least plastic. Because the population-level average response is negative, the most negative BLUP slope estimates represent steeper reaction norm slopes and hence greater plasticity, and more positive BLUP slope estimates represent flatter reaction norms and less plasticity in FFD in response to temperatures.



Another way to visualise the plasticity rank for negative data is by adding the BLUP slope values to the population-level average effect of temperature.



Important note! BLUPs estimated from linear mixed-effects regression models fitted in lme4 are single point estimates that do not have associated measures of uncertainty. As a result, any derived statistics or formal interpretation of plasticity based on these BLUPs is potentially very dangerous and anticonservative without properly accounting for estimation uncertainty. For using BLUPs beyond simple ranking (e.g., of the least to most plastic genotypes), it is strongly encouraged to read the references provided here to avoid the misuse of BLUPs by using a Bayesian MCMC framework (e.g., by using the MCMCglmm package in R) to generate estimates of uncertainty around BLUPs (Hadfield, 2010; Hadfield et al., 2010; Houslay & Wilson, 2017).

Merge reaction norm parameters with previous data

```
id_data$id_text<-as.factor(row.names(id_data))
data_4yrs<-data_4yrs %>%
   right_join(id_data[c(2:4,6)], by = c("id" = "id_text"))
```