

Model Selection

Gan Yao

2023-01-20

This analysis conducts model selection by AIC on models that are of our particular interest.

Aaster graph

root → *flCt* → *flCtNotConsumed* → *flCtUndamaged* → *capsuleCt* → *isHarvested* → *ovuleCt* → *embryoCt*(*fitness*)

Load library and data, transfer data

```
library(aster)
library(tidyverse)

data <- read.csv("data/output/remLilium2021Data30Nov2022.csv")
names(data)

## [1] "id"          "site"        "year"
## [4] "Ax"          "Ly"          "flCt"
## [7] "capsuleCt"   "nCapsulesHarvested" "ovuleCt"
## [10] "embryoCt"    "nn1Dist"     "nn2Dist"
## [13] "nn3Dist"     "nn4Dist"     "nn5Dist"
## [16] "nn6Dist"     "nn7Dist"     "nn8Dist"
## [19] "nn9Dist"     "nn10Dist"    "nn1DistNotConsumed"
## [22] "nn2DistNotConsumed" "nn3DistNotConsumed" "nn4DistNotConsumed"
## [25] "nn5DistNotConsumed" "nn6DistNotConsumed" "nn7DistNotConsumed"
## [28] "nn8DistNotConsumed" "nn9DistNotConsumed" "nn10DistNotConsumed"
## [31] "fecundity"    "flCtNotConsumed" "flCtUndamaged"

data <- data[data$site != "lf",]
data <- data[data$site != "wrrx",]
data[is.na(data$nCapsulesHarvested), 'nCapsulesHarvested'] <- 0
data[is.na(data$ovuleCt), 'ovuleCt'] <- 0
data[is.na(data$embryoCt), 'embryoCt'] <- 0
names(data)[names(data) == 'nCapsulesHarvested'] <- 'isHarvested'

pred <- c(0,1,2,3,4,5,6)
fam <- c(2,1,1,1,1,2,1)
vars <- c("flCt", "flCtNotConsumed", "flCtUndamaged", "capsuleCt",
          "isHarvested", "ovuleCt", "embryoCt")
test <- data %>% mutate(nn5Dist_s = nn5Dist/1000,
                       nn5DistNotConsumed = replace_na(nn5DistNotConsumed, 0)) %>%
  mutate(nn5DistNotConsumed_s = nn5DistNotConsumed/1000)

redata <- reshape(test, varying = list(vars), direction="long", timevar="varb",
```

```

times = as.factor(vars), v.names="resp")

redata <- data.frame(redata, root = 1)
redata$fit <- as.numeric(redata$varb == "embryoCt")
redata$Nid <- as.numeric(gsub("[^0-9.-]", "", redata$id))

names(redata)

## [1] "id" "site" "year"
## [4] "Ax" "Ly" "nn1Dist"
## [7] "nn2Dist" "nn3Dist" "nn4Dist"
## [10] "nn5Dist" "nn6Dist" "nn7Dist"
## [13] "nn8Dist" "nn9Dist" "nn10Dist"
## [16] "nn1DistNotConsumed" "nn2DistNotConsumed" "nn3DistNotConsumed"
## [19] "nn4DistNotConsumed" "nn5DistNotConsumed" "nn6DistNotConsumed"
## [22] "nn7DistNotConsumed" "nn8DistNotConsumed" "nn9DistNotConsumed"
## [25] "nn10DistNotConsumed" "fecundity" "nn5Dist_s"
## [28] "nn5DistNotConsumed_s" "varb" "resp"
## [31] "root" "fit" "Nid"

redata$Deer <- as.numeric(redata$varb=="flCtNotConsumed")
redata$Pollination <- as.numeric(is.element(redata$varb,
c("capsuleCt", "isHarvested", "ovuleCt", "embryoCt")))

names(redata)

## [1] "id" "site" "year"
## [4] "Ax" "Ly" "nn1Dist"
## [7] "nn2Dist" "nn3Dist" "nn4Dist"
## [10] "nn5Dist" "nn6Dist" "nn7Dist"
## [13] "nn8Dist" "nn9Dist" "nn10Dist"
## [16] "nn1DistNotConsumed" "nn2DistNotConsumed" "nn3DistNotConsumed"
## [19] "nn4DistNotConsumed" "nn5DistNotConsumed" "nn6DistNotConsumed"
## [22] "nn7DistNotConsumed" "nn8DistNotConsumed" "nn9DistNotConsumed"
## [25] "nn10DistNotConsumed" "fecundity" "nn5Dist_s"
## [28] "nn5DistNotConsumed_s" "varb" "resp"
## [31] "root" "fit" "Nid"
## [34] "Deer" "Pollination"

```

Null Model

```

model.null <- aster(resp ~ -1 + varb,
  pred, fam,varb,id,root,data=redata)
#summary(model3, info.tol=1e-12)
summary(model.null)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb, pred = pred, fam = fam,
##   varvar = varb, idvar = id, root = root, data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt   -1.135e+00  1.703e-01  -6.666 2.64e-11 ***
## varbembryoCt    -5.784e-01  1.189e-02 -48.645 < 2e-16 ***

```

```
## varbflCt          -2.478e-01  6.063e-02  -4.087  4.36e-05 ***
## varbflCtNotConsumed -6.847e-02  1.075e-01  -0.637   0.524
## varbflCtUndamaged   1.366e+00  1.090e-01  12.531  < 2e-16 ***
## varbisHarvested    -3.215e+02  1.855e+00 -173.370  < 2e-16 ***
## varbovuleCt        6.027e+00  7.127e-03  845.591  < 2e-16 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

eigen(model.null$fisher)$val

## [1] 1.125517e+07 6.263672e+03 1.179341e+03 1.694817e+02 5.191389e+01
## [6] 3.340106e+01 2.906923e-01
```

Previous Models

```
model.P1 <- aster(resp ~ -1 + varb + fit:nn5Dist_s,
                  pred, fam,varb,id,root,data=redata)

summary(model.P1)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb + fit:nn5Dist_s, pred = pred,
##   fam = fam, varvar = varb, idvar = id, root = root, data = redata)
##
##              Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt    -1.135e+00  1.703e-01  -6.666 2.64e-11 ***
## varbembryoCt     -5.796e-01  1.191e-02 -48.664  < 2e-16 ***
## varbflCt         -2.478e-01  6.063e-02  -4.087  4.36e-05 ***
## varbflCtNotConsumed -6.847e-02  1.075e-01  -0.637   0.5241
## varbflCtUndamaged   1.366e+00  1.090e-01  12.531  < 2e-16 ***
## varbisHarvested    -3.215e+02  1.855e+00 -173.358  < 2e-16 ***
## varbovuleCt        6.027e+00  7.127e-03  845.591  < 2e-16 ***
## fit:nn5Dist_s      8.131e-02  3.454e-02   2.354   0.0186 *
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

model.P2 <- aster(resp ~ -1 + varb + fit:nn5DistNotConsumed_s,
                  pred, fam,varb,id,root,data=redata)

summary(model.P2)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb + fit:nn5DistNotConsumed_s,
##   pred = pred, fam = fam, varvar = varb, idvar = id, root = root,
##   data = redata)
##
##              Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt    -1.135e+00  1.703e-01  -6.666 2.64e-11 ***
## varbembryoCt     -5.809e-01  1.192e-02 -48.732  < 2e-16 ***
## varbflCt         -2.478e-01  6.063e-02  -4.087  4.36e-05 ***
## varbflCtNotConsumed -6.847e-02  1.075e-01  -0.637   0.524
## varbflCtUndamaged   1.366e+00  1.090e-01  12.531  < 2e-16 ***
## varbisHarvested    -3.215e+02  1.855e+00 -173.324  < 2e-16 ***
```

```
## varbovuleCt          6.027e+00  7.127e-03  845.591 < 2e-16 ***
## fit:nn5DistNotConsumed_s 1.066e-01  2.188e-02    4.875 1.09e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

model.P3 <- aster(resp ~ -1 + varb + fit:nn5Dist_s + fit:nn5DistNotConsumed_s,
                  pred, fam,varb,id,root,data=redata)

summary(model.P3)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb + fit:nn5Dist_s + fit:nn5DistNotConsumed_s,
##               pred = pred, fam = fam, varvar = varb, idvar = id, root = root,
##               data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt      -1.135e+00  1.703e-01  -6.666 2.64e-11 ***
## varbembryoCt       -5.807e-01  1.192e-02 -48.709 < 2e-16 ***
## varbflCt          -2.478e-01  6.063e-02  -4.087 4.36e-05 ***
## varbflCtNotConsumed -6.847e-02  1.075e-01  -0.637  0.524
## varbflCtUndamaged   1.366e+00  1.090e-01  12.531 < 2e-16 ***
## varbisHarvested    -3.215e+02  1.855e+00 -173.321 < 2e-16 ***
## varbovuleCt         6.027e+00  7.127e-03  845.591 < 2e-16 ***
## fit:nn5Dist_s       -4.136e-02  4.278e-02  -0.967  0.334
## fit:nn5DistNotConsumed_s 1.250e-01  2.702e-02   4.627 3.71e-06 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Model1: *Deer* ~ 1, *Pollination* ~ *nnA*

```
model1 <- aster(resp ~ -1 + varb + Pollination:nn5Dist_s,
                 pred, fam,varb,id,root,data=redata)

summary(model1)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb + Pollination:nn5Dist_s,
##               pred = pred, fam = fam, varvar = varb, idvar = id, root = root,
##               data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt      -1.136e+00  1.703e-01  -6.667 2.60e-11 ***
## varbembryoCt       -5.789e-01  1.189e-02 -48.677 < 2e-16 ***
## varbflCt          -2.478e-01  6.063e-02  -4.087 4.36e-05 ***
## varbflCtNotConsumed -6.847e-02  1.075e-01  -0.637 0.524080
## varbflCtUndamaged   1.366e+00  1.090e-01  12.531 < 2e-16 ***
## varbisHarvested    -3.215e+02  1.855e+00 -173.348 < 2e-16 ***
## varbovuleCt         6.026e+00  7.130e-03  845.191 < 2e-16 ***
## Pollination:nn5Dist_s 2.720e-02  8.037e-03   3.384 0.000714 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
eigen(model1$fisher)$val
```

```
## [1] 1.126056e+07 1.548561e+04 6.262111e+03 1.179341e+03 1.694817e+02
## [6] 5.191390e+01 3.340110e+01 2.906938e-01
```

Model2: $Deer \sim 1$ and $Pollination \sim nnB$

```
model2 <- aster(resp ~ -1 + varb + Pollination:nn5DistNotConsumed_s,
                pred, fam,varb,id,root,data=redata)
summary(model2)
```

```
##
## Call:
## aster.formula(formula = resp ~ -1 + varb + Pollination:nn5DistNotConsumed_s,
##               pred = pred, fam = fam, varvar = varb, idvar = id, root = root,
##               data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt      -1.136e+00  1.703e-01   -6.668 2.60e-11 ***
## varbembryoCt       -5.793e-01  1.189e-02  -48.706 < 2e-16 ***
## varbflCt          -2.478e-01  6.063e-02   -4.087 4.36e-05 ***
## varbflCtNotConsumed -6.847e-02  1.075e-01   -0.637  0.524
## varbflCtUndamaged   1.366e+00  1.090e-01   12.531 < 2e-16 ***
## varbisHarvested     -3.214e+02  1.855e+00 -173.303 < 2e-16 ***
## varbovuleCt         6.026e+00  7.132e-03  844.981 < 2e-16 ***
## Pollination:nn5DistNotConsumed_s 3.214e-02  5.346e-03   6.013 1.83e-09 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
eigen(model2$fisher)$val
```

```
## [1] 1.126807e+07 3.500802e+04 6.261913e+03 1.179341e+03 1.694817e+02
## [6] 5.191390e+01 3.340111e+01 2.906958e-01
```

Model3: $Deer \sim nnA$ and $Pollination \sim 1$

```
model3 <- aster(resp ~ -1 + varb + Deer:nn5Dist_s,
                pred, fam,varb,id,root,data=redata)
summary(model3)
```

```
##
## Call:
## aster.formula(formula = resp ~ -1 + varb + Deer:nn5Dist_s, pred = pred,
##               fam = fam, varvar = varb, idvar = id, root = root, data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt      -1.135e+00  1.703e-01   -6.666 2.64e-11 ***
## varbembryoCt       -5.784e-01  1.189e-02  -48.645 < 2e-16 ***
## varbflCt          -2.478e-01  6.063e-02   -4.087 4.36e-05 ***
## varbflCtNotConsumed -1.423e-01  1.109e-01   -1.283  0.19932
## varbflCtUndamaged   1.366e+00  1.090e-01   12.531 < 2e-16 ***
## varbisHarvested     -3.215e+02  1.855e+00 -173.370 < 2e-16 ***
## varbovuleCt         6.027e+00  7.127e-03  845.591 < 2e-16 ***
```

```
## Deer:nn5Dist_s      6.122e+00  2.061e+00    2.970  0.00298 **
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

eigen(model3$fisher)$val

## [1] 1.125517e+07 6.263672e+03 1.179404e+03 1.694867e+02 5.191860e+01
## [6] 3.340114e+01 2.906923e-01 2.352754e-01
```

Model4: $Deer \sim nnA$ and $Pollination \sim nnA$

```
model4 <- aster(resp ~ -1 + varb + Deer:nn5Dist_s + Pollination:nn5Dist_s,
               pred, fam,varb,id,root,data=redata)
summary(model4, info.tol=1e-9)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb + Deer:nn5Dist_s + Pollination:nn5Dist_s,
##               pred = pred, fam = fam, varvar = varb, idvar = id, root = root,
##               data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt      -1.136e+00  1.703e-01   -6.667 2.61e-11 ***
## varbembryoCt       -5.787e-01  1.189e-02  -48.663 < 2e-16 ***
## varbflCt          -2.478e-01  6.063e-02   -4.087 4.36e-05 ***
## varbflCtNotConsumed -1.166e-01  1.119e-01   -1.043  0.2971
## varbflCtUndamaged   1.366e+00  1.090e-01   12.531 < 2e-16 ***
## varbisHarvested    -3.215e+02  1.855e+00 -173.356 < 2e-16 ***
## varbovuleCt         6.027e+00  7.130e-03  845.198 < 2e-16 ***
## Deer:nn5Dist_s      4.117e+00  2.500e+00    1.647  0.0997 .
## nn5Dist_s:Pollination 1.794e-02  9.818e-03    1.827  0.0677 .
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

eigen(model4$fisher)$val

## [1] 1.126054e+07 1.545341e+04 6.262534e+03 1.179397e+03 1.694860e+02
## [6] 5.191802e+01 3.340115e+01 2.907060e-01 1.599076e-01
```

Model5: $Deer \sim nnA$ and $Pollination \sim nnB$

```
model5 <- aster(resp ~ -1 + varb + Deer:nn5Dist_s + Pollination:nn5DistNotConsumed_s,
               pred, fam,varb,id,root,data=redata)
summary(model5)

##
## Call:
## aster.formula(formula = resp ~ -1 + varb + Deer:nn5Dist_s + Pollination:nn5DistNotConsumed_s,
##               pred = pred, fam = fam, varvar = varb, idvar = id, root = root,
##               data = redata)
##
##               Estimate Std. Error  z value Pr(>|z|)
## varbcapsuleCt      -1.136e+00  1.703e-01   -6.668 2.60e-11 ***
## varbembryoCt       -5.792e-01  1.189e-02  -48.699 < 2e-16 ***
```

```
## varbflCt -2.478e-01 6.063e-02 -4.087 4.36e-05 ***
## varbflCtNotConsumed -9.378e-02 1.109e-01 -0.846 0.398
## varbflCtUndamaged 1.366e+00 1.090e-01 12.531 < 2e-16 ***
## varbisHarvested -3.214e+02 1.855e+00 -173.303 < 2e-16 ***
## varbovuleCt 6.026e+00 7.132e-03 844.966 < 2e-16 ***
## Deer:nn5Dist_s 2.222e+00 2.334e+00 0.952 0.341
## Pollination:nn5DistNotConsumed_s 2.939e-02 6.157e-03 4.774 1.80e-06 ***
## ---
## Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

eigen(model5$fisher)$val

## [1] 1.126806e+07 3.506450e+04 6.262052e+03 1.179391e+03 1.694856e+02
## [6] 5.191758e+01 3.340118e+01 2.907462e-01 1.834819e-01
```

Model Selection

Wrap up a function for computing AIC of aster models.

```
aster_AIC <- function(mod) {
  return(mod$deviance + 2*length(mod$coefficients))
}
```

```
## resp ~ -1 + varb
## [1] -249884.7
## resp ~ -1 + varb + fit:nn5Dist_s
## [1] -249886.9
## resp ~ -1 + varb + fit:nn5DistNotConsumed_s
## [1] -249899.3
## resp ~ -1 + varb + fit:nn5Dist_s + fit:nn5DistNotConsumed_s
## [1] -249898.2
## resp ~ -1 + varb + Pollination:nn5Dist_s
## [1] -249890.7
## resp ~ -1 + varb + Pollination:nn5DistNotConsumed_s
## [1] -249906.9
## resp ~ -1 + varb + Deer:nn5Dist_s
## [1] -249890.1
## resp ~ -1 + varb + Deer:nn5Dist_s + Pollination:nn5Dist_s
## [1] -249891.1
## resp ~ -1 + varb + Deer:nn5Dist_s + Pollination:nn5DistNotConsumed_s
## [1] -249905.8
```

Model	Formula	AIC	P-values of interaction terms
1	$resp \sim varb$	-249884.7	NA
2	$resp \sim varb + Pollination : nnA$	-249890.7	0.7e-3
3	$resp \sim varb + Pollination : nnB$	-249906.9	1.83e-9
4	$resp \sim varb + Deer : nnA$	-249890.1	0.00298
5	$resp \sim varb + Deer : nnA + Pollination : nnA$	-249891.1	0.0997, 0.0667
6	$resp \sim varb + Deer : nnA + Pollination : nnB$	-249905.8	0.341, 1.8e-6
7	$resp \sim varb + fit : nnA$	-249886.9	0.0186
8	$resp \sim varb + fit : nnB$	-249899.3	1.09e-6
9	$resp \sim varb + fit : nnA + fit : nnB$	-249898.2	0.334, 3.71e-6