Meta-Analysis for Longevity Summary Excluding HUM251

Fay Frost

1. Summary

This document reports the process taken in the model fitting stage of the meta-analysis in thermal longevity.

2. Setup

We first read in our data and select all of the effect sizes related to longevity. We do this using the following code.

```r
### Read in effect size data
effectdata <- read.csv("Survival project all pairwise.es.csv")
longdata_warm <- subset(effectdata, Trait.category == "Longevity" &
    warm.cool == "Warm")
longdata_cool <- subset(effectdata, Trait.category == "Longevity" &
    warm.cool == "Cool")

alllong <- rbind(longdata_warm, longdata_cool)

### select data for analysis
rdata <- alllong

rdata <- subset(rdata, Paper.code != "HUM251")

rdata <- rdata %>%
    mutate(c_treattemp = treattemp - 25)
```

Next we create new columns in our dataframe which will serve as random factors in our multi-level meta analysis models. The following initialises four new columns, namely "obs", "study_code", "Species.phylo" and "species". Lastly, we create a column name "precision" which is equal to the inverse standard error.

```r
### Create random factors into data frame
rdata$obs <- factor(c(1:nrow(rdata)))  # Unique observation code
rdata$study_code <- factor(rdata$Paper.code)  # Model requires column names study_code
rdata$Species.phylo <- factor(rdata$Species.latin)  # Species names for phylo matrix
rdata$species <- factor(rdata$Species.latin)  # Another species column for random factor

precision <- sqrt(1/rdata$v)  # inverse standard error
rdata[, "precision"] <- precision
```

The number of species and total number of studies present in the data are as follows.

```r
nlevels(rdata$species)  # Check number of species
```

```
## [1] 289
```

```r
nlevels(rdata$study_code)  # Check number of studies
```

```
## [1] 315
```

The final stage in the setup is to import a phylogentic tree of the data. Below is the code used to produce the tree and a plot of the tree itself.

```
## import tree from map
tree1 <- read.nexus("all_longevity_excHUM251_tree.nex")
tree_grafen = compute.brlen(tree1, method = "Grafen", power = 1)
phylo_matrix <- vcv(tree_grafen, cor = TRUE, model = "Brownian")  # Make phylogenetic matrix
```

character(0)
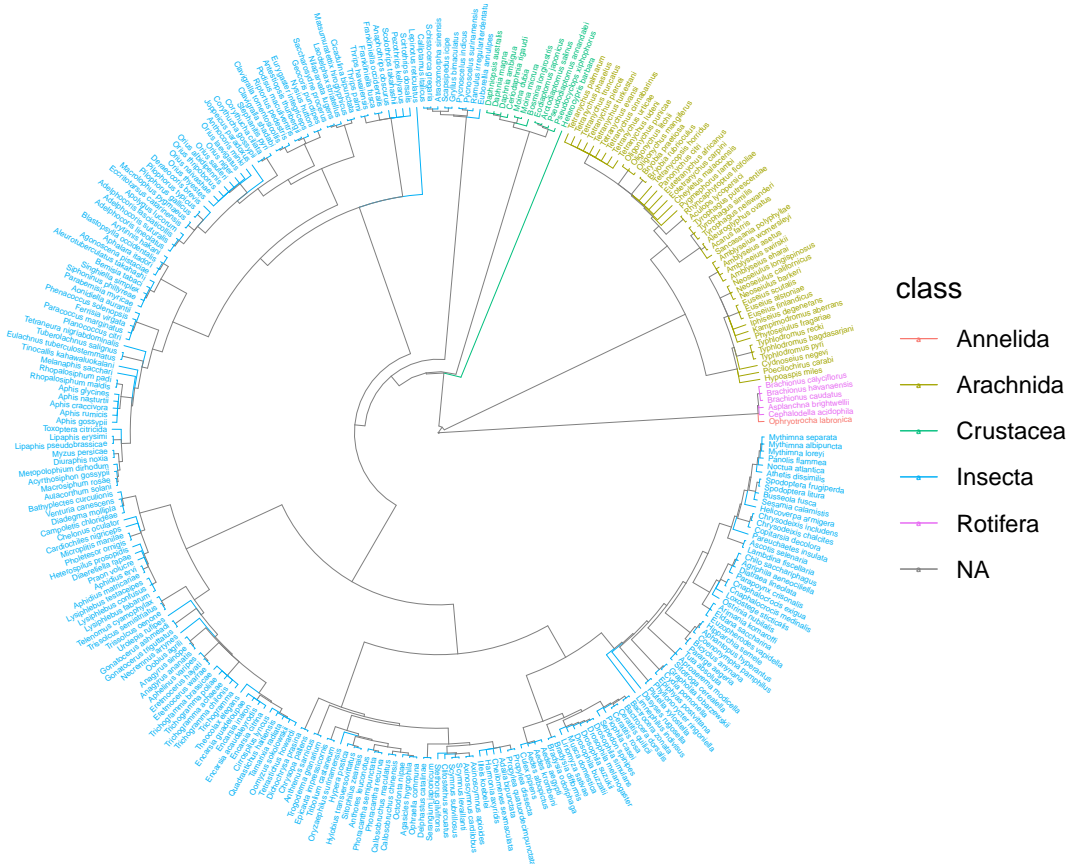
3. Random effects models

In this section we determine which random effects to include in our model. For each model I have provided the code used to specify the structure of the model and a summary of the results. We begin with a model that includes all of the random factors we created earlier.

```
# Adding four random factors
meta2 <- rma.mv(es, v, random = list(~1 | Species.phylo, ~1 |
    species, ~1 | study_code, ~1 | obs), R = list(Species.phylo = phylo_matrix),
    test = "t", dfs = "contain", data = rdata, method = "REML")
```

```
summary(meta2)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik     Deviance          AIC          BIC         AICc
## -3561.6881    7123.3763    7133.3763    7159.5110    7133.4201
##
## Variance Components:
##
##            estim    sqrt   nlvls  fixed          factor    R
## sigma^2.1  0.0000  0.0000    289     no   Species.phylo  yes
## sigma^2.2  0.0000  0.0003    289     no         species   no
## sigma^2.3  0.9600  0.9798    315     no      study_code   no
## sigma^2.4  7.6881  2.7727   1377     no             obs   no
##
## Test for Heterogeneity:
## Q(df = 1376) = 51290.4596, p-val < .0001
##
## Model Results:
##
## estimate      se     zval     pval    ci.lb    ci.ub
##  -0.1761  0.0983  -1.7919   0.0732  -0.3687   0.0165   .
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
i2_ml(meta2, method = c("ratio"))  # Heterogeneity at each random factor level
```

```
##        I2_Total I2_Species.phylo      I2_species   I2_study_code        I2_obs
##    9.908349e+01     2.668549e-08    8.400439e-07    1.099874e+01  8.808476e+01
```

Accounting for non-independence of data points from the same experiment

The data has a nested structure. Each study (study_code) may have a number of experiments (effect.size.code) which share a common control temperature. Each effect size has its own unique code, obs. Effect sizes from the same experiment which share a control temperature are thought to be non-independent. The following code create a covariance matrix "VCV_shared" which assumes a correlation of 0.5 between effect sizes from the same experiment. We include this structure in our proceeding models.

```
rdata$shared_control <- factor(rdata$Effect.size.code)
vcv_shared <- impute_covariance_matrix(vi = rdata$v, cluster = rdata$shared_control,
    r = 0.5)
```

```
# Add new variance matrix into the mixed-effects
# meta-analysis model
```

```r
meta3 <- rma.mv(es, vcv_shared, random = list(~1 | Species.phylo,
    ~1 | species, ~1 | study_code, ~1 | obs), test = "t", dfs = "contain",
    R = list(Species.phylo = phylo_matrix), data = rdata, method = "REML")
```

```r
summary(meta3)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3550.7026   7101.4053   7111.4053   7137.5399   7111.4491
##
## Variance Components:
##
##               estim     sqrt   nlvls   fixed           factor     R
## sigma^2.1    0.0000   0.0000     289      no   Species.phylo   yes
## sigma^2.2    0.0000   0.0003     289      no         species    no
## sigma^2.3    0.6854   0.8279     315      no      study_code    no
## sigma^2.4    7.9521   2.8200    1377      no             obs    no
##
## Test for Heterogeneity:
## Q(df = 1376) = 69509.3558, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval     ci.lb     ci.ub
##  -0.1880   0.0950  -1.9800   0.0477   -0.3742   -0.0019  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```r
i2_ml(meta3, method = c("ratio"))   # Heterogeneity at each random factor level
```

```
##        I2_Total I2_Species.phylo      I2_species   I2_study_code          I2_obs
##    9.908239e+01     6.530133e-11    7.656187e-07    7.862756e+00    9.121964e+01
```

Model without phylogeny

The variance-covariance matrix for phylogenetic relatedness of included species has now been excluded as a random effect in the model (Chamberlain et al., 2012) as its inclusion did not improve model fit and the phylogenetic signal was very weak.

```
## without phylogeny
meta5 <- rma.mv(es, VCV_shared, random = list(~1 | species, ~1 |
    study_code, ~1 | obs), test = "t", dfs = "contain", data = rdata,
    method = "REML")
```

```
summary(meta5)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3550.7026   7101.4053   7109.4053   7130.3130   7109.4344
##
## Variance Components:
##
##             estim     sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000   0.0003    289     no     species
## sigma^2.2  0.6854   0.8279    315     no  study_code
## sigma^2.3  7.9521   2.8200   1377     no         obs
##
## Test for Heterogeneity:
## Q(df = 1376) = 69509.3558, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb    ci.ub
##  -0.1880   0.0950  -1.9800   0.0477  -0.3742  -0.0019  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
i2_ml(meta5, method = c("ratio"))   # Heterogeneity at each random factor level
```

```
##      I2_Total    I2_species I2_study_code        I2_obs
## 9.908239e+01  9.359070e-07  7.862751e+00  9.121964e+01
```

Model without phylogeny or species

```
## without phylogeny or species
meta4 <- rma.mv(es, VCV_shared, random = list(~1 | study_code,
    ~1 | obs), data = rdata, test = "t", dfs = "contain", method = "REML")
```

```
summary(meta4)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC          BIC        AICc
## -3550.7026   7101.4053   7107.4053   7123.0861   7107.4227
##
## Variance Components:
##
##              estim    sqrt  nlvls  fixed      factor
## sigma^2.1   0.6854  0.8279    315     no  study_code
## sigma^2.2   7.9521  2.8200   1377     no         obs
##
## Test for Heterogeneity:
## Q(df = 1376) = 69509.3558, p-val < .0001
##
## Model Results:
##
## estimate      se     zval     pval    ci.lb     ci.ub
##  -0.1880  0.0950  -1.9800   0.0477  -0.3742   -0.0019  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
i2_ml(meta4, method = c("ratio"))  # Heterogeneity at each random factor level
```

```
##      I2_Total I2_study_code       I2_obs
##     99.082392      7.862751    91.219641
```

Model without phylogeny, species or study_code

```
## without phylogeny, species or study_code
meta7 <- rma.mv(es, VCV_shared, random = list(~1 | obs), data = rdata,
    test = "t", dfs = "contain", method = "REML")
```

```
summary(meta7)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3569.6777   7139.3555   7143.3555   7153.8094   7143.3642
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed  factor
## sigma^2    8.7259  2.9540   1377     no     obs
##
## Test for Heterogeneity:
## Q(df = 1376) = 69509.3558, p-val < .0001
##
## Model Results:
##
## estimate       se     zval     pval    ci.lb     ci.ub
##  -0.2031   0.0819  -2.4800   0.0131  -0.3637  -0.0426  *
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
i2_ml(meta7, method = c("ratio"))  # Heterogeneity at each random factor level
```

```
## I2_Total    I2_obs
##  99.0916   99.0916
```

We can see from the above that the best fitting model according to AIC is "meta4" which includes only the study code and the unique effect size code, obs. There is a AIC difference of 4 between the model meta4 and the next best model meta3 . We continue our analysis using meta4 as our base model.

4. Meta-regressions

Starting with the best fitting random-effect model from Section 3, "meta4" we now include single factors as a fixed effect. We initially explore the fixed factors

- **reftemp**: The experiment's control (reference) temperature.

- **treattemp**: The treatment temperature, which we expect to havea non-linear relationship to longevity.

- **warm.cool** : A categorical variable indicating whether treatment is warmer or cooler than the reference temperature

- **diff**: The difference between the reference and treatment temperature.

Reference temperature

```
meta_trait_ref <- rma.mv(es, VCV_shared, mod = ~reftemp, random = list(~1 |
    study_code, ~1 | obs), test = "t", dfs = "contain", data = rdata,
    method = "REML")
```

```
summary(meta_trait_ref)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3547.8024   7095.6047   7103.6047   7124.5095   7103.6339
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed       factor
## sigma^2.1  0.6777  0.8232    315     no   study_code
## sigma^2.2  7.9551  2.8205   1377     no          obs
##
## Test for Residual Heterogeneity:
## QE(df = 1375) = 69447.2265, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 1.4962, p-val = 0.2213
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt   -1.2889  0.9049  -1.4243  0.1544  -3.0626  0.4848
## reftemp    0.0443  0.0362   1.2232  0.2213  -0.0267  0.1152
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Treatment temperature

```
meta_trait_treattemp <- rma.mv(es, VCV_shared, mod = ~c_treattemp,
    random = list(~1 | study_code, ~1 | obs), test = "t", dfs = "contain",
    data = rdata, method = "REML")
```

```
summary(meta_trait_treattemp)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3348.5884   6697.1767   6705.1767   6726.0816   6705.2059
##
## Variance Components:
##
##            estim    sqrt  nlvls  fixed      factor
## sigma^2.1  1.1371  1.0664    315     no  study_code
## sigma^2.2  5.3222  2.3070   1377     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1375) = 48746.5750, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 490.4429, p-val < .0001
##
## Model Results:
##
##              estimate      se     zval     pval    ci.lb    ci.ub
## intrcpt       -0.3183  0.0929  -3.4280   0.0006  -0.5003  -0.1363  ***
## c_treattemp   -0.1917  0.0087 -22.1459   <.0001  -0.2086  -0.1747  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Warm vs Cool

```
meta_trait_warm <- rma.mv(es, VCV_shared, mod = ~warm.cool, random = list(~1 |
    study_code, ~1 | obs), test = "t", dfs = "contain", data = rdata,
    method = "REML")
```

```
summary(meta_trait_warm)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3318.5464   6637.0928   6645.0928   6665.9977   6645.1220
##
## Variance Components:
##
##            estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.6289  0.7930    315     no  study_code
## sigma^2.2  5.2918  2.3004   1377     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1375) = 46521.5595, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 555.7125, p-val < .0001
##
## Model Results:
##
##               estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt         1.3480  0.1050  12.8393  <.0001   1.1422   1.5537  ***
## warm.coolWarm  -3.1023  0.1316 -23.5736  <.0001  -3.3602  -2.8443  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We model warm versus cool without and intercept so we can visualise the estimates easier.

```
meta_trait_warm_nointer <- rma.mv(es, VCV_shared, mod = ~warm.cool -
    1, random = list(~1 | study_code, ~1 | obs), data = rdata,
    test = "t", dfs = "contain", method = "REML")
```

```
summary(meta_trait_warm_nointer)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3318.5464   6637.0928   6645.0928   6665.9977   6645.1220
##
## Variance Components:
##
##            estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.6289  0.7930    315     no  study_code
## sigma^2.2  5.2918  2.3004   1377     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1375) = 46521.5595, p-val < .0001
##
## Test of Moderators (coefficients 1:2):
## QM(df = 2) = 561.2102, p-val < .0001
##
## Model Results:
##
##                 estimate       se      zval     pval    ci.lb    ci.ub
## warm.coolCool     1.3480   0.1050   12.8393   <.0001   1.1422   1.5537  ***
## warm.coolWarm    -1.7543   0.1055  -16.6214   <.0001  -1.9612  -1.5474  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Difference

```
meta_trait_diff <- rma.mv(es, VCV_shared, mod = ~diff, random = list(~1 |
    study_code, ~1 | obs), data = rdata, test = "t", dfs = "contain",
    method = "REML")
```

```
summary(meta_trait_diff)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3334.7459   6669.4918   6677.4918   6698.3967   6677.5210
##
## Variance Components:
##
##              estim    sqrt  nlvls  fixed      factor
## sigma^2.1   0.9900  0.9950    315     no  study_code
## sigma^2.2   5.2579  2.2930   1377     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1375) = 46413.4259, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 522.6694, p-val < .0001
##
## Model Results:
##
##           estimate      se      zval    pval    ci.lb    ci.ub
## intrcpt    -0.2886  0.0897   -3.2190  0.0013  -0.4644  -0.1129    **
## diff       -0.1981  0.0087  -22.8620  <.0001  -0.2151  -0.1812   ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Treatment temperature as a cubic effect

We expect that treatment temperature will have a non-linear effect on longevity. We expect that the relationship will be cubic. More specifically, we expect that longevity will increase at cooler temperature and decrease at more extreme higher temperature. We also assume that this optimum temperature will be close to the control temperature. That said, we investigate **treattemp** as a cubic fixed effect.

```
meta_trait_treat3 <- rma.mv(es, VCV_shared, mod = ~poly(c_treattemp,
    degree = 3, raw = TRUE), random = list(~1 | study_code, ~1 |
    obs), test = "t", dfs = "contain", data = rdata, method = "REML")
```

```
summary(meta_trait_treat3)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3281.5199    6563.0398   6575.0398   6606.3883   6575.1013
##
## Variance Components:
##
##            estim     sqrt   nlvls   fixed      factor
## sigma^2.1  0.6689   0.8178     315      no   study_code
## sigma^2.2  5.0068   2.2376    1377      no          obs
##
## Test for Residual Heterogeneity:
## QE(df = 1373) = 45334.9453, p-val < .0001
##
## Test of Moderators (coefficients 2:4):
## QM(df = 3) = 649.9766, p-val < .0001
##
## Model Results:
##
##                                        estimate      se      zval     pval     ci.lb     ci.ub
## intrcpt                                 -0.0646  0.0989   -0.6530   0.5137   -0.2585    0.1293
## poly(c_treattemp, degree = 3, raw = TRUE)1   -0.2895  0.0125  -23.1796   <.0001   -0.3139   -0.2650  **
## poly(c_treattemp, degree = 3, raw = TRUE)2   -0.0032  0.0009   -3.7662   0.0002   -0.0049   -0.0015  **
## poly(c_treattemp, degree = 3, raw = TRUE)3    0.0006  0.0001   10.4464   <.0001    0.0005    0.0008  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

For completeness we also investigate treatment temperature as a quadratic effect.

```
meta_trait_treat2 <- rma.mv(es, VCV_shared, mod = ~poly(c_treattemp,
    degree = 2, raw = TRUE), random = list(~1 | study_code, ~1 |
    obs), test = "t", dfs = "contain", data = rdata, method = "REML")
```

```
summary(meta_trait_treat2)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3332.9224    6665.8449   6675.8449   6701.9723   6675.8887
##
## Variance Components:
```

```
## 
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  1.1454  1.0702    315     no  study_code
## sigma^2.2  5.2016  2.2807   1377     no         obs
## 
## Test for Residual Heterogeneity:
## QE(df = 1374) = 48507.7186, p-val < .0001
## 
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 528.7874, p-val < .0001
## 
## Model Results:
## 
##                                        estimate      se     zval    pval   ci.lb   ci.ub
## intrcpt                                 -0.0112  0.1094  -0.1021  0.9187  -0.2257   0.2033
## poly(c_treattemp, degree = 2, raw = TRUE)1    -0.1990  0.0087  -22.9314  <.0001  -0.2160  -0.1820  **
## poly(c_treattemp, degree = 2, raw = TRUE)2    -0.0047  0.0009   -5.2603  <.0001  -0.0065  -0.0030  **
## 
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Below is a bubble plot of the fitted cubic model.
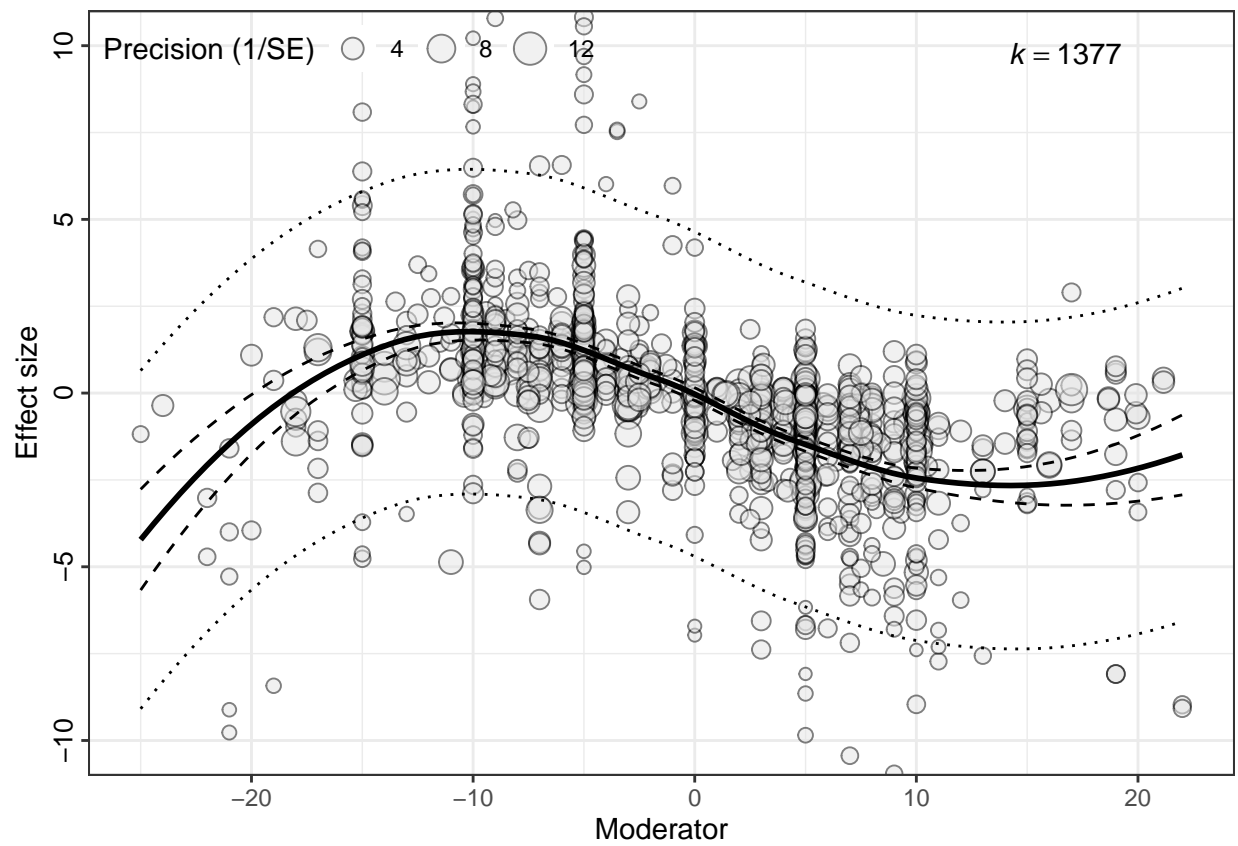
Modelling response with binned treatment temperatures

The last model we will try is one with categorised or 'binned' treatment temperatures. We categories our effect sizes into one seven bins depending on the treatment temperature. The bins and the number of effect sizes in each bin are given below

```
table(rdata$bin.temp)
```

```
##
##  <15  >40 15-20 20-25 25-30 30-35 35-40
##  110   53   249   315   188   354   108
```

```
meta_trait_bintemp <- rma.mv(es, VCV_shared, mod = ~bin.temp -
    1, random = list(~1 | study_code, ~1 | obs), data = rdata,
    method = "REML")
```

```
summary(meta_trait_bintemp)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3292.4098   6584.8197   6602.8197   6649.8228   6602.9520
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed     factor
## sigma^2.1  0.7464  0.8639    315     no  study_code
## sigma^2.2  5.1064  2.2597   1377     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1370) = 46239.8539, p-val < .0001
##
## Test of Moderators (coefficients 1:7):
## QM(df = 7) = 610.3853, p-val < .0001
##
## Model Results:
##
##                estimate      se     zval    pval    ci.lb    ci.ub
## bin.temp<15      0.9653  0.2508   3.8488  0.0001   0.4737   1.4568  ***
## bin.temp>40     -2.2404  0.3892  -5.7568  <.0001  -3.0032  -1.4777  ***
## bin.temp15-20    1.8022  0.1602  11.2503  <.0001   1.4883   2.1162  ***
## bin.temp20-25    1.1169  0.1439   7.7587  <.0001   0.8347   1.3990  ***
## bin.temp25-30   -0.8035  0.1830  -4.3900  <.0001  -1.1622  -0.4448  ***
## bin.temp30-35   -1.7818  0.1369 -13.0177  <.0001  -2.0501  -1.5136  ***
## bin.temp35-40   -2.7822  0.2407 -11.5565  <.0001  -3.2540  -2.3103  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
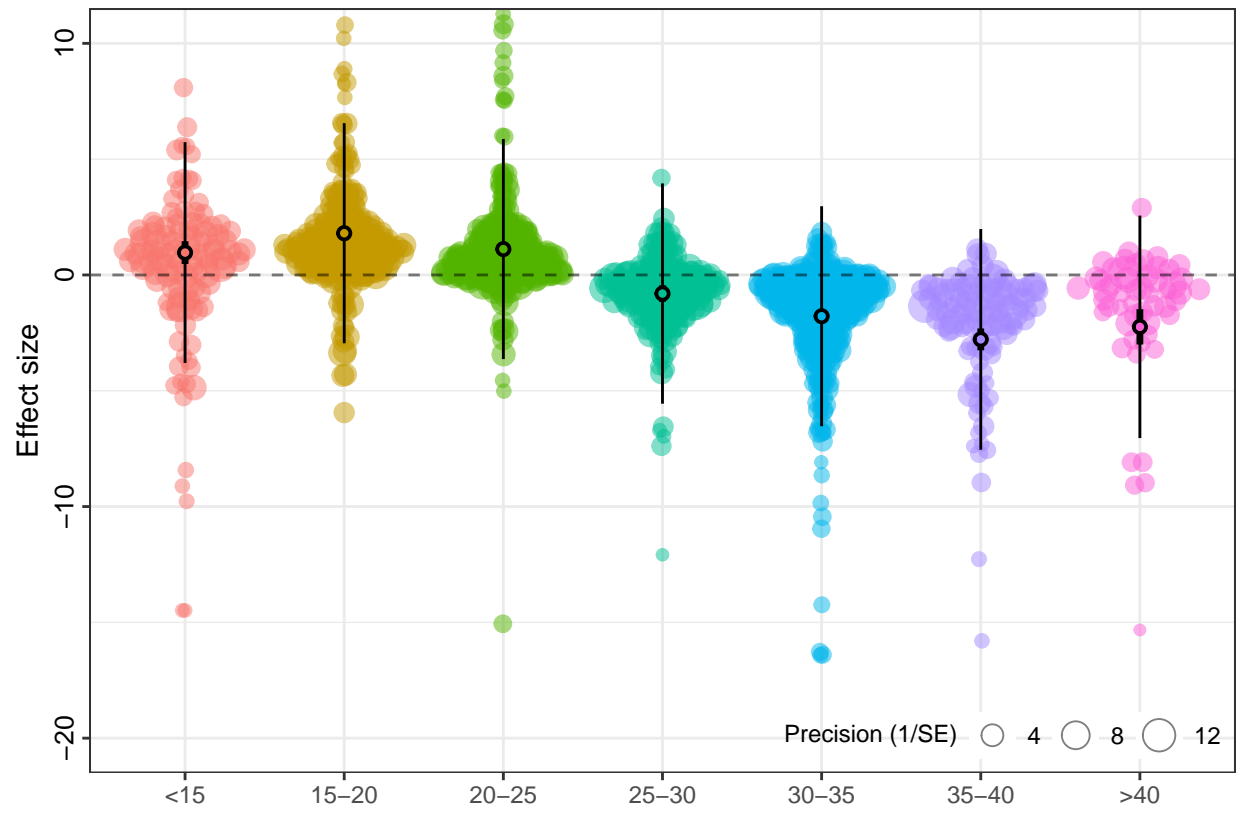
Other fixed effects

Here are tabled number of counts for each level of our moderators.

```
table(rdata$Class)
```

```
##
## Annelida Arachnida Crustacea  Insecta  Rotifera
##        2       219        28     1102        26
```

```
table(rdata$Habitat)
```

```
##
##    Aquatic        Both Terrestrial
##         59          63        1255
```

```
table(rdata$Sex.exposed)
```

```
##
##            Both         Female          Male Parthenogenetic         Unsure
##             569            599             4             202              3
```

```
table(rdata$Fertilisation.mode)
```

```
##
## External Internal
##        2     1160
```

```
table(rdata$Agricultural.importance)
```

```
##
## Control agent          No        Pest       Vector
##           425         198         708           46
```

```
table(rdata$Lab.or.field)
```

```
##
##        Field         Lab        Mix Semi-natural
##            8        1346          6           17
```

```
table(rdata$Exposure.duration)
```

```
##
##       < 24 hours      1 to 5 days      2 to 5 days          Mix  More than 5 days Natural va
##               70               19                2           17             1263
```

```
table(rdata$Life.stage.of.animal)
```

```
##
##    Adult      Egg   Embryo Juvenile   Larvae      Mix    Pupae
##      481       14        3       49       13      794       23
```

Given the imbalance in number of effect sizes in each level, taxonomic class and habitat moderators do not seem like they are going to be useful.

The moderators with the good spreads are (providing we do some recategorisation) are: 1. Sex exposed.

2. Life stage.

Sex exposed

We could lump categories so that we have cases where males are included (Both, Male), versus cases with just females (Female, Parthenogenetic), with Unsure removed. I would predict that the 'Both' category would show the biggest drop for reproduction, but there will be no difference for lifespan

```
table(new_data$Sex.exposed)
```

```
##
##    Both Female
##     573    801
```

```
meta_treat_sex <- rma.mv(es, VCV_shared_sex, mod = ~poly(c_treattemp,
    degree = 2, raw = TRUE) * Sex.exposed, random = list(~1 |
    study_code, ~1 | obs), data = new_data, method = "REML")
```

```
summary(meta_treat_sex)
```

```
##
## Multivariate Meta-Analysis Model (k = 1374; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3268.4626    6536.9251   6556.9251   6609.1215   6557.0875
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.6795  0.8243    314     no  study_code
## sigma^2.2  5.0357  2.2440   1374     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1366) = 45001.9395, p-val < .0001
##
## Test of Moderators (coefficients 2:8):
## QM(df = 7) = 645.6823, p-val < .0001
##
## Model Results:
##
##                                                      estimate      se      zval     pval     (
## intrcpt                                               -0.0703  0.1547   -0.4546   0.6494   -0
## poly(c_treattemp, degree = 3, raw = TRUE)1            -0.2889  0.0190  -15.2130   <.0001   -0
## poly(c_treattemp, degree = 3, raw = TRUE)2            -0.0031  0.0012   -2.4436   0.0145   -0
## poly(c_treattemp, degree = 3, raw = TRUE)3             0.0006  0.0001    7.5112   <.0001    0
## Sex.exposedFemale                                      0.0123  0.2020    0.0607   0.9516   -0
## poly(c_treattemp, degree = 3, raw = TRUE)1:Sex.exposedFemale  -0.0017  0.0255   -0.0673   0.9464   -0
## poly(c_treattemp, degree = 3, raw = TRUE)2:Sex.exposedFemale  -0.0003  0.0017   -0.1952   0.8453   -0
## poly(c_treattemp, degree = 3, raw = TRUE)3:Sex.exposedFemale   0.0000  0.0001    0.0936   0.9254   -0
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Life-stage

We could lump categories so that we have cases where only adults were exposed (Adult), versus cases where immature stages were exposed (Juvenile, Larvae, Pupae, Mix)- perhaps after excluding 'Egg' and 'Embryo' because these categories are a bit weird. I would predict that exposure of juveniles is worse than just exposure

of adults

```
table(ls_data$Life.stage.of.animal)
```

```
##
##    Adult Immature
##     481     879
```

```
meta_treat_ls <- rma.mv(es, VCV_shared_life, mod = ~poly(c_treattemp,
    degree = 3, raw = TRUE) * Life.stage.of.animal, random = list(~1 |
    study_code, ~1 | obs), data = ls_data, method = "REML")
```

```
summary(meta_treat_ls)
```

```
##
## Multivariate Meta-Analysis Model (k = 1360; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3240.1388   6480.2777   6500.2777   6552.3711   6500.4417
##
## Variance Components:
##
##             estim     sqrt  nlvls  fixed      factor
## sigma^2.1  0.6393   0.7995    314     no  study_code
## sigma^2.2  5.0931   2.2568   1360     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1352) = 44701.2594, p-val < .0001
##
## Test of Moderators (coefficients 2:8):
## QM(df = 7) = 644.8931, p-val < .0001
##
## Model Results:
##
##                                                               estimate      se      zval
## intrcpt                                                        -0.2700  0.1677   -1.6104
## poly(c_treattemp, degree = 3, raw = TRUE)1                     -0.3043  0.0205  -14.8172
## poly(c_treattemp, degree = 3, raw = TRUE)2                     -0.0029  0.0012   -2.3560
## poly(c_treattemp, degree = 3, raw = TRUE)3                      0.0007  0.0001    8.3334
## Life.stage.of.animalImmature                                   0.3181  0.2083    1.5270
## poly(c_treattemp, degree = 3, raw = TRUE)1:Life.stage.of.animalImmature   0.0259  0.0262    0.9866
## poly(c_treattemp, degree = 3, raw = TRUE)2:Life.stage.of.animalImmature  -0.0006  0.0017   -0.3258
## poly(c_treattemp, degree = 3, raw = TRUE)3:Life.stage.of.animalImmature  -0.0002  0.0001   -1.4164
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Sub-analysis on pest species

```
pest_data <- subset(rdata, Agricultural.importance == "Pest")




VCV_shared_pest <- impute_covariance_matrix(vi = pest_data$v,
    cluster = pest_data$shared_control, r = 0.5)

meta_pest <- rma.mv(es, VCV_shared_pest, mod = ~poly(c_treattemp,
    degree = 2, raw = TRUE), random = list(~1 | study_code, ~1 |
    obs), data = pest_data, method = "REML")
```

```
summary(meta_pest)
```

```
##
## Multivariate Meta-Analysis Model (k = 708; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -1763.1418   3526.2837   3536.2837   3559.0747   3536.3695
##
## Variance Components:
##
##             estim     sqrt   nlvls  fixed       factor
## sigma^2.1   0.2567   0.5067     156     no   study_code
## sigma^2.2   6.9862   2.6431     708     no          obs
##
## Test for Residual Heterogeneity:
## QE(df = 705) = 26704.6508, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 290.7654, p-val < .0001
##
## Model Results:
##
##                                       estimate       se      zval     pval     ci.lb     ci.ub
## intrcpt                                -0.0753   0.1442   -0.5225   0.6013   -0.3580    0.2073
## poly(c_treattemp, degree = 2, raw = TRUE)1   -0.2183   0.0129  -16.9012   <.0001   -0.2436   -0.1929   **
## poly(c_treattemp, degree = 2, raw = TRUE)2   -0.0000   0.0014   -0.0026   0.9979   -0.0028    0.0027
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Publication Bias.

We fit meta-regression model with year as a moderator to see whether there is an publication bias. In other words, is effect size strongly correlated with publication year. We found now evidence to suggest that year has an effect on the report effect sizes. A summary of the model is given below.

```
meta_year <- rma.mv(es, VCV_shared, mod = ~Publication.year,
    random = list(~1 | study_code, ~1 | obs), data = rdata, test = "t",
    dfs = "contain", method = "REML")
```

```
summary(meta_year)
```

```
##
## Multivariate Meta-Analysis Model (k = 1377; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3548.1669   7096.3338   7104.3338   7125.2386   7104.3630
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.6965  0.8346    315     no  study_code
## sigma^2.2  7.9493  2.8195   1377     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1375) = 69422.4286, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 0.6979, p-val = 0.4035
##
## Model Results:
##
##                    estimate       se     zval    pval     ci.lb    ci.ub
## intrcpt             16.5861  20.0795   0.8260  0.4088  -22.7690  55.9413
## Publication.year    -0.0083   0.0100  -0.8354  0.4035   -0.0279   0.0112
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
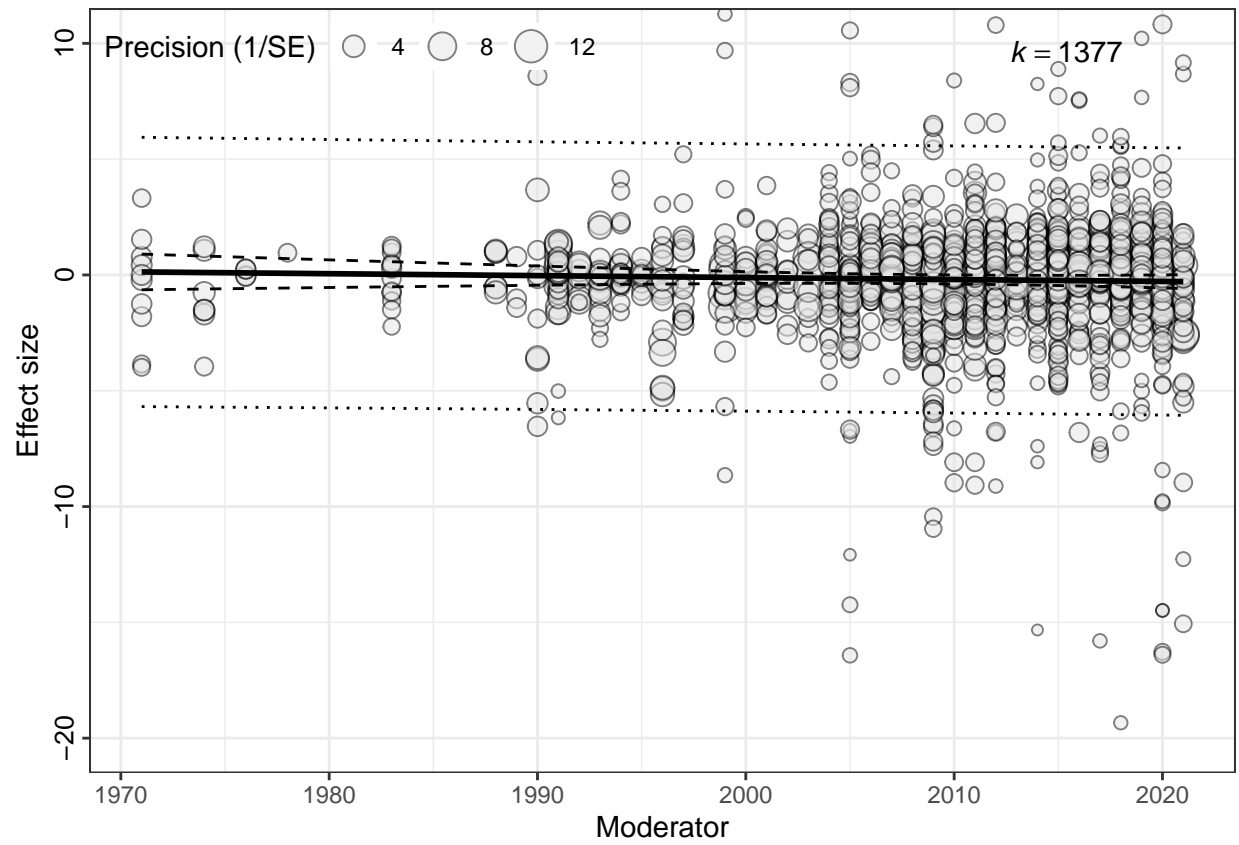
Sensitivty Analysis

Here, we perform a sensitivity analysis by removing the smallest and largest 2.5% of effect sizes.

The 2.5th and 97.5th percentiles are given below.

```
print(minq)
```

```
##      2.5%
## -6.911142
```

```
print(maxq)
```

```
##     97.5%
## 6.443133
```

The sumamry of the effect sizes is now

```
summary(sdata$es)
```

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -6.82667 -1.23094 -0.06942 -0.18585  0.98969  6.37946
```

We run the cubic treatment temperature and the binned temperature models again with our new subsetted data i.e. only with data that is between the 2.5Th and 97.5Th percentile.

Treatment temperature as a cubic effect (sesnsitivity analysis)

We re-create the variance-covariance matrix with our new subsetted data, which we name sdata. Then we run the meta analysis model again with treatment temperature as quadratic fixed effect.

```
# recreate vcv_shared matrix
VCV_shared_sa <- impute_covariance_matrix(vi = sdata$v, cluster = sdata$shared_control,
    r = 0.5)
```
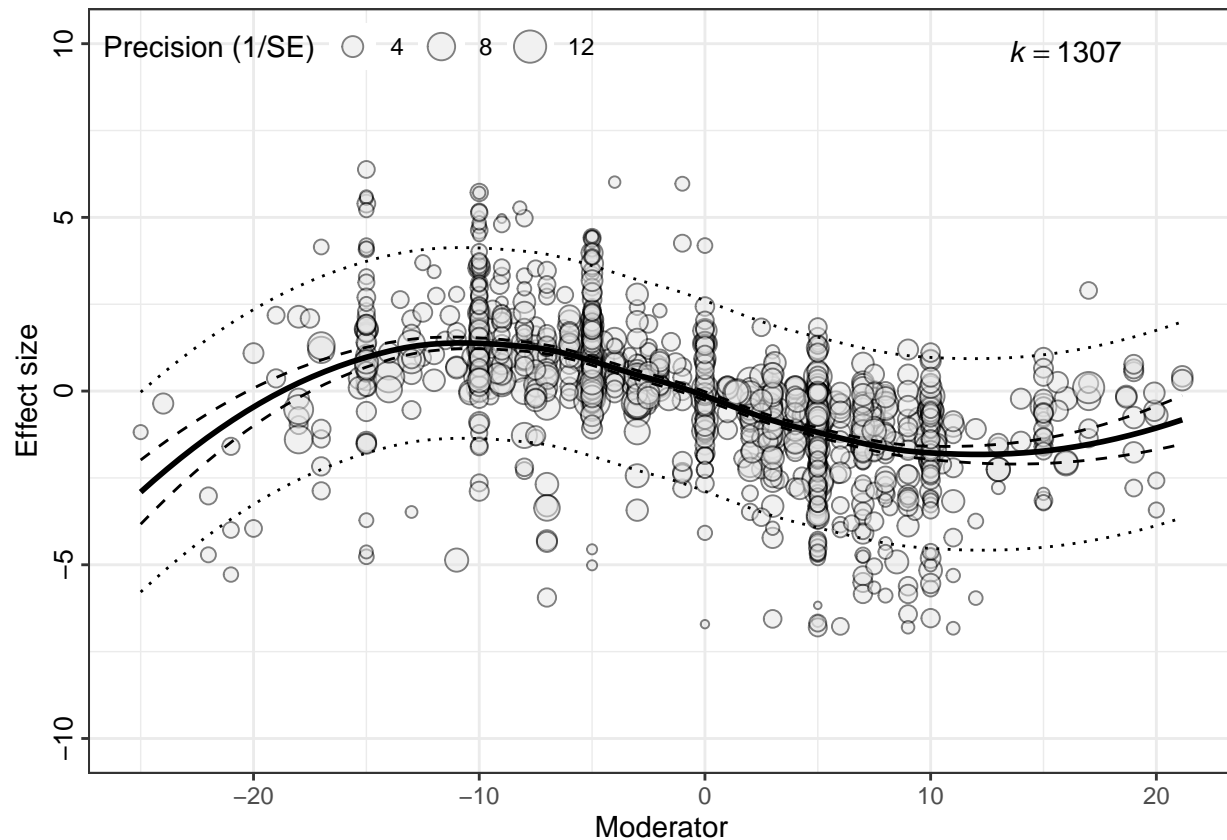
```
meta_sa_treat3 <- rma.mv(es, VCV_shared_sa, mod = ~poly(c_treattemp,
    degree = 3, raw = TRUE), random = list(~1 | study_code, ~1 |
    obs), test = "t", dfs = "contain", data = sdata, method = "REML")
```

```
summary(meta_sa_treat3)
```

```
##
## Multivariate Meta-Analysis Model (k = 1307; method: REML)
##
##     logLik    Deviance        AIC         BIC        AICc
## -2298.2358   4596.4717   4608.4717   4639.5062   4608.5365
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed     factor
## sigma^2.1  0.5199  0.7210    312     no  study_code
## sigma^2.2  1.4353  1.1980   1307     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1303) = 28210.1296, p-val < .0001
##
## Test of Moderators (coefficients 2:4):
## QM(df = 3) = 1101.0657, p-val < .0001
##
```

```
## Model Results:
##
##                                       estimate      se      zval    pval    ci.lb     ci.ub
## intrcpt                                -0.1581  0.0669   -2.3628  0.0181  -0.2892  -0.0269
## poly(c_treattemp, degree = 3, raw = TRUE)1   -0.2228  0.0073  -30.3373  <.0001  -0.2372  -0.2084  ***
## poly(c_treattemp, degree = 3, raw = TRUE)2   -0.0010  0.0005   -1.7879  0.0738  -0.0021   0.0001
## poly(c_treattemp, degree = 3, raw = TRUE)3    0.0005  0.0000   13.1738  <.0001   0.0004   0.0006  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Below is a bubble plot of the fitted cubic model to the subsetted data.

Modelling response with binned treatment temperatures

Lastly, we rerun the model where our effect sizes are categorised into one seven bins depending on the treatment temperature. The bins and the number of effect sizes in each bin of the new subsetted data are given below

```
table(sdata$bin.temp)

##
##    <15    >40 15-20 20-25 25-30 30-35 35-40
##    102     48   233   298   185   342    99

meta_sa_bintemp <- rma.mv(es, VCV_shared_sa, mod = ~bin.temp -
    1, random = list(~1 | study_code, ~1 | obs), test = "t",
    dfs = "contain", data = sdata, method = "REML")

summary(meta_sa_bintemp)

##
## Multivariate Meta-Analysis Model (k = 1307; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -2320.3142   4640.6285   4658.6285   4705.1596   4658.7680
##
## Variance Components:
##
```

```
##              estim    sqrt  nlvls  fixed       factor
## sigma^2.1  0.5486  0.7407    312     no  study_code
## sigma^2.2  1.5040  1.2264   1307     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1300) = 29271.2184, p-val < .0001
##
## Test of Moderators (coefficients 1:7):
## QM(df = 7) = 1013.9200, p-val < .0001
##
## Model Results:
##
##                estimate      se      zval    pval    ci.lb    ci.ub
## bin.temp<15      0.9801  0.1543    6.3535  <.0001   0.6777   1.2824  ***
## bin.temp>40     -1.5756  0.2513   -6.2692  <.0001  -2.0682  -1.0830  ***
## bin.temp15-20    1.2745  0.0993   12.8376  <.0001   1.0799   1.4691  ***
## bin.temp20-25    0.8331  0.0897    9.2831  <.0001   0.6572   1.0090  ***
## bin.temp25-30   -0.6769  0.1098   -6.1633  <.0001  -0.8922  -0.4617  ***
## bin.temp30-35   -1.4249  0.0853  -16.7068  <.0001  -1.5921  -1.2578  ***
## bin.temp35-40   -1.9770  0.1470  -13.4521  <.0001  -2.2651  -1.6890  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```
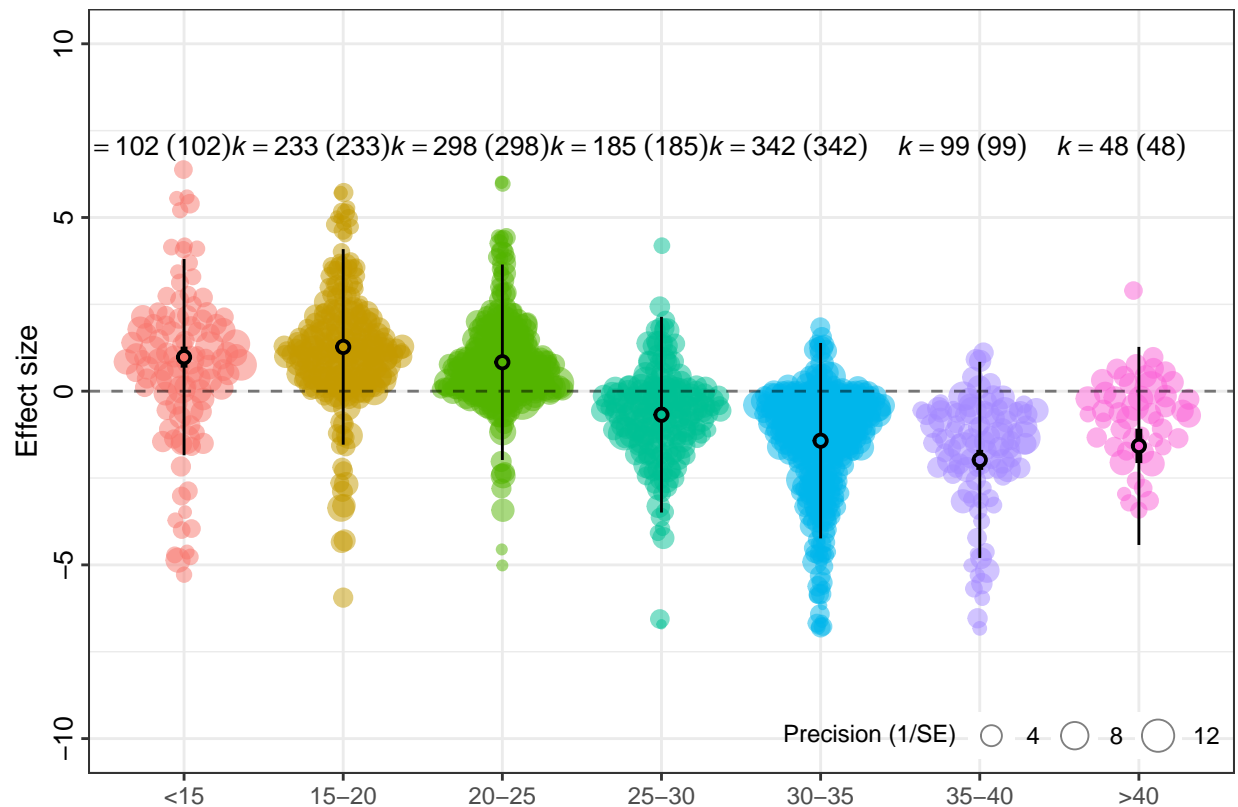
Geary;'s Test

Here we use Geary's test to determine which effect sizes to remove. According to Hedges et al 1999, effect sizes are deemed valid and accurate when the standardized mean of either the control or the treatment group is ≥ 3. In other words, if

$$\frac{\bar{X}}{SD}\sqrt{N} \geq 3$$

.

As a sensitivity analysis we will determine which of our effect sizes do not meet this threshold i.e. are < 3. We then rerun our best fitting models with these effect sizes removed.

```r
# Read in data with calculated standardised mean
gearydata <- read.csv("../Data/Gearys_test_data.csv")

# view the first fews rows of selected columns
gearydata %>%
    dplyr::select(Experiment.code, Trait.category, reftemp, treattemp,
        es, v, gtest) %>%
    head(., 30)

# How many effect sozes have a standardised mean < 3 ?
outliers_lon <- subset(gearydata, Trait.category == "Longevity" &
    gtest < 3)  #82
outliers_sur <- subset(gearydata, Trait.category == "Survival" &
    gtest < 3)  # 3
outliers_rep <- subset(gearydata, Trait.category == "Reproduction" &
    gtest < 3)  # 171
```

There are 256 effect sizes in total that have a gtest < 3. 82 of these are longevity effect sizes, 3 are survival and 171 are reproduction.

We remove these effect sizes and rerun the best fitting model from earlier which is the model meta_trait_treat2, where temperature is modelled as a quadratic effect on reproduction.

Treatment temperature as a quadratic effect with Geary's outliers removed.

```r
# Remove outliers
s.gearydata <- gearydata %>%
    subset(gtest > 3)

# Resubset reproduction dat
gdata_lon_warm <- subset(s.gearydata, Trait.category == "Longevity" &
    warm.cool == "Warm")
gdata_lon_cool <- subset(s.gearydata, Trait.category == "Longevity" &
    warm.cool == "Cool")

all_gdata <- rbind(gdata_lon_warm, gdata_lon_cool)

all_gdata <- all_gdata %>%
    mutate(c_treattemp = treattemp - 25)



### Create random factors into data frame
all_gdata$obs <- factor(c(1:nrow(all_gdata)))  # Unique observation code
all_gdata$study_code <- factor(all_gdata$Paper.code)  # Model requires column names study_code
```

```r
all_gdata$Species.phylo <- factor(all_gdata$Species.latin)  # Species names for phylo matrix
all_gdata$species <- factor(all_gdata$Species.latin)  # Another species column for random factor

precision <- sqrt(1/all_gdata$v)  # inverse standard error
all_gdata[, "precision"] <- precision

# recalculate v matrix
all_gdata$shared_control <- factor(all_gdata$Effect.size.code)
VCV_shared_ga <- impute_covariance_matrix(vi = all_gdata$v, cluster = all_gdata$shared_control,
    r = 0.5)
```

```r
meta_ga_treat3 <- rma.mv(es, VCV_shared_ga, mod = ~poly(c_treattemp,
    degree = 3, raw = TRUE), random = list(~1 | study_code, ~1 |
    obs), test = "t", dfs = "contain", data = all_gdata, method = "REML")
```

```r
summary(meta_ga_treat3)
```

```
##
## Multivariate Meta-Analysis Model (k = 1310; method: REML)
##
##      logLik    Deviance         AIC         BIC        AICc
## -3195.6635   6391.3269   6403.3269   6434.3753   6403.3916
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.6337  0.7960    311     no  study_code
## sigma^2.2  5.1381  2.2667   1310     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1306) = 44067.8973, p-val < .0001
##
## Test of Moderators (coefficients 2:4):
## F(df1 = 3, df2 = 1306) = 215.6771, p-val < .0001
##
## Model Results:
##
##                                              estimate      se      tval    df    pval    ci.lb    ci.u
## intrcpt                                       -0.0537  0.1011   -0.5316   307  0.5954  -0.2526   0.145
## poly(c_treattemp, degree = 3, raw = TRUE)1    -0.3003  0.0130  -23.1438  1306  <.0001  -0.3258  -0.274
## poly(c_treattemp, degree = 3, raw = TRUE)2    -0.0030  0.0009   -3.4836  1306  0.0005  -0.0047  -0.00
## poly(c_treattemp, degree = 3, raw = TRUE)3     0.0007  0.0001   10.5347  1306  <.0001   0.0005   0.000
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Below is a bubble plot of the fitted quadratic model.