Meta-Analysis Reproduction Summary Excluding HUM251

Fay Frost

1. Summary

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

2. Setup

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

```
### Read in effect size data
effectdata <- read.csv("Data/Survival project all pairwise.es.csv")
repdata_warm <- subset(effectdata, Trait.category == "Reproduction" &
    warm.cool == "Warm")
repdata_cool <- subset(effectdata, Trait.category == "Reproduction" &
    warm.cool == "Cool")

allrep <- rbind(repdata_warm, repdata_cool)

### select data for analysis
rdata <- allrep
```

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.

```
### 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.

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

```
## [1] 306
```

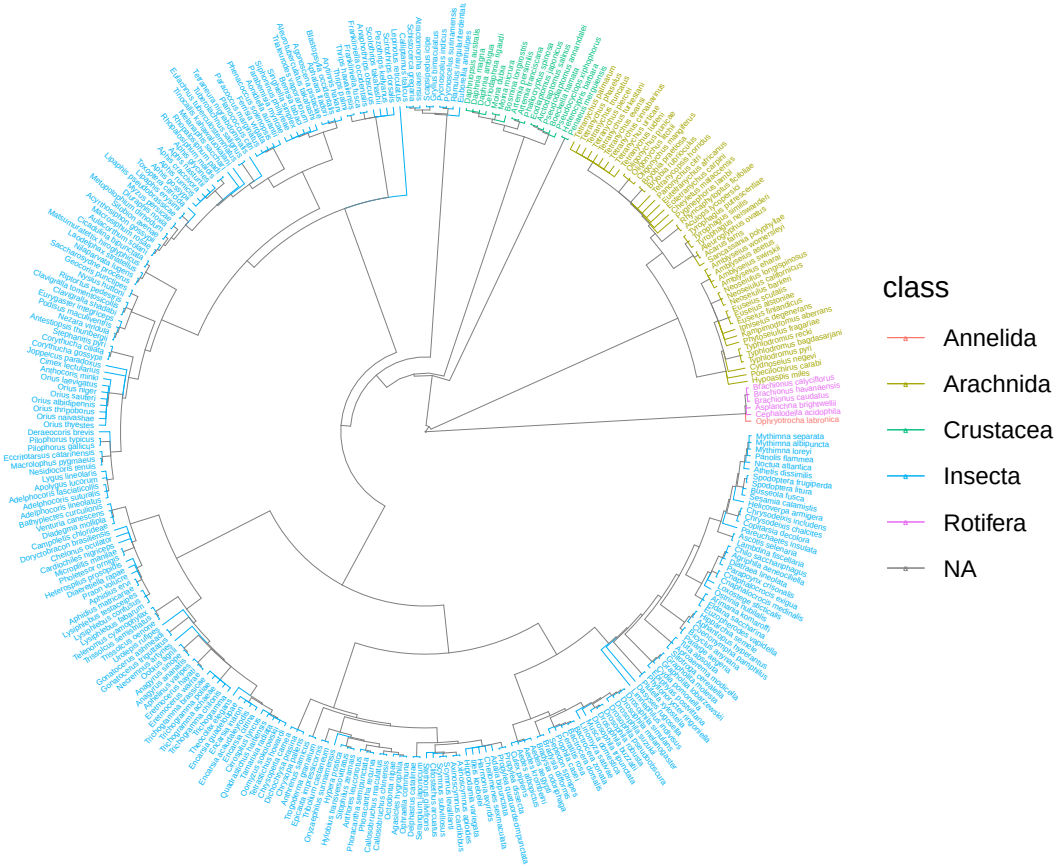
```
nlevels(rdata$study_code)  # Check number of studies
```

```
## [1] 339
```

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_reproduction_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), test = "t", dfs = "contain",
    R = list(Species.phylo = phylo_matrix), data = rdata, method = "REML")
```

```
summary(meta2)
```

```
##
## Multivariate Meta-Analysis Model (k = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -4048.7191   8097.4382   8107.4382   8133.6090   8107.4816
##
## Variance Components:
##
##              estim    sqrt  nlvls  fixed         factor    R
## sigma^2.1   0.0000  0.0000    306     no  Species.phylo  yes
## sigma^2.2   0.0000  0.0003    306     no        species   no
## sigma^2.3  10.4420  3.2314    339     no     study_code   no
## sigma^2.4   5.2505  2.2914   1387     no            obs   no
##
## Test for Heterogeneity:
## Q(df = 1386) = 37155.0413, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -1.5543  0.1907  -8.1485  <.0001  -1.9282  -1.1805  ***
##
## ---
## 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
##     9.946178e+01     7.043695e-10     7.683821e-07     6.618307e+01     3.327870e+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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3770.1104   7540.2208   7550.2208   7576.3917   7550.2643
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed         factor    R
## sigma^2.1  0.0000  0.0000    306     no  Species.phylo  yes
## sigma^2.2  0.0000  0.0002    306     no        species   no
## sigma^2.3  5.1474  2.2688    339     no     study_code   no
## sigma^2.4  4.4602  2.1119   1387     no            obs   no
##
## Test for Heterogeneity:
## Q(df = 1386) = 38656.3912, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -1.3401  0.1414  -9.4771  <.0001  -1.6172  -1.0629  ***
##
## ---
## 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
##     9.912389e+01     1.238319e-09     6.017065e-07     5.310681e+01     4.601707e+01
```

Model without phylogeny

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

```
## without phylogeny but with shared control
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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3770.1104   7540.2208   7548.2208   7569.1575   7548.2497
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  0.0000  0.0002    306     no     species
## sigma^2.2  5.1474  2.2688    339     no  study_code
## sigma^2.3  4.4602  2.1119   1387     no         obs
##
## Test for Heterogeneity:
## Q(df = 1386) = 38656.3912, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -1.3401  0.1414  -9.4771  <.0001  -1.6172  -1.0629  ***
##
## ---
## 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
##  9.912389e+01  3.936585e-07  5.310681e+01  4.601707e+01
```

Model without phylogeny or species

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

```
summary(meta4)
```

```
##
## Multivariate Meta-Analysis Model (k = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3770.1104   7540.2208   7546.2208   7561.9233   7546.2382
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.1474  2.2688    339     no  study_code
## sigma^2.2  4.4602  2.1119   1387     no         obs
##
## Test for Heterogeneity:
## Q(df = 1386) = 38656.3912, p-val < .0001
##
## Model Results:
##
## estimate      se     zval    pval    ci.lb    ci.ub
##  -1.3401  0.1414  -9.4771  <.0001  -1.6172  -1.0629  ***
##
## ---
## 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
##      99.12389      53.10681      46.01708
```

Model without phylogeny, species or study_code

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

```
summary(meta7)
```

```
##
## Multivariate Meta-Analysis Model (k = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3926.3572   7852.7144   7856.7144   7867.1827   7856.7230
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed  factor
## sigma^2    7.0479  2.6548   1387     no     obs
##
## Test for Heterogeneity:
## Q(df = 1386) = 38656.3912, p-val < .0001
##
## Model Results:
##
## estimate      se      zval    pval    ci.lb    ci.ub
##  -1.2769  0.0743  -17.1774  <.0001  -1.4226  -1.1312  ***
##
## ---
## 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
## 98.80947 98.80947
```

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 meta5 . We continue our analysis using meta4 as our base model.

4. Meta-regressions

Starting with the best fitting random-effect model from Section 3, "meta8" 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
- **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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3766.7483   7533.4967   7541.4967   7562.4305   7541.5257
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.0944  2.2571    339     no  study_code
## sigma^2.2  4.4713  2.1145   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1385) = 38565.0269, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 1.8362, p-val = 0.1754
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb   ci.ub
## intrcpt    0.2198  1.1597   0.1896  0.8497  -2.0531  2.4927
## reftemp   -0.0631  0.0466  -1.3551  0.1754  -0.1545  0.0282
##
## ---
## 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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3759.9690   7519.9380   7527.9380   7548.8718   7527.9670
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.1557  2.2706    339     no  study_code
## sigma^2.2  4.3805  2.0930   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1385) = 38615.4347, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 16.5055, p-val < .0001
##
## Model Results:
##
##              estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt       -1.3584  0.1413  -9.6123  <.0001  -1.6354  -1.0814  ***
## c_treattemp   -0.0354  0.0087  -4.0627  <.0001  -0.0524  -0.0183  ***
##
## ---
## 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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3760.9334   7521.8668   7529.8668   7550.8006   7529.8958
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.1402  2.2672    339     no  study_code
## sigma^2.2  4.3888  2.0950   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1385) = 38400.3073, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 14.8227, p-val = 0.0001
##
## Model Results:
##
##                estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt         -1.0819  0.1560  -6.9363  <.0001  -1.3876  -0.7762  ***
## warm.coolWarm   -0.5061  0.1315  -3.8500  0.0001  -0.7638  -0.2485  ***
##
## ---
## 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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3760.9334   7521.8668   7529.8668   7550.8006   7529.8958
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.1402  2.2672    339     no  study_code
## sigma^2.2  4.3888  2.0950   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1385) = 38400.3073, p-val < .0001
##
## Test of Moderators (coefficients 1:2):
## QM(df = 2) = 104.7301, p-val < .0001
##
## Model Results:
##
##                estimate      se      zval    pval    ci.lb    ci.ub
## warm.coolCool   -1.0819  0.1560   -6.9363  <.0001  -1.3876  -0.7762  ***
## warm.coolWarm   -1.5880  0.1553  -10.2230  <.0001  -1.8924  -1.2835  ***
##
## ---
## 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), test = "t", dfs = "contain", data = rdata,
    method = "REML")
```

```
summary(meta_trait_diff)
```

```
##
## Multivariate Meta-Analysis Model (k = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3760.7161   7521.4323   7529.4323   7550.3661   7529.4612
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.2122  2.2830    339     no  study_code
## sigma^2.2  4.3759  2.0919   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1385) = 38574.5945, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 15.0667, p-val = 0.0001
##
## Model Results:
##
##          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt   -1.3478  0.1418  -9.5026  <.0001  -1.6258  -1.0698  ***
## diff      -0.0343  0.0088  -3.8816  0.0001  -0.0517  -0.0170  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Treatment temperature as a quadratic effect

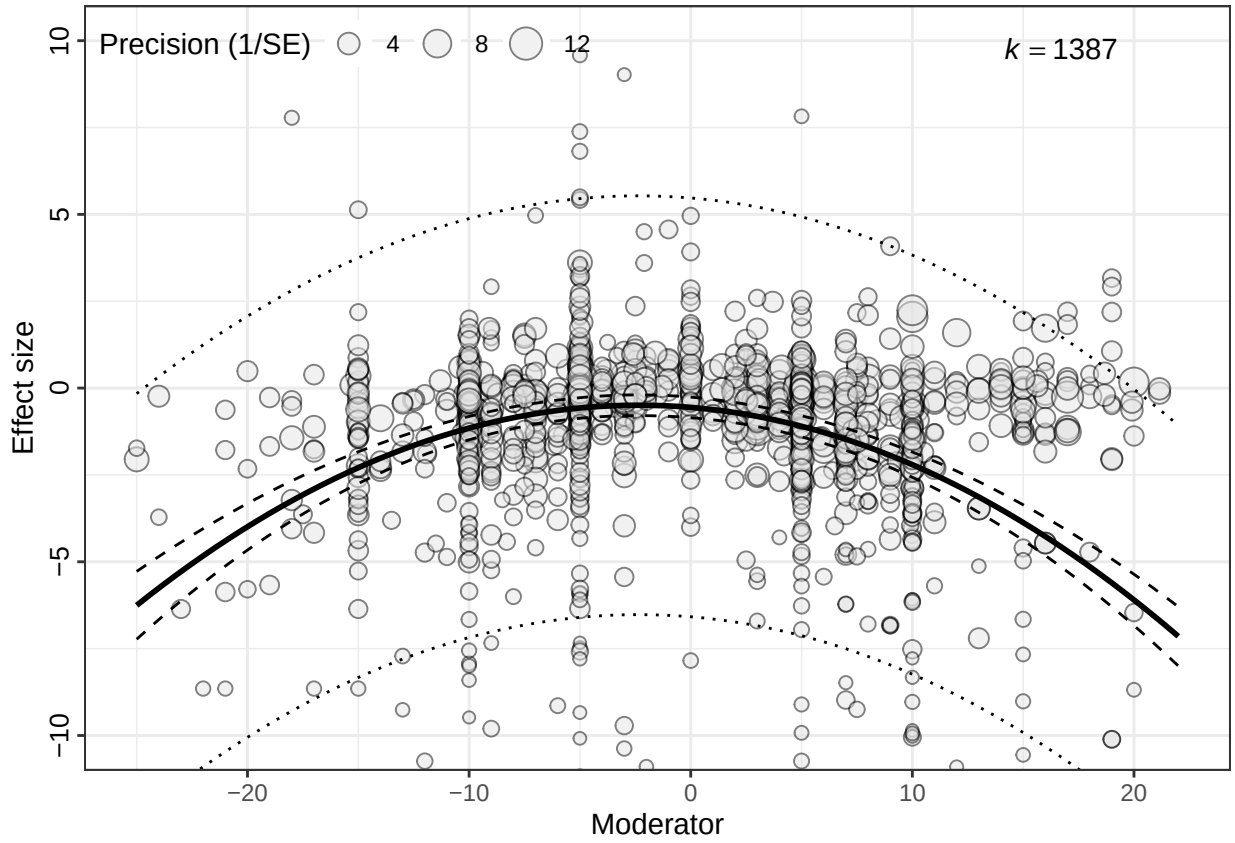
We expect that treatment temperature will have a quadratic effect on reproduction. In other words, we expect that reproduction will decrease either side of an optimum temperature for each species. We also assume that this optimum temperature will be close to the control temperature. That said, we investigate **treattemp** as a quadratic fixed 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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3682.3878   7364.7755   7374.7755   7400.9392   7374.8191
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.8064  2.4096    339     no  study_code
## sigma^2.2  3.6290  1.9050   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1384) = 36417.9019, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 185.2754, p-val < .0001
##
## Model Results:
##
##                                          estimate      se      zval    pval    ci.lb    ci.ub
## intrcpt                                   -0.5552  0.1579   -3.5172  0.0004  -0.8646  -0.2458  **
## poly(c_treattemp, degree = 2, raw = TRUE)1   -0.0526  0.0082   -6.4356  <.0001  -0.0687  -0.0366  **
## poly(c_treattemp, degree = 2, raw = TRUE)2   -0.0112  0.0009  -12.8809  <.0001  -0.0129  -0.0095  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Below is a bubble plot of the fitted quadratic 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

```
## [1] "<15"   ">40"   "15-20" "20-25" "25-30" "30-35" "35-40"
```

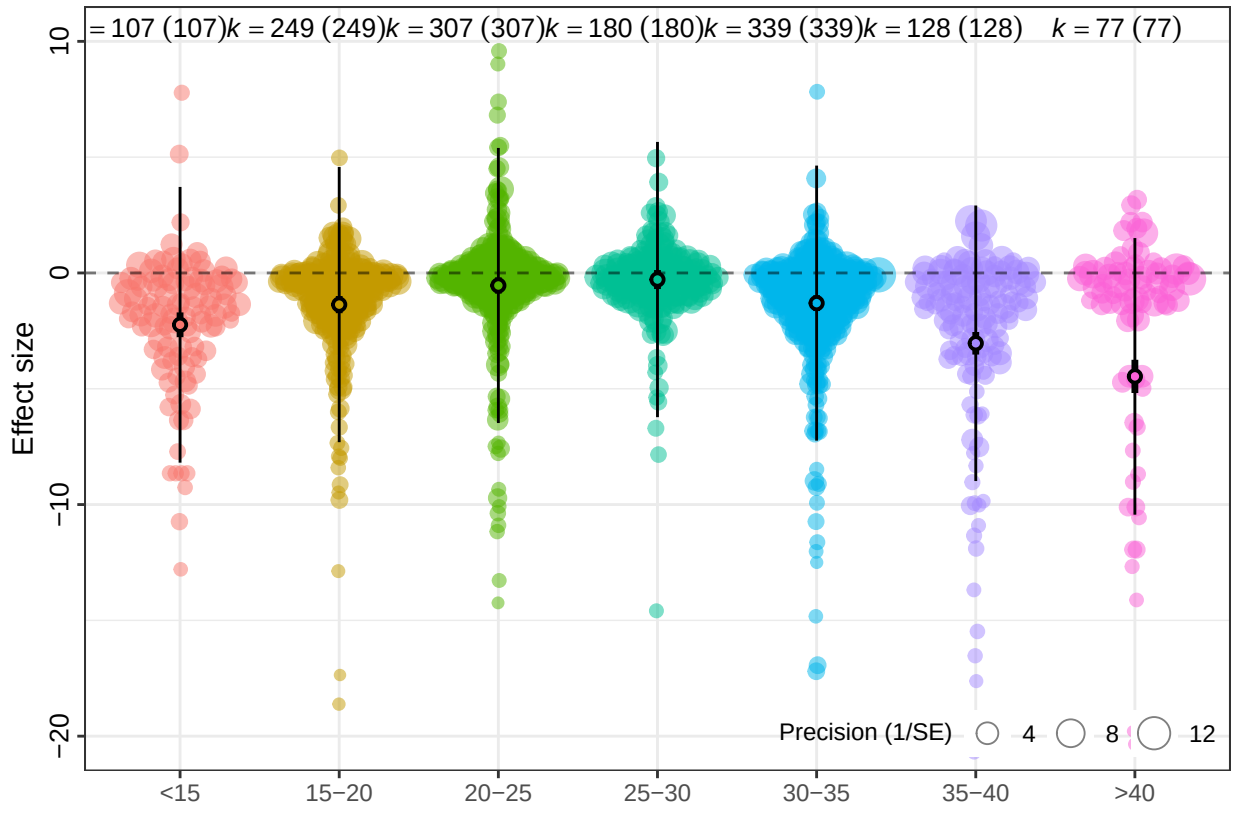
```
table(rdata$bin.temp)
```

```
##
##   <15   >40 15-20 20-25 25-30 30-35 35-40
##   107    77   249   307   180   339   128
```

```
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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3663.6641   7327.3282   7345.3282   7392.3967   7345.4595
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.5364  2.3530    339     no  study_code
## sigma^2.2  3.6126  1.9007   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1380) = 36020.9128, p-val < .0001
##
## Test of Moderators (coefficients 1:7):
## QM(df = 7) = 294.0378, p-val < .0001
##
## Model Results:
##
##                estimate      se      zval    pval    ci.lb    ci.ub
## bin.temp<15     -2.2382  0.2727   -8.2077  <.0001  -2.7727  -1.7037  ***
## bin.temp>40     -4.4663  0.3662  -12.1967  <.0001  -5.1840  -3.7486  ***
## bin.temp15-20   -1.3672  0.1891   -7.2292  <.0001  -1.7378  -0.9965  ***
## bin.temp20-25   -0.5414  0.1783   -3.0357  0.0024  -0.8909  -0.1918   **
## bin.temp25-30   -0.2847  0.2094   -1.3599  0.1739  -0.6951   0.1256
## bin.temp30-35   -1.3030  0.1744   -7.4720  <.0001  -1.6447  -0.9612  ***
## bin.temp35-40   -3.0364  0.2484  -12.2229  <.0001  -3.5233  -2.5495  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


= 107 (107)k = 249 (249)k = 307 (307)k = 180 (180)k = 339 (339)k = 128 (128) k = 77 (77)
Effect size
10
0
−10
−20
Precision (1/SE)
4
8
12
<15
15−20
20−25
25−30
30−35
35−40
>40

Other fixed effects

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

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

```
##
##  Annelida Arachnida Crustacea   Insecta  Rotifera
##         2       212        40      1107        26
```

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

```
##
##     Aquatic        Both Terrestrial
##          71          41        1275
```

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

```
##
##            Both          Female            Male Parthenogenetic          Unsure
##             871             310               6             197               3
```

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

```
##
## External Internal
##        2     1174
```

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

```
##
## Control agent            No          Pest        Vector
##           422           212           725            28
```

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

```
##
##        Field          Lab          Mix Semi-natural
##            8         1356            6           17
```

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

```
##
##        < 24 hours       1 to 5 days               Mix  More than 5 days Natural variation
##                86                45                22              1226                 8
```

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

```
##
##    Adult      Egg   Embryo Juvenile   Larvae      Mix    Pupae
##      504       14        3       52       13      773       28
```

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
##    877    507
```

```
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 = 1384; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3655.4228   7310.8456   7326.8456   7368.6727   7326.9508
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.8495  2.4186    338     no  study_code
## sigma^2.2  3.5461  1.8831   1384     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1378) = 36058.6623, p-val < .0001
##
## Test of Moderators (coefficients 2:6):
## QM(df = 5) = 218.3540, p-val < .0001
##
## Model Results:
##
##                                                          estimate      se     zval    pval    c
## intrcpt                                                   -0.6954  0.1951  -3.5645  0.0004  -1.
## poly(c_treattemp, degree = 2, raw = TRUE)1                -0.0242  0.0100  -2.4196  0.0155  -0.
## poly(c_treattemp, degree = 2, raw = TRUE)2                -0.0099  0.0010  -9.5679  <.0001  -0.
## Sex.exposedFemale                                          0.3501  0.3154   1.1100  0.2670  -0.
## poly(c_treattemp, degree = 2, raw = TRUE)1:Sex.exposedFemale  -0.0785  0.0159  -4.9443  <.0001  -0.
## poly(c_treattemp, degree = 2, raw = TRUE)2:Sex.exposedFemale  -0.0040  0.0017  -2.3941  0.0167  -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

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

```
##
##    Adult Immature
##      504      866
```

```r
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")
```

```r
summary(meta_treat_ls)
```

```
##
## Multivariate Meta-Analysis Model (k = 1370; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3562.7153   7125.4306   7145.4306   7197.5977   7145.5934
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.7222  2.3921    338     no  study_code
## sigma^2.2  3.2170  1.7936   1370     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1362) = 34641.6909, p-val < .0001
##
## Test of Moderators (coefficients 2:8):
## QM(df = 7) = 222.4540, p-val < .0001
##
## Model Results:
##
##                                                                    estimate      se     zval
## intrcpt                                                             -0.7501  0.2504  -2.9953
## poly(c_treattemp, degree = 3, raw = TRUE)1                          -0.0005  0.0181  -0.0250
## poly(c_treattemp, degree = 3, raw = TRUE)2                          -0.0071  0.0012  -5.7647
## poly(c_treattemp, degree = 3, raw = TRUE)3                          -0.0001  0.0001  -1.3180
## Life.stage.of.animalImmature                                         0.3226  0.3084   1.0459
## poly(c_treattemp, degree = 3, raw = TRUE)1:Life.stage.of.animalImmature  -0.0450  0.0229  -1.9700
## poly(c_treattemp, degree = 3, raw = TRUE)2:Life.stage.of.animalImmature  -0.0082  0.0017  -4.9611
## poly(c_treattemp, degree = 3, raw = TRUE)3:Life.stage.of.animalImmature  -0.0001  0.0001  -1.1485
##
## ---
## 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 = 725; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -1636.4749   3272.9498   3282.9498   3305.8599   3283.0336
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  2.0370  1.4273    166     no  study_code
## sigma^2.2  3.0406  1.7437    725     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 722) = 19002.3449, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 53.5416, p-val < .0001
##
## Model Results:
##
##                                          estimate      se     zval    pval    ci.lb    ci.ub
## intrcpt                                   -0.5542  0.1529  -3.6246  0.0003  -0.8538  -0.2545  ***
## poly(c_treattemp, degree = 2, raw = TRUE)1   -0.0260  0.0097  -2.6797  0.0074  -0.0450  -0.0070   **
## poly(c_treattemp, degree = 2, raw = TRUE)2   -0.0077  0.0011  -7.1422  <.0001  -0.0098  -0.0056  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

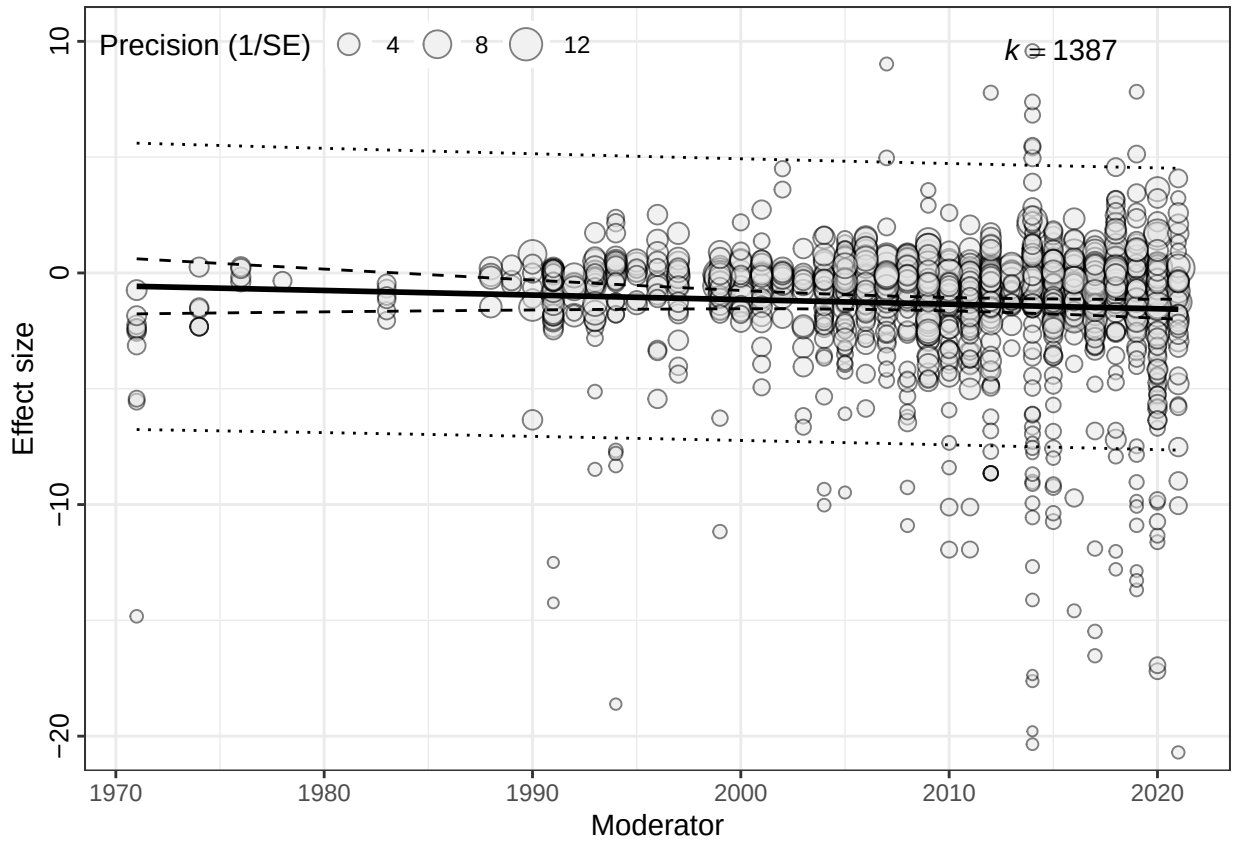
Publication Bias.

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 = 1387; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3766.6712   7533.3425   7541.3425   7562.2763   7541.3715
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  5.1258  2.2640    339     no  study_code
## sigma^2.2  4.4583  2.1115   1387     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1385) = 38642.9810, p-val < .0001
##
## Test of Moderators (coefficient 2):
## QM(df = 1) = 1.6764, p-val = 0.1954
##
## Model Results:
##
##                   estimate       se     zval    pval     ci.lb    ci.ub
## intrcpt            38.3496  30.6547   1.2510  0.2109  -21.7325  98.4317
## Publication.year   -0.0198   0.0153  -1.2947  0.1954   -0.0496   0.0101
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```


Precision (1/SE)
4
8
12
k = 1387
10
0
−10
−20
Effect size
1970
1980
1990
2000
2010
2020
Moderator

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%
## -14.36496
```

```
print(maxq)
```

```
##    97.5%
## 2.362038
```

The summary of the effective sizes is now

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

```
##     Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
## -14.2452  -1.8325  -0.6242  -1.3739   0.0000   2.3584
```

We run the quadratic 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 quadratic effect (sesnsitivity analysis)

We re-create the variance-covariance matrix with our new substted 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)
```

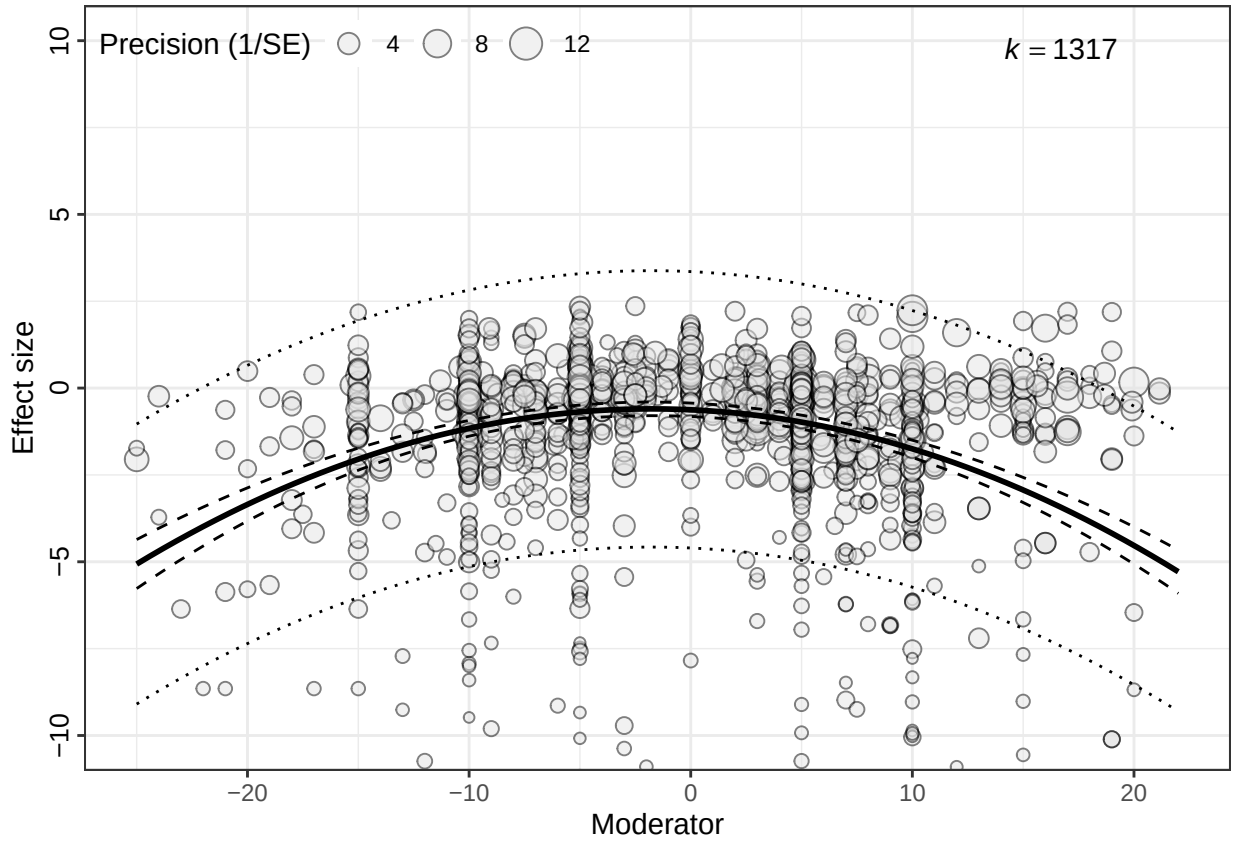
```
meta_sa_treat2 <- rma.mv(es, VCV_shared_sa, mod = ~poly(c_treattemp,
    degree = 2, raw = TRUE), random = list(~1 | study_code, ~1 |
    obs), test = "t", dfs = "contain", data = sdata, method = "REML")
```

```
summary(meta_sa_treat2)
```

```
##
## Multivariate Meta-Analysis Model (k = 1317; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -2688.7673   5377.5346   5387.5346   5413.4388   5387.5805
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  2.1891  1.4796    333     no  study_code
## sigma^2.2  1.9202  1.3857   1317     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1314) = 28346.3766, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## QM(df = 2) = 176.1132, p-val < .0001
##
```

```
## Model Results:
##
##                                            estimate      se      zval    pval    ci.lb    ci.ub
## intrcpt                                     -0.6247  0.1043   -5.9910  <.0001  -0.8290  -0.4203  **
## poly(c_treattemp, degree = 2, raw = TRUE)1  -0.0296  0.0061   -4.8815  <.0001  -0.0414  -0.0177  **
## poly(c_treattemp, degree = 2, raw = TRUE)2  -0.0083  0.0006  -13.0113  <.0001  -0.0095  -0.0070  **
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Below is a bubble plot of the fitted quadratic 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

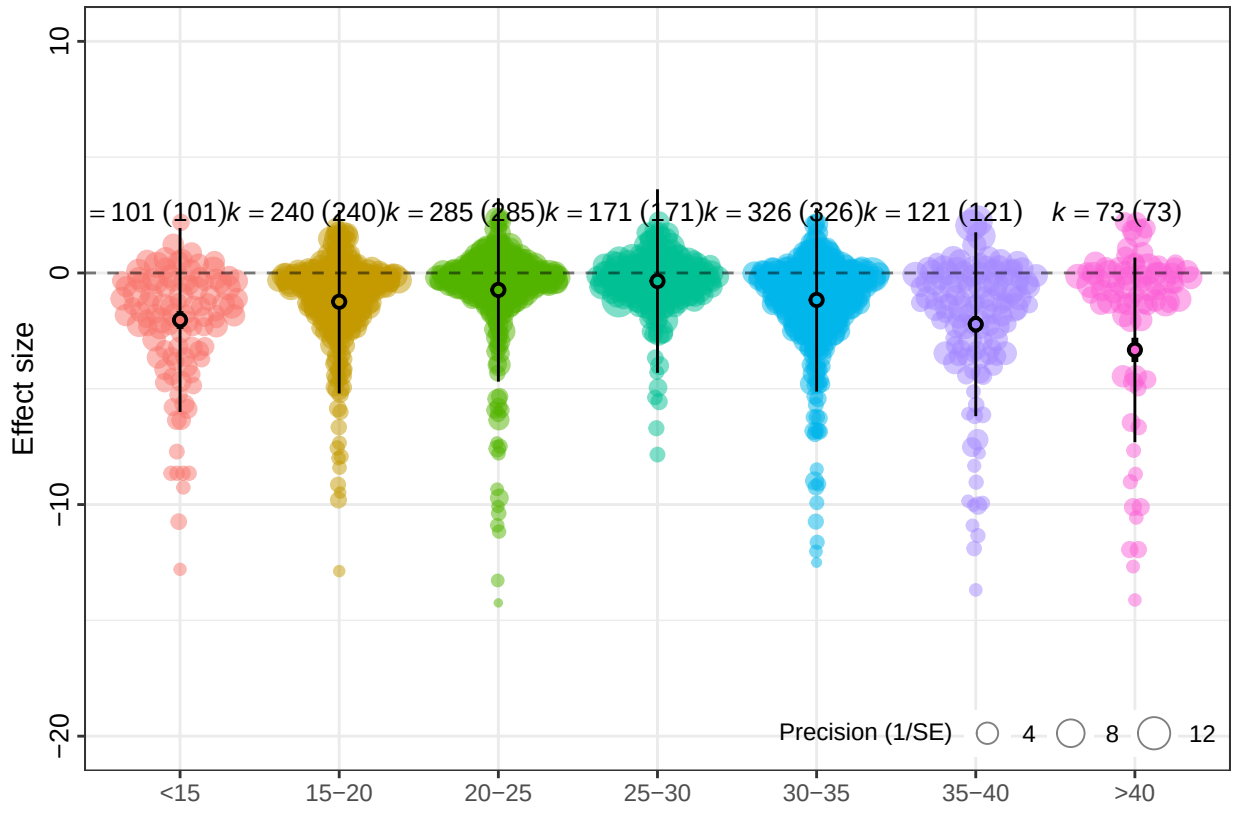
```
table(sdata$bin.temp)
```

```
##
##   <15   >40 15-20 20-25 25-30 30-35 35-40
##   101    73   240   285   171   326   121
```

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

```
summary(meta_sa_bintemp)
```

```
##
## Multivariate Meta-Analysis Model (k = 1317; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -2682.1397   5364.2794   5382.2794   5428.8795   5382.4179
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  2.1228  1.4570    333     no  study_code
## sigma^2.2  1.9424  1.3937   1317     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1310) = 28196.0335, p-val < .0001
##
## Test of Moderators (coefficients 1:7):
## QM(df = 7) = 339.9113, p-val < .0001
##
## Model Results:
##
##                estimate      se      zval    pval    ci.lb    ci.ub
## bin.temp<15     -2.0314  0.1958  -10.3731  <.0001  -2.4153  -1.6476  ***
## bin.temp>40     -3.3215  0.2678  -12.4015  <.0001  -3.8464  -2.7965  ***
## bin.temp15-20   -1.2425  0.1308   -9.4974  <.0001  -1.4989  -0.9861  ***
## bin.temp20-25   -0.7339  0.1241   -5.9157  <.0001  -0.9770  -0.4907  ***
## bin.temp25-30   -0.3498  0.1483   -2.3581  0.0184  -0.6406  -0.0591    *
## bin.temp30-35   -1.1655  0.1194   -9.7619  <.0001  -1.3994  -0.9315  ***
## bin.temp35-40   -2.2142  0.1793  -12.3461  <.0001  -2.5657  -1.8627  ***
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```



Now we completely remove any study that has an effect size in the highest or lowest 2.5%.

```
summary(meta_sub_bintemp)
```

```
##
## Multivariate Meta-Analysis Model (k = 1217; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -2316.9699   4633.9398   4651.9398   4697.8252   4652.0898
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  1.4339  1.1974    309     no  study_code
## sigma^2.2  1.5158  1.2312   1217     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1210) = 24351.0817, p-val < .0001
##
## Test of Moderators (coefficients 1:7):
## QM(df = 7) = 305.6725, p-val < .0001
##
## Model Results:
##
```

```
##                estimate      se      zval    pval    ci.lb    ci.ub
## bin.temp<15     -1.9502  0.1740  -11.2057  <.0001  -2.2914  -1.6091  ***
## bin.temp>40     -2.2312  0.2590   -8.6148  <.0001  -2.7388  -1.7236  ***
## bin.temp15-20   -1.1827  0.1162  -10.1757  <.0001  -1.4105  -0.9549  ***
## bin.temp20-25   -0.6228  0.1101   -5.6567  <.0001  -0.8386  -0.4070  ***
## bin.temp25-30   -0.3410  0.1318   -2.5873  0.0097  -0.5993  -0.0827   **
## bin.temp30-35   -1.1553  0.1061  -10.8840  <.0001  -1.3634  -0.9473  ***
## bin.temp35-40   -1.7366  0.1649  -10.5340  <.0001  -2.0597  -1.4135  ***
##
## ---
## 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.

```
# 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)  # 171
outliers_rep <- subset(gearydata, Trait.category == "Reproduction" &
    gtest < 3)  # 3
```

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.

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

# Resubset reproduction dat
gdata_rep_warm <- subset(s.gearydata, Trait.category == "Reproduction" &
```

```
    warm.cool == "Warm")
gdata_rep_cool <- subset(s.gearydata, Trait.category == "Reproduction" &
    warm.cool == "Cool")

all_gdata <- rbind(gdata_rep_warm, gdata_rep_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$Paper.code)  # Model requires column names study_code
all_gdata$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)
```

```
summary(meta_ga_treat2)
```

```
##
## Multivariate Meta-Analysis Model (k = 1198; method: REML)
##
##     logLik    Deviance         AIC         BIC        AICc
## -3224.0379   6448.0758   6458.0758   6483.5053   6458.1263
##
## Variance Components:
##
##             estim    sqrt  nlvls  fixed      factor
## sigma^2.1  6.2549  2.5010    330     no  study_code
## sigma^2.2  4.0534  2.0133   1198     no         obs
##
## Test for Residual Heterogeneity:
## QE(df = 1195) = 33797.9756, p-val < .0001
##
## Test of Moderators (coefficients 2:3):
## F(df1 = 2, df2 = 1195) = 71.8333, p-val < .0001
##
## Model Results:
##
##                                          estimate      se      tval    df    pval    ci.lb    ci.u
## intrcpt                                   -0.4916  0.1698   -2.8947   327  0.0041  -0.8257  -0.157
## poly(c_treattemp, degree = 2, raw = TRUE)1  -0.0580  0.0097   -5.9797  1195  <.0001  -0.0770  -0.039
## poly(c_treattemp, degree = 2, raw = TRUE)2  -0.0114  0.0010  -11.2330  1195  <.0001  -0.0134  -0.009
##
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Below is a bubble plot of the fitted quadratic model.