

Mixed-effects models

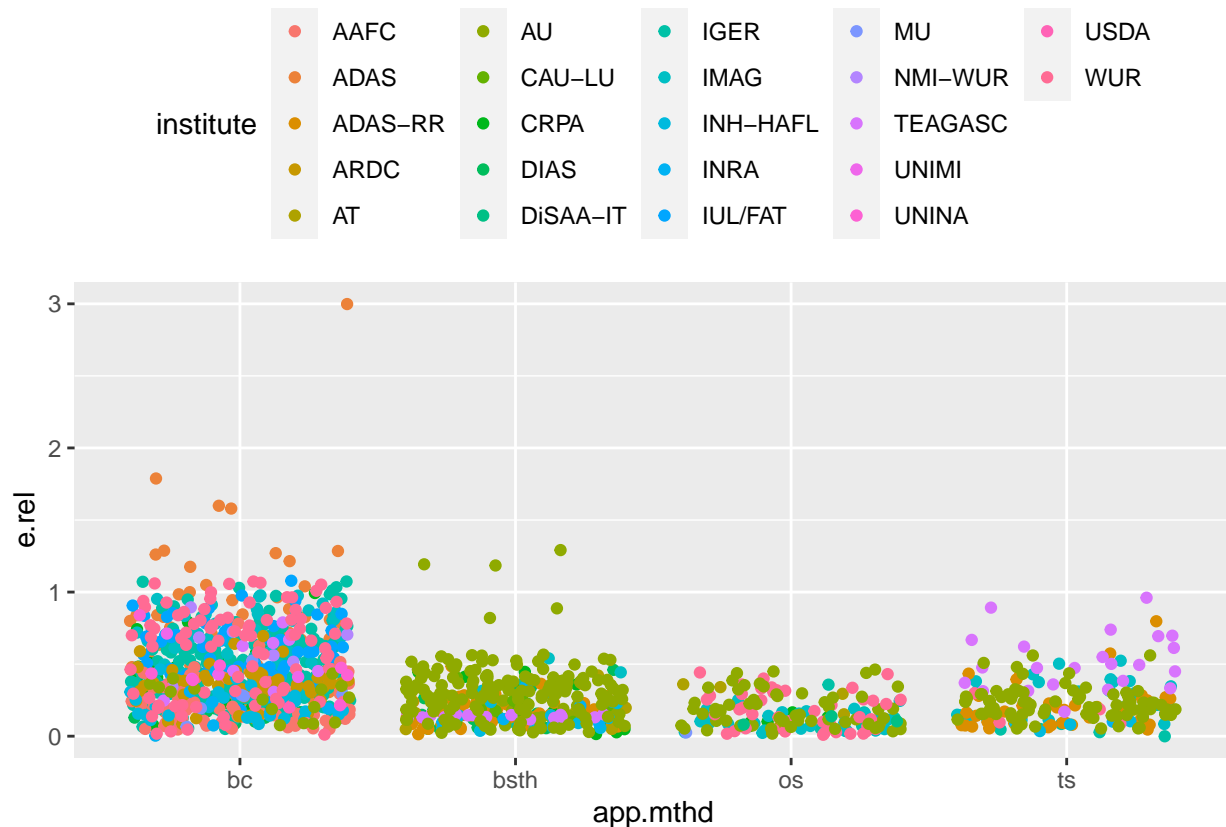
Sasha D. Hafner

12 May, 2023 13:39

1. Take a look

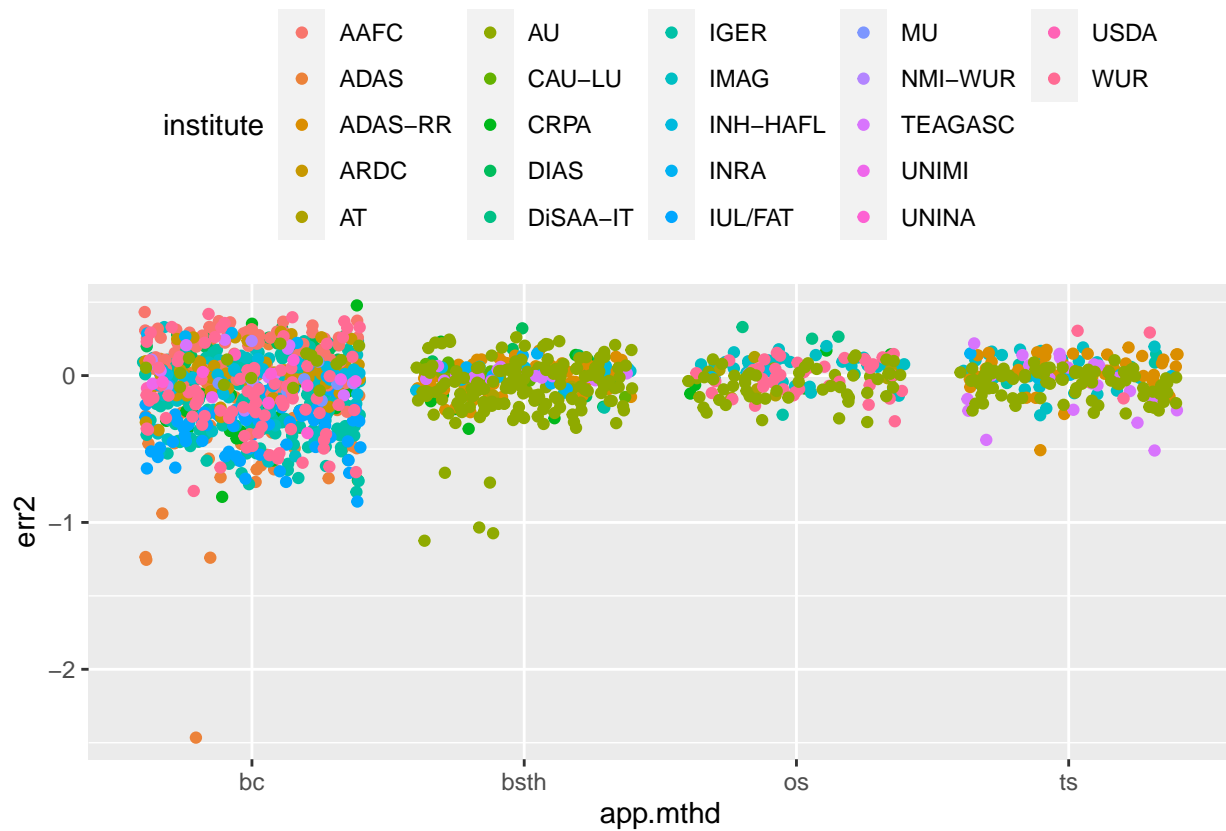
```
ggplot(dfinal, aes(app.mthd, e.rel, colour = institute)) +  
  geom_jitter(height = 0) +  
  theme(legend.position = 'top')
```

Warning: Removed 33 rows containing missing values (`geom_point()`).



```
ggplot(dfinal, aes(app.mthd, err2, colour = institute)) +  
  geom_jitter(height = 0) +  
  theme(legend.position = 'top')
```

Warning: Removed 33 rows containing missing values (`geom_point()`).



Drop values > 100% applied TAN.

```
dim(dfinal)
```

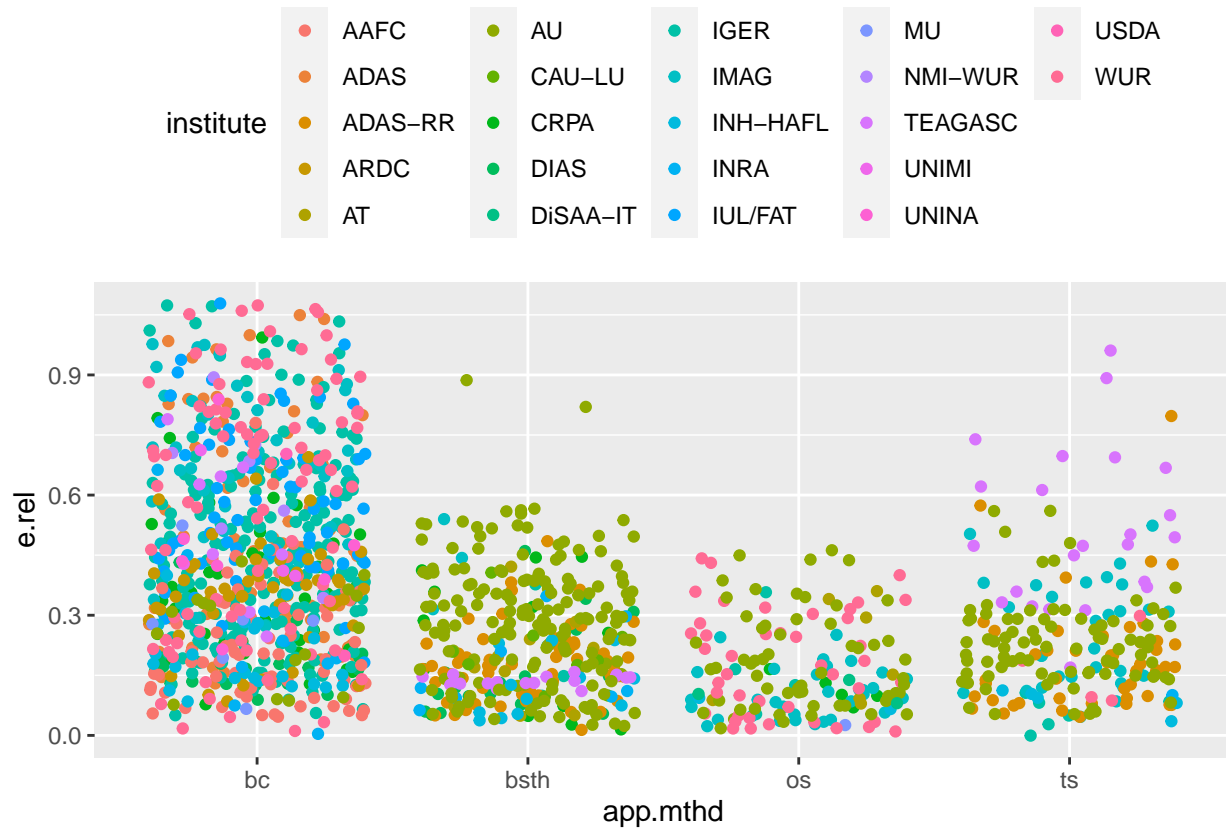
```
## [1] 1487 279
```

```
length(unique(dfinal$pmid))
```

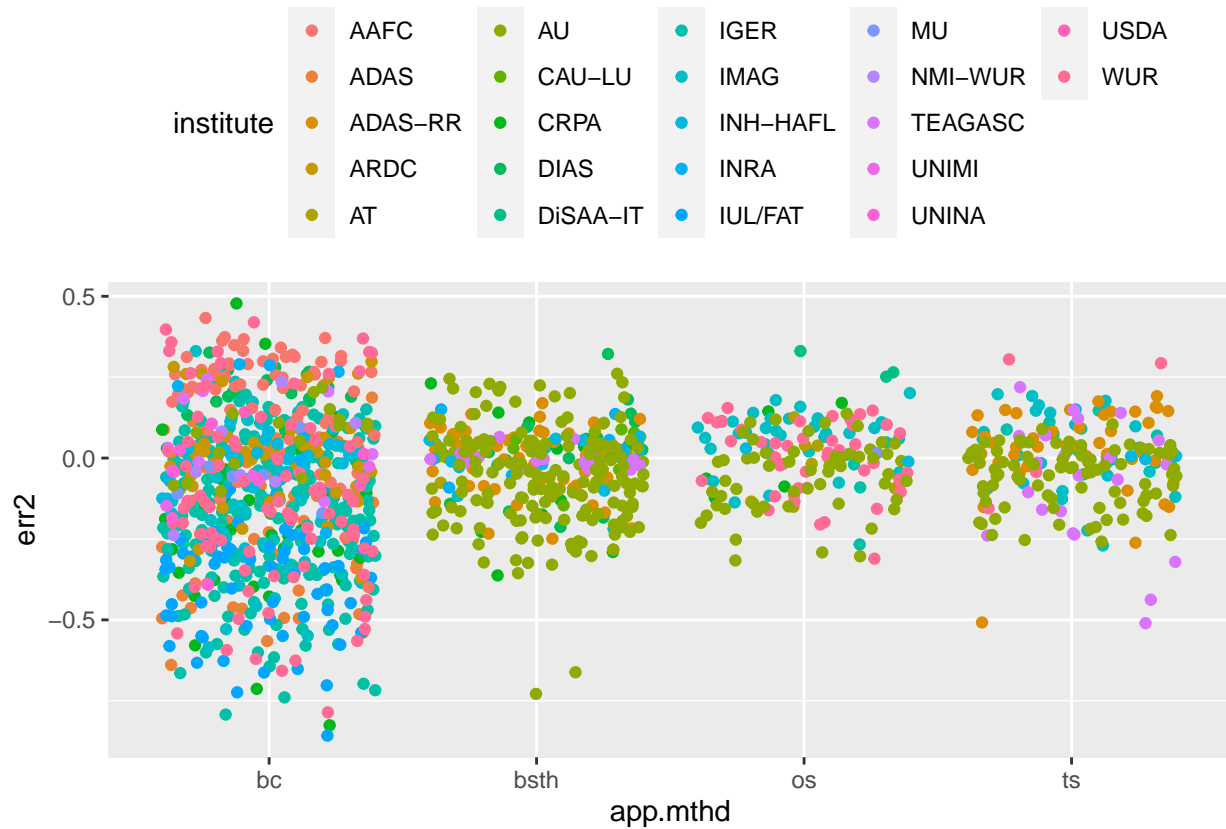
```
## [1] 1487
```

```
dfinal <- dfinal[e.rel < 1.1, ]
```

```
ggplot(dfinal, aes(app.mthd, e.rel, colour = institute)) +  
  geom_jitter(height = 0) +  
  theme(legend.position = 'top')
```



```
ggplot(dfinal, aes(app.mthd, err2, colour = institute)) +
  geom_jitter(height = 0) +
  theme(legend.position = 'top')
```



Total counts.

```
dim(dfinal)
```

```
## [1] 1441 279
```

```
length(unique(dfinal$pmid))
```

```
## [1] 1441
```

```
length(unique(dfinal$country))
```

```
## [1] 11
```

```
length(unique(dfinal$inst))
```

```
## [1] 22
```

```
unique(dfinal$country)
```

```
## [1] "UK" "IT" "DK" "NL" "CH" "CA" "DE" "FR" "IE" "US" "SE"
```

```
unique(dfinal$institute)
```

```
## [1] "ADAS" "CRPA" "DIAS" "IGER" "IMAG" "IUL/FAT"
## [7] "AAFC" "ADAS-RR" "ARDC" "AT" "AU" "CAU-LU"
## [13] "INH-HAFL" "INRA" "MU" "NMI-WUR" "TEAGASC" "USDA"
## [19] "WUR" "DiSAA-IT" "UNIMI" "UNINA"
```

```
unique(dfinal$inst)
```

```
## [1] 101 103 104 105 106 107 201 202 203 204 205 206 207 208 209 210 212 213 214
## [20] 303 304 305
```

2. Data prep

```
dfinal <- droplevels(dfinal[!is.na(e.rel), ])  
dfinal$inst <- factor(dfinal$inst)  
dfinal$inst.meas.tech <- interaction(dfinal$institute, dfinal$meas.tech)  
dfinal$app.mthd <- factor(dfinal$app.mthd)
```

Get subset without crazy broadcast

```
dfinalb <- dfinal[app.mthd != 'bc', ]
```

3. Basic variability and comparison of simplest predictors

```
m0 <- lmer(e.rel ~ (1|inst.meas.tech), data = dfinal)
```

```
m1 <- lmer(e.rel ~ app.mthd + (1|inst.meas.tech), data = dfinal)
```

```
m2 <- lm(e.rel ~ app.mthd, data = dfinal)
```

```
AIC(m0, m1, m2)
```

```
##      df      AIC  
## m0   3 -421.8459  
## m1   6 -678.7292  
## m2   5 -438.8121
```

```
summary(m0)
```

```
## Linear mixed model fit by REML ['lmerMod']  
## Formula: e.rel ~ (1 | inst.meas.tech)  
## Data: dfinal  
##  
## REML criterion at convergence: -427.8  
##  
## Scaled residuals:  
##      Min       1Q   Median       3Q      Max   
## -2.1406 -0.6467 -0.1080  0.5530  3.3263   
##  
## Random effects:  
## Groups      Name      Variance Std.Dev.  
## inst.meas.tech (Intercept) 0.01762  0.1327  
## Residual                0.04119  0.2030  
## Number of obs: 1441, groups: inst.meas.tech, 37  
##  
## Fixed effects:  
##              Estimate Std. Error t value  
## (Intercept)  0.32878    0.02539   12.95
```

```
summary(m1)
```

```
## Linear mixed model fit by REML ['lmerMod']  
## Formula: e.rel ~ app.mthd + (1 | inst.meas.tech)  
## Data: dfinal  
##  
## REML criterion at convergence: -690.7  
##
```

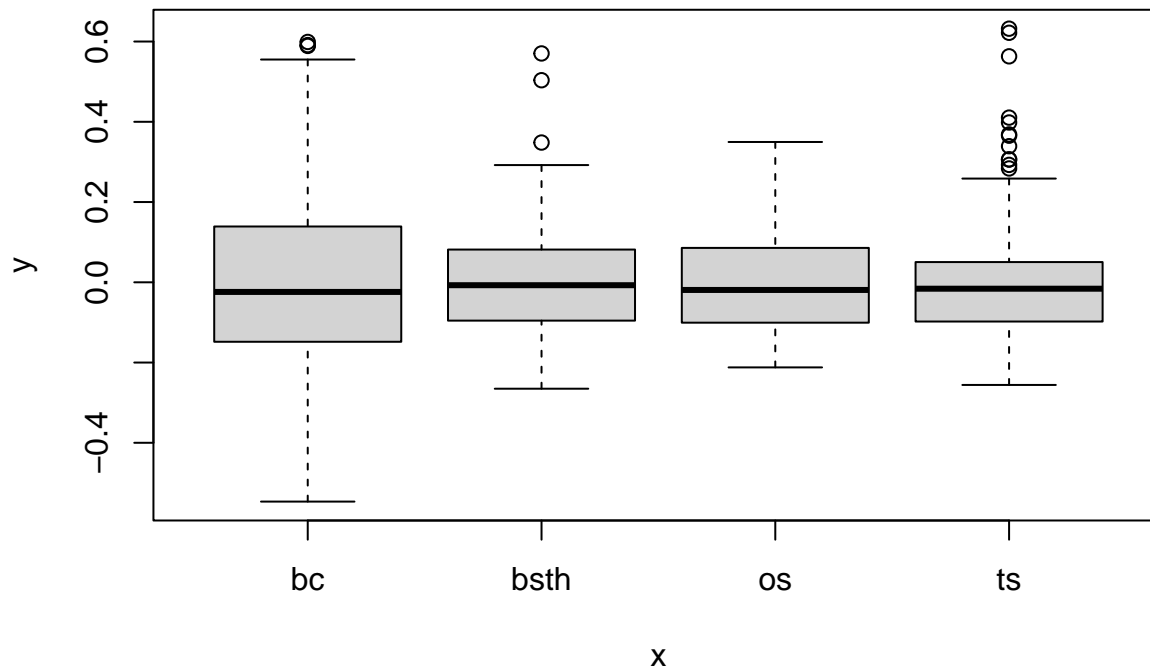
```
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.9593 -0.6684 -0.0838  0.5823  3.4226
##
## Random effects:
##   Groups             Name             Variance Std.Dev.
##   inst.meas.tech (Intercept) 0.01128  0.1062
##   Residual                0.03410  0.1847
## Number of obs: 1441, groups:  inst.meas.tech, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)   0.43034    0.02195   19.60
## app.mthdbsth -0.19508    0.01911  -10.21
## app.mthdos    -0.33559    0.01954  -17.17
## app.mthdts    -0.21475    0.02057  -10.44
##
## Correlation of Fixed Effects:
##              (Intr) app.mthdb app.mthds
## app.mthdbsth -0.291
## app.mthdos   -0.227  0.436
## app.mthdts   -0.236  0.616    0.401
```

```
summary(m2)
```

```
##
## Call:
## lm(formula = e.rel ~ app.mthd, data = dfinal)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.44340 -0.13681 -0.03545  0.12378  0.71341
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)   0.447225    0.007547   59.26  <2e-16 ***
## app.mthdbsth -0.212243    0.013699  -15.49  <2e-16 ***
## app.mthdos    -0.284969    0.018189  -15.67  <2e-16 ***
## app.mthdts    -0.199741    0.016491  -12.11  <2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.2074 on 1437 degrees of freedom
## Multiple R-squared:  0.235, Adjusted R-squared:  0.2334
## F-statistic: 147.1 on 3 and 1437 DF,  p-value: < 2.2e-16
```

So, institute x measurement technique effect is around 12% of applied TAN (from model m1). Residuals are large, around 20% of applied TAN. Presumably residuals are smaller for injection.

```
res <- resid(m1)
plot(dfinal$app.mthd, res)
```



Perhaps, but could be worse.

Repeat without broadcast

```
m0b <- lmer(e.rel ~ (1|inst.meas.tech), data = dfinalb)
```

```
m1b <- lmer(e.rel ~ app.mthd + (1|inst.meas.tech), data = dfinalb)
```

```
m2b <- lm(e.rel ~ app.mthd, data = dfinalb)
```

```
AIC(m0b, m1b, m2b)
```

```
##      df      AIC
## m0b  3 -722.8347
## m1b  5 -733.1104
## m2b  4 -702.5938
```

```
summary(m0b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ (1 | inst.meas.tech)
## Data: dfinalb
##
## REML criterion at convergence: -728.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.7406 -0.7082 -0.1489  0.5329  4.5316
##
## Random effects:
## Groups      Name                Variance Std.Dev.
## inst.meas.tech (Intercept) 0.003508 0.05922
## Residual                0.019267 0.13881
## Number of obs: 686, groups: inst.meas.tech, 19
##
```

```
## Fixed effects:
##           Estimate Std. Error t value
## (Intercept)  0.19900    0.01659     12

summary(m1b)

## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ app.mthd + (1 | inst.meas.tech)
## Data: dfinalb
##
## REML criterion at convergence: -743.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.8037 -0.7134 -0.1433  0.5686  4.5847
##
## Random effects:
## Groups           Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.003163 0.05624
## Residual                0.018580 0.13631
## Number of obs: 686, groups:  inst.meas.tech, 19
##
## Fixed effects:
##           Estimate Std. Error t value
## (Intercept)  0.221512    0.016987  13.040
## app.mthdos   -0.077669    0.016031  -4.845
## app.mthdts    0.004581    0.013043   0.351
##
## Correlation of Fixed Effects:
##           (Intr) app.mthds
## app.mthdos -0.330
## app.mthdts -0.240  0.347
```

```
summary(m2b)
```

```
##
## Call:
## lm(formula = e.rel ~ app.mthd, data = dfinalb)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -0.24808 -0.10570 -0.03087  0.07986  0.71341
##
## Coefficients:
##           Estimate Std. Error t value Pr(>|t|)
## (Intercept)  0.234982    0.007965  29.502 < 2e-16 ***
## app.mthdos   -0.072725    0.014014  -5.190 2.78e-07 ***
## app.mthdts    0.012503    0.012954   0.965  0.335
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 0.1445 on 683 degrees of freedom
## Multiple R-squared:  0.04971, Adjusted R-squared:  0.04693
## F-statistic: 17.86 on 2 and 683 DF, p-value: 2.741e-08
```

Less variability without broadcast.

4. ALFAM2 model residuals

```
m3 <- lmer(err2 ~ (1|inst.meas.tech), data = dfinal)
```

```
m4 <- lmer(err2 ~ app.mthd + (1|inst.meas.tech), data = dfinal)
```

```
summary(m0)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ (1 | inst.meas.tech)
## Data: dfinal
##
## REML criterion at convergence: -427.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.1406 -0.6467 -0.1080  0.5530  3.3263
##
## Random effects:
## Groups          Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.01762  0.1327
## Residual              0.04119  0.2030
## Number of obs: 1441, groups: inst.meas.tech, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.32878    0.02539   12.95
```

```
summary(m1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ app.mthd + (1 | inst.meas.tech)
## Data: dfinal
##
## REML criterion at convergence: -690.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.9593 -0.6684 -0.0838  0.5823  3.4226
##
## Random effects:
## Groups          Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.01128  0.1062
## Residual              0.03410  0.1847
## Number of obs: 1441, groups: inst.meas.tech, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.43034    0.02195   19.60
## app.mthdbsth -0.19508    0.01911  -10.21
## app.mthdos   -0.33559    0.01954  -17.17
## app.mthdts   -0.21475    0.02057  -10.44
##
## Correlation of Fixed Effects:
##              (Intr) app.mthdb app.mthds
```

```
## app.mthdbst -0.291
## app.mthdos -0.227 0.436
## app.mthdts -0.236 0.616 0.401
```

```
summary(m3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: err2 ~ (1 | inst.meas.tech)
## Data: dfinal
##
## REML criterion at convergence: -902.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.1054 -0.5388  0.1000  0.6074  3.5407
##
## Random effects:
## Groups           Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.01207  0.1098
## Residual              0.02965  0.1722
## Number of obs: 1441, groups: inst.meas.tech, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.03566    0.02111  -1.689
```

```
summary(m4)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: err2 ~ app.mthd + (1 | inst.meas.tech)
## Data: dfinal
##
## REML criterion at convergence: -904.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -3.9679 -0.5474  0.0938  0.5880  3.6548
##
## Random effects:
## Groups           Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.01161  0.1077
## Residual              0.02929  0.1711
## Number of obs: 1441, groups: inst.meas.tech, 37
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.05802    0.02173  -2.670
## app.mthdbsth  0.03633    0.01783   2.037
## app.mthdos    0.08261    0.01818   4.544
## app.mthdts    0.04962    0.01915   2.592
##
## Correlation of Fixed Effects:
##              (Intr) app.mthdb app.mthds
## app.mthdbst -0.272
## app.mthdos  -0.214  0.439
```

```
## app.mthdts -0.220 0.619 0.404
```

```
AIC(m3, m4)
```

```
##      df      AIC
## m3  3 -896.5180
## m4  6 -892.7441
```

Reassuring that m3 is actually a better model than m4, meaning adding application method on top of ALFAM2 predictions doesn't help.

Again, exclude broadcast.

```
m3b <- lmer(err2 ~ (1|inst.meas.tech), data = dfinalb)
```

```
m4b <- lmer(err2 ~ app.mthd + (1|inst.meas.tech), data = dfinalb)
```

```
summary(m0b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ (1 | inst.meas.tech)
##      Data: dfinalb
##
## REML criterion at convergence: -728.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.7406 -0.7082 -0.1489  0.5329  4.5316
##
## Random effects:
##   Groups             Name           Variance Std.Dev.
##   inst.meas.tech (Intercept) 0.003508 0.05922
##   Residual                0.019267 0.13881
## Number of obs: 686, groups:  inst.meas.tech, 19
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept)  0.19900    0.01659     12
```

```
summary(m1b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ app.mthd + (1 | inst.meas.tech)
##      Data: dfinalb
##
## REML criterion at convergence: -743.1
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -1.8037 -0.7134 -0.1433  0.5686  4.5847
##
## Random effects:
##   Groups             Name           Variance Std.Dev.
##   inst.meas.tech (Intercept) 0.003163 0.05624
##   Residual                0.018580 0.13631
## Number of obs: 686, groups:  inst.meas.tech, 19
##
## Fixed effects:
```

```
##           Estimate Std. Error t value
## (Intercept)  0.221512   0.016987  13.040
## app.mthdos   -0.077669   0.016031  -4.845
## app.mthdts   0.004581   0.013043   0.351
##
## Correlation of Fixed Effects:
##           (Intr) app.mthds
## app.mthdos -0.330
## app.mthdts -0.240  0.347
```

```
summary(m3b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: err2 ~ (1 | inst.meas.tech)
## Data: dfinalb
##
## REML criterion at convergence: -900.2
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.8657 -0.5597  0.1550  0.6170  2.5881
##
## Random effects:
## Groups           Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.00298  0.05459
## Residual              0.01498  0.12239
## Number of obs: 686, groups:  inst.meas.tech, 19
##
## Fixed effects:
##           Estimate Std. Error t value
## (Intercept) -0.008732   0.015141  -0.577
```

```
summary(m4b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: err2 ~ app.mthd + (1 | inst.meas.tech)
## Data: dfinalb
##
## REML criterion at convergence: -886.8
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.8666 -0.5765  0.1516  0.6033  2.6082
##
## Random effects:
## Groups           Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.002823  0.05313
## Residual              0.015029  0.12259
## Number of obs: 686, groups:  inst.meas.tech, 19
##
## Fixed effects:
##           Estimate Std. Error t value
## (Intercept) -0.012550   0.015805  -0.794
## app.mthdos   0.009106   0.014461   0.630
## app.mthdts   0.005060   0.011744   0.431
```

```
##
## Correlation of Fixed Effects:
##      (Intr) app.mthds
## app.mthdos -0.321
## app.mthdts -0.232  0.348
```

```
AIC(m3b, m4b)
```

```
##      df      AIC
## m3b  3 -894.2066
## m4b  5 -876.8290
```

Here too, m3b is better.

5. “Institution effect”

Our best estimate of an “institution effect” is from m3, where we have corrected for different application methods, manure DM, pH, and weather using the ALFAM2 model.

```
VarCorr(m1)
```

```
## Groups      Name      Std.Dev.
## inst.meas.tech (Intercept) 0.10621
## Residual              0.18467
```

```
VarCorr(m1b)
```

```
## Groups      Name      Std.Dev.
## inst.meas.tech (Intercept) 0.056243
## Residual              0.136309
```

```
VarCorr(m3)
```

```
## Groups      Name      Std.Dev.
## inst.meas.tech (Intercept) 0.10984
## Residual              0.17220
```

```
VarCorr(m3b)
```

```
## Groups      Name      Std.Dev.
## inst.meas.tech (Intercept) 0.054594
## Residual              0.122390
```

```
summary(m1)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: e.rel ~ app.mthd + (1 | inst.meas.tech)
##      Data: dfinal
##
## REML criterion at convergence: -690.7
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -2.9593 -0.6684 -0.0838  0.5823  3.4226
##
## Random effects:
##      Groups      Name      Variance Std.Dev.
## inst.meas.tech (Intercept) 0.01128  0.1062
## Residual              0.03410  0.1847
```

```
## Number of obs: 1441, groups: inst.meas.tech, 37
##
## Fixed effects:
##           Estimate Std. Error t value
## (Intercept)  0.43034    0.02195   19.60
## app.mthdbsth -0.19508    0.01911  -10.21
## app.mthdos   -0.33559    0.01954  -17.17
## app.mthdts   -0.21475    0.02057  -10.44
##
## Correlation of Fixed Effects:
##           (Intr) app.mthdb app.mthds
## app.mthdbsth -0.291
## app.mthdos   -0.227  0.436
## app.mthdts   -0.236  0.616    0.401
```

```
VarCorr(m1b)
```

```
## Groups          Name          Std.Dev.
## inst.meas.tech (Intercept) 0.056243
## Residual              0.136309
```

```
summary(m3)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: err2 ~ (1 | inst.meas.tech)
## Data: dfinal
##
## REML criterion at convergence: -902.5
##
## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.1054 -0.5388  0.1000  0.6074  3.5407
##
## Random effects:
## Groups          Name          Variance Std.Dev.
## inst.meas.tech (Intercept) 0.01207  0.1098
## Residual              0.02965  0.1722
## Number of obs: 1441, groups: inst.meas.tech, 37
##
## Fixed effects:
##           Estimate Std. Error t value
## (Intercept) -0.03566    0.02111  -1.689
```

```
VarCorr(m3b)
```

```
## Groups          Name          Std.Dev.
## inst.meas.tech (Intercept) 0.054594
## Residual              0.122390
```

```
summary(m3b)
```

```
## Linear mixed model fit by REML ['lmerMod']
## Formula: err2 ~ (1 | inst.meas.tech)
## Data: dfinalb
##
## REML criterion at convergence: -900.2
##
```

```

## Scaled residuals:
##      Min       1Q   Median       3Q      Max
## -4.8657 -0.5597  0.1550  0.6170  2.5881
##
## Random effects:
##   Groups             Name             Variance Std.Dev.
##  inst.meas.tech (Intercept) 0.00298  0.05459
##   Residual                  0.01498  0.12239
## Number of obs: 686, groups:  inst.meas.tech, 19
##
## Fixed effects:
##              Estimate Std. Error t value
## (Intercept) -0.008732  0.015141  -0.577

```