Model Selection Report

Loïc Pages & Eric Imbert

2026-02-26

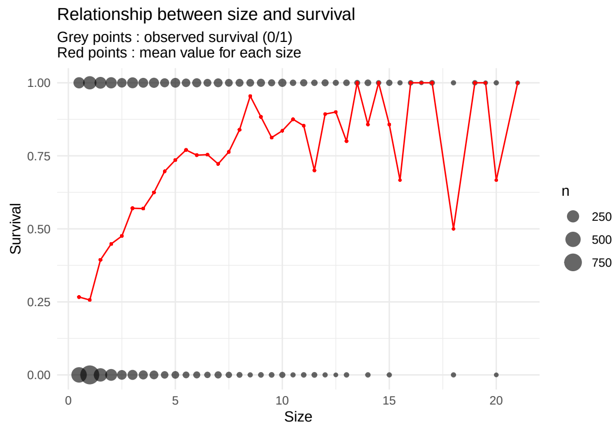
First, graphical representations of raw data and the final selected model for each vital rate Summary of each model and (quasi) explicit formula for splines

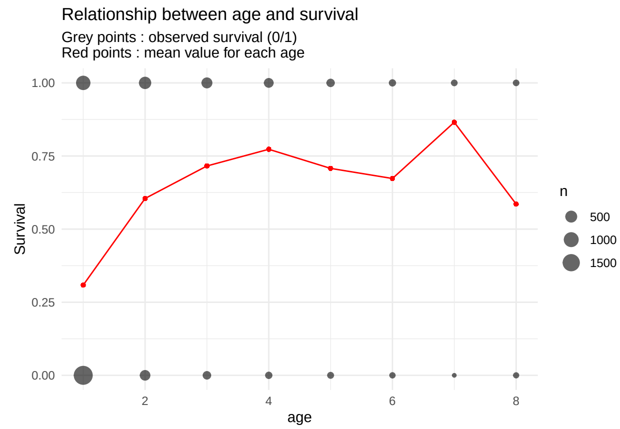
Final step : each model is renamed and a R object is created to be imported in the IPM script

Import data, models

```r
rm(list=ls())

library(knitr)
library(spaMM)
library(tidyverse)
library(foreach)
library(doParallel)
library(splines)

### Required packages, including  `SplinesUtils` not on CRAN
if ( ! requireNamespace("SplinesUtils", quietly=TRUE)) {
  if ( ! requireNamespace("devtools", quietly=TRUE)) install.packages("devtools")
  devtools::install_github("ZheyuanLi/SplinesUtils")
}

#setpath and #getdefSplinesAsPoly
pwd <- paste0(getwd(),"/")
source(paste0(pwd, "SplinesAsPolySpaMM.R"))

IPM_data <- read.csv("../data/donneesIPM.csv")
centauree_data <- IPM_data[!is.na(IPM_data$Size0Mars) & !is.na(IPM_data$Age),]
centauree_data$Age[centauree_data$Age > 8] <- 8

load("SelectedModels.RData")
spaMM.options(separation_max=70)
```

Survival Probability

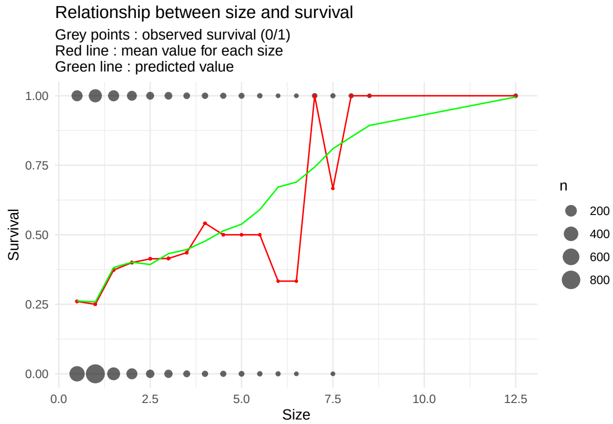




Seedling survival Selected model for seedling survival Seedling_survAIC[[1]]

```
##
## Attachement du package : 'e1071'

## L'objet suivant est masqué depuis 'package:ggplot2':
##
##     element

## If the 'ROI.plugin.glpk' package were installed,
## spaMM could properly check (quasi-)separation in binary regression problem.
## See help('external-libraries') if you have troubles installing 'ROI.plugin.glpk'.
## Alternative procedure 'e1071::svm()' will be used for checking separation.

## formula: Survie ~ bs(Size0Mars, df = 4, degree = 2) + (Size0Mars | year) +
##     1
## Estimation of ranCoefs by ML (p_v approximation of logL).
## Estimation of fixed effects by h-likelihood approximation.
## family: binomial( link = logit )
##  ------------ Fixed effects (beta) ------------
##                                     Estimate Cond. SE t-value
## (Intercept)                          -1.3258   0.2560  -5.179
## bs(Size0Mars, df = 4, degree = 2)1  -0.4160   0.2095  -1.986
## bs(Size0Mars, df = 4, degree = 2)2   0.7485   0.1590   4.709
## bs(Size0Mars, df = 4, degree = 2)3   0.9162   0.8042   1.139
## bs(Size0Mars, df = 4, degree = 2)4   7.1514   3.0511   2.344
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##         --- Random-coefficients Cov matrices:
##  Group        Term    Var.   Corr.
##   year (Intercept)   1.582
##   year   Size0Mars 0.02833 -0.9827
## # of obs: 2842; # of groups: year, 27
##  ------------- Likelihood values  -------------
##                         logLik
##       h-likelihood: -1620.909
## logL       (p_v(h)): -1606.418

## 3 piecewise polynomials of degree 2 are constructed!
## Use 'summary' to export all of them.
## The first 3 are printed below.
## 1.28e-16 - 1.66 * (x - 0.5) + 3.99 * (x - 0.5) ^ 2
```

```
## 0.166 + 2.33 * (x - 1) - 2.3 * (x - 1) ^ 2
## 0.756 + 0.0292 * (x - 1.5) + 0.0502 * (x - 1.5) ^ 2
```



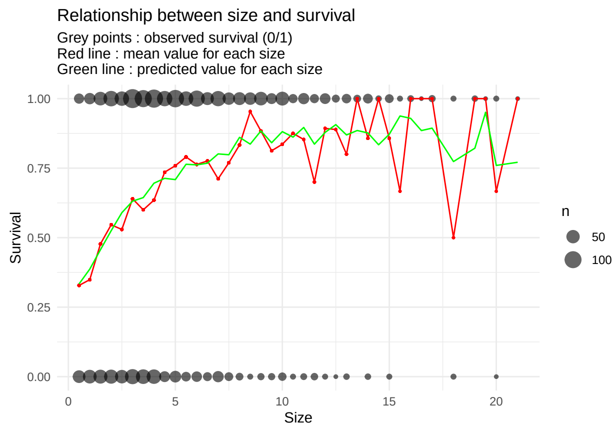
Age > 1 Selected model for plant survival Plant_survAIC[[2]]

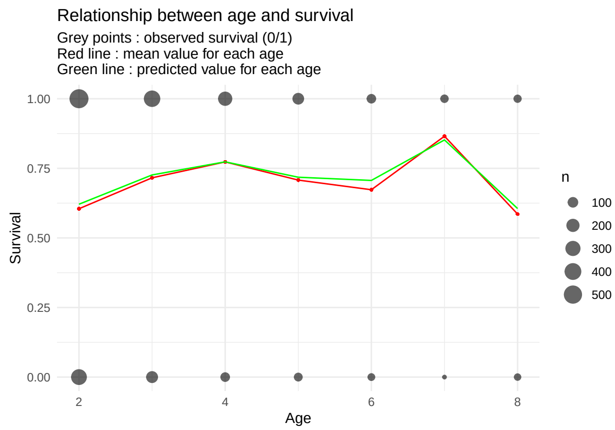
```
## formula: Survie ~ bs(Age, degree = 3, knots = 6.5) + bs(Size0Mars, df = 3,
##     degree = 2) + (1 | Pop) + (Age | year) + 1
## Estimation of lambda and ranCoefs by ML (p_v approximation of logL).
## Estimation of fixed effects by h-likelihood approximation.
## family: binomial( link = logit )
##  ------------ Fixed effects (beta) ------------
##                                   Estimate Cond. SE t-value
## (Intercept)                         -1.426   0.3201  -4.455
## bs(Age, degree = 3, knots = 6.5)1    1.070   0.3859   2.772
## bs(Age, degree = 3, knots = 6.5)2   -1.802   0.6620  -2.722
## bs(Age, degree = 3, knots = 6.5)3    1.083   0.6949   1.559
## bs(Age, degree = 3, knots = 6.5)4   -1.050   0.3301  -3.180
## bs(Size0Mars, df = 3, degree = 2)1    2.024   0.2895   6.991
## bs(Size0Mars, df = 3, degree = 2)2    4.546   0.3978  11.429
## bs(Size0Mars, df = 3, degree = 2)3    3.385   0.8440   4.010
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##          --- Random-coefficients Cov matrices:
##  Group        Term    Var.   Corr.
##   year (Intercept)   1.443
##   year         Age 0.01105 -0.8682
##            --- Variance parameters ('lambda'):
## lambda = var(u) for u ~ Gaussian;
##    Pop  :  0.07594
##              --- Coefficients for log(lambda):
##  Group        Term Estimate Cond.SE
##    Pop (Intercept)   -2.578  0.6928
## # of obs: 2156; # of groups: Pop, 6; year, 26
##  ------------- Likelihood values  -------------
##                         logLik
##         h-likelihood: -1100.556
## logL       (p_v(h)): -1088.777
```

```
## 2 piecewise polynomials of degree 3 are constructed!
## Use 'summary' to export all of them.
## The first 2 are printed below.
```

```
## 3.77e-15 + 0.713 * (x - 2) - 0.478 * (x - 2) ^ 2 + 0.0709 * (x - 2) ^ 3
## 0.000524 + 0.723 * (x - 6.5) + 0.48 * (x - 6.5) ^ 2 - 0.952 * (x - 6.5) ^ 3
```

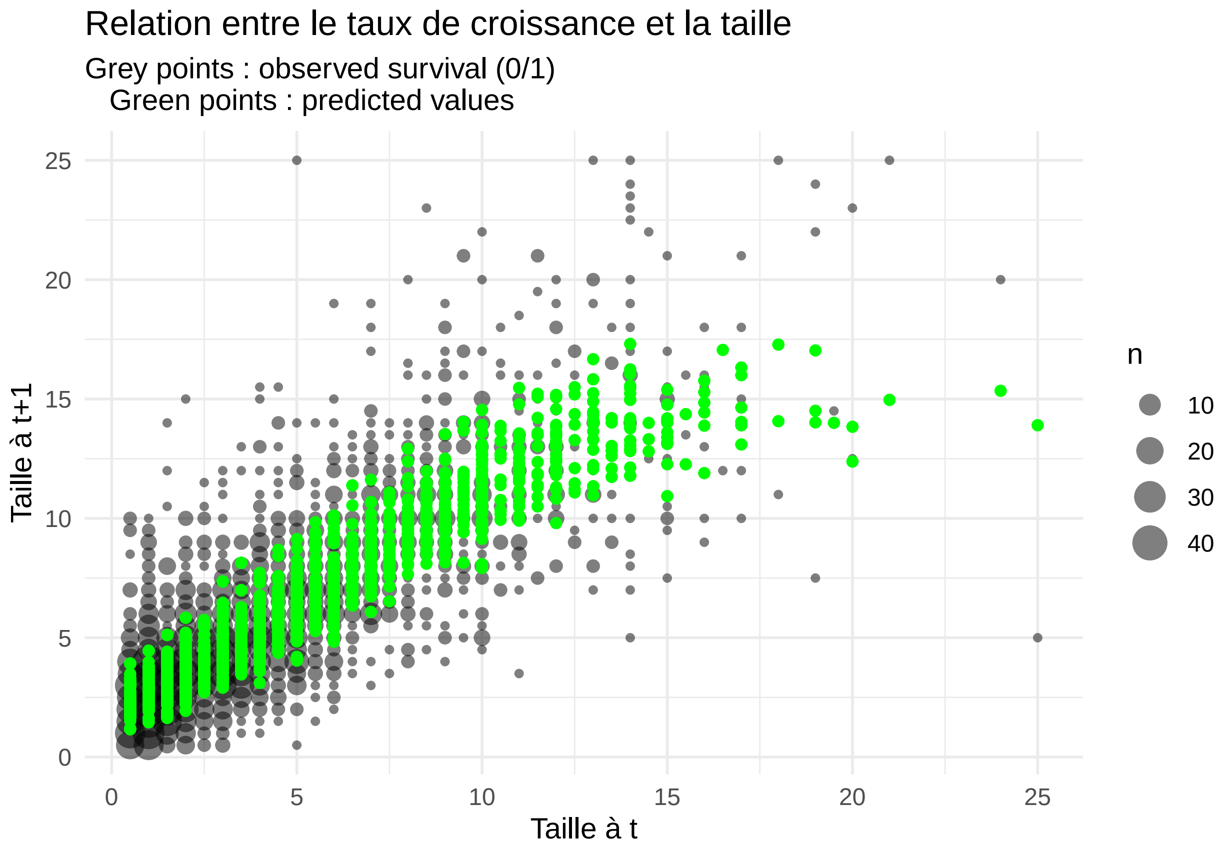
```
## 2 piecewise polynomials of degree 2 are constructed!
## Use 'summary' to export all of them.
## The first 2 are printed below.
## -5.13e-16 + 1.01 * (x - 0.5) - 0.0958 * (x - 0.5) ^ 2
## 2.52 + 0.246 * (x - 4.5) - 0.0117 * (x - 4.5) ^ 2
```





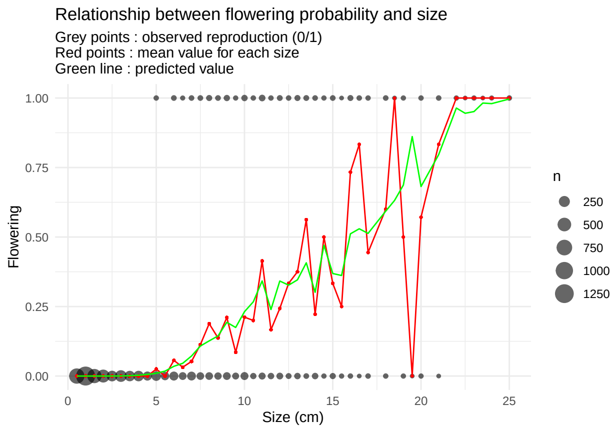
Growth Selected model Growthglm1

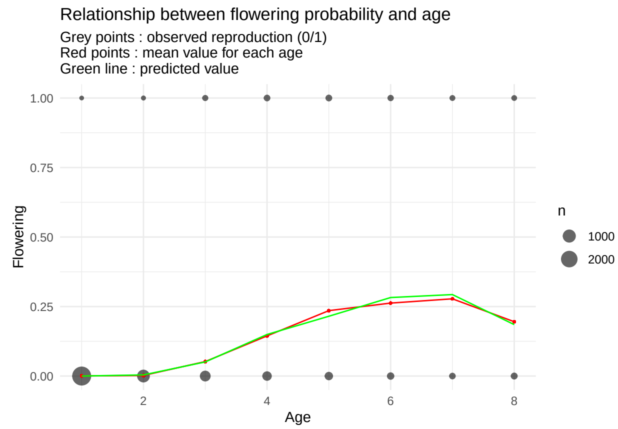
```r
growthdata <- centauree_data[!is.na(centauree_data$Size1Mars), ]
growthdata <- growthdata[growthdata$Size1Mars != 0, ]
Growthpredict <- predict(Growthglm1, newdata = growthdata)[,1]
Growthpredict <- exp(log(Growthpredict)/0.4343354) #back transformation


growthdata %>%
  mutate(predicted_value = Growthpredict) %>%
  ggplot(aes(y = Size1Mars, x = Size0Mars)) +
  geom_count(alpha=0.5) +
  geom_point(aes(y=predicted_value), col="green") +
  labs(title = "Relation entre le taux de croissance et la taille",
       y = "Taille à t+1",
       x = "Taille à t",
     subtitle = "Grey points : observed survival (0/1)
 Green points : predicted values") +
  theme_minimal()
```



Flowering Probability Selected model flowmodAIC[[1]]

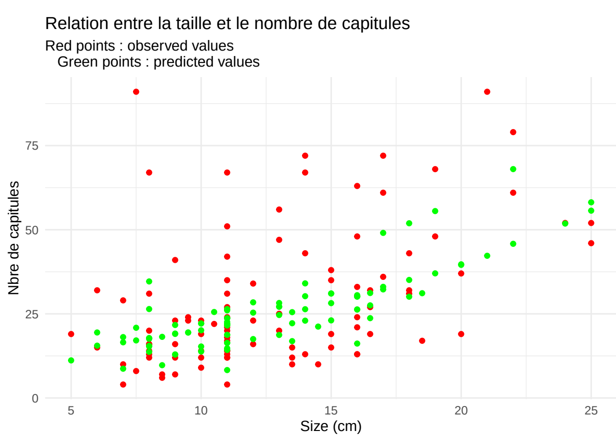
```
## formula: Flowering ~ poly(Age, 2) + poly(Size0Mars, 3) + (Age | Pop) +
##     1
## Estimation of ranCoefs by ML (p_v approximation of logL).
## Estimation of fixed effects by h-likelihood approximation.
## family: binomial( link = logit )
##  ------------ Fixed effects (beta) ------------
##                      Estimate Cond. SE t-value
## (Intercept)            -10.21   0.9991 -10.222
## poly(Age, 2)1          133.48  22.3423   5.974
## poly(Age, 2)2          -54.57   8.8053  -6.197
## poly(Size0Mars, 3)1   226.93  34.0557   6.663
## poly(Size0Mars, 3)2   -76.57  17.1002  -4.478
## poly(Size0Mars, 3)3    39.27  11.3976   3.445
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##          --- Random-coefficients Cov matrices:
##  Group        Term    Var.   Corr.
##    Pop (Intercept)   2.252
##    Pop          Age 0.09228 -0.9823
## # of obs: 5320; # of groups: Pop, 6
##  ------------- Likelihood values  -------------
##                        logLik
##         h-likelihood: -414.4578
## logL       (p_v(h)): -411.1534
```

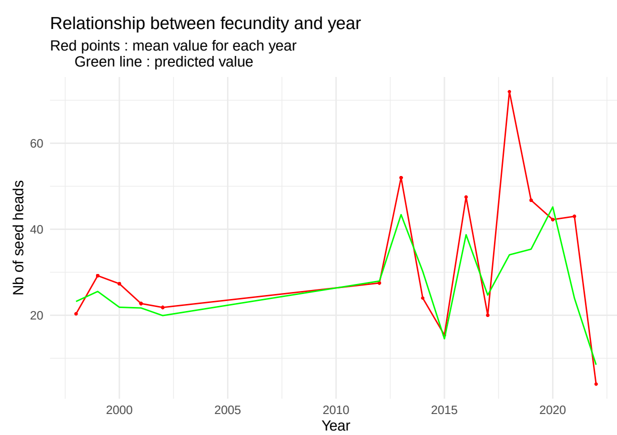




Fecundity

```
## formula: log(Capitule) ~ 1 + Size0Mars + (Age | year)
## ML: Estimation of ranCoefs and phi by ML.
##     Estimation of fixed effects by ML.
## Estimation of phi by 'outer' ML, maximizing logL.
## family: gaussian( link = identity )
##  ------------ Fixed effects (beta) ------------
##              Estimate Cond. SE t-value
## (Intercept)   2.32734   0.2001  11.633
## Size0Mars     0.06752   0.0136   4.963
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##           --- Random-coefficients Cov matrices:
##  Group        Term   Var. Corr.
##   year (Intercept) 0.2127
##   year         Age 0.0238    -1
## # of obs: 95; # of groups: year, 16
##  -------------- Residual variance  ------------
## phi estimate was 0.282245
##  ------------- Likelihood values  -------------
##                         logLik
## logL       (p_v(h)): -82.96256
```

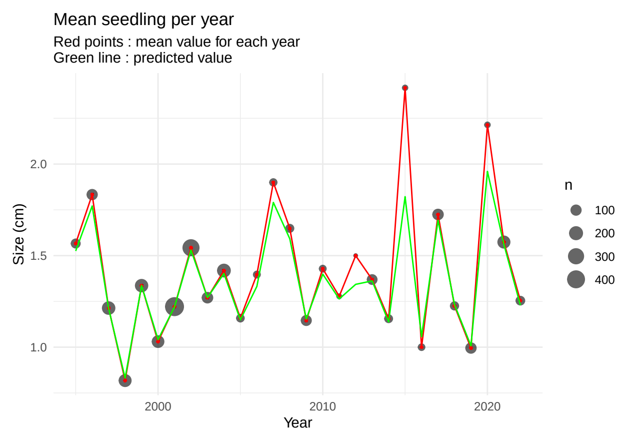


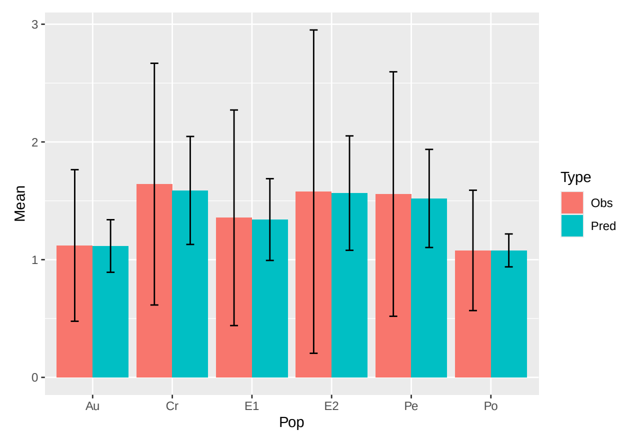


Seedling Size

```
## formula: Size0Mars ~ 1 + (1 | year) + (1 | Pop) + (1 | Pop:year)
```

```
## Estimation of fixed effects by h-likelihood approximation.
## Estimation of lambda and phi by 'outer' ML, maximizing logL.
## family: Gamma( link = log )
##  ------------ Fixed effects (beta) ------------
##            Estimate Cond. SE t-value
## (Intercept)   0.2765  0.07698   3.592
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##            --- Variance parameters ('lambda'):
## lambda = var(u) for u ~ Gaussian;
##    year  :  0.02496
##    Pop  :  0.02486
##    Pop:year  :  0.07811
## # of obs: 2904; # of groups: year, 28; Pop, 6; Pop:year, 142
##  --- Residual variation ( var = phi * mu^2 )  --
## phi estimate was 0.257654
##  ------------- Likelihood values  -------------
##                          logLik
##        h-likelihood: -2509.662
## logL      (P_v(h)):  -2715.887
```






```
## NULL
```

Establishement rate

```
#First replace NA values for number of capitula using predicted values from selected model
cptldata_predi <- IPM_data[IPM_data$Flowering==1,] %>%
  mutate(Capitule = ifelse(is.na(Capitule),exp(predict(Cptlglm1)),Capitule))

#Number of seedlings
plt <- IPM_data %>%
  filter(Age==1) %>%
  group_by(Quadrat,year,Pop) %>%
  summarize(NombrePlantules = sum(Age))
```

```
## `summarise()` has regrouped the output.
## i Summaries were computed grouped by Quadrat, year, and Pop.
## i Output is grouped by Quadrat and year.
## i Use `summarise(.groups = "drop_last")` to silence this message.
## i Use `summarise(.by = c(Quadrat, year, Pop))` for per-operation grouping
##   (`?dplyr::dplyr_by`) instead.
```

```
cptl <- cptldata_predi %>%
  group_by(Quadrat,year,Pop) %>%
  summarize(Capitule = sum(Capitule))
```

```
## `summarise()` has regrouped the output.
## i Summaries were computed grouped by Quadrat, year, and Pop.
## i Output is grouped by Quadrat and year.
## i Use `summarise(.groups = "drop_last")` to silence this message.
## i Use `summarise(.by = c(Quadrat, year, Pop))` for per-operation grouping
##   (`?dplyr::dplyr_by`) instead.
```

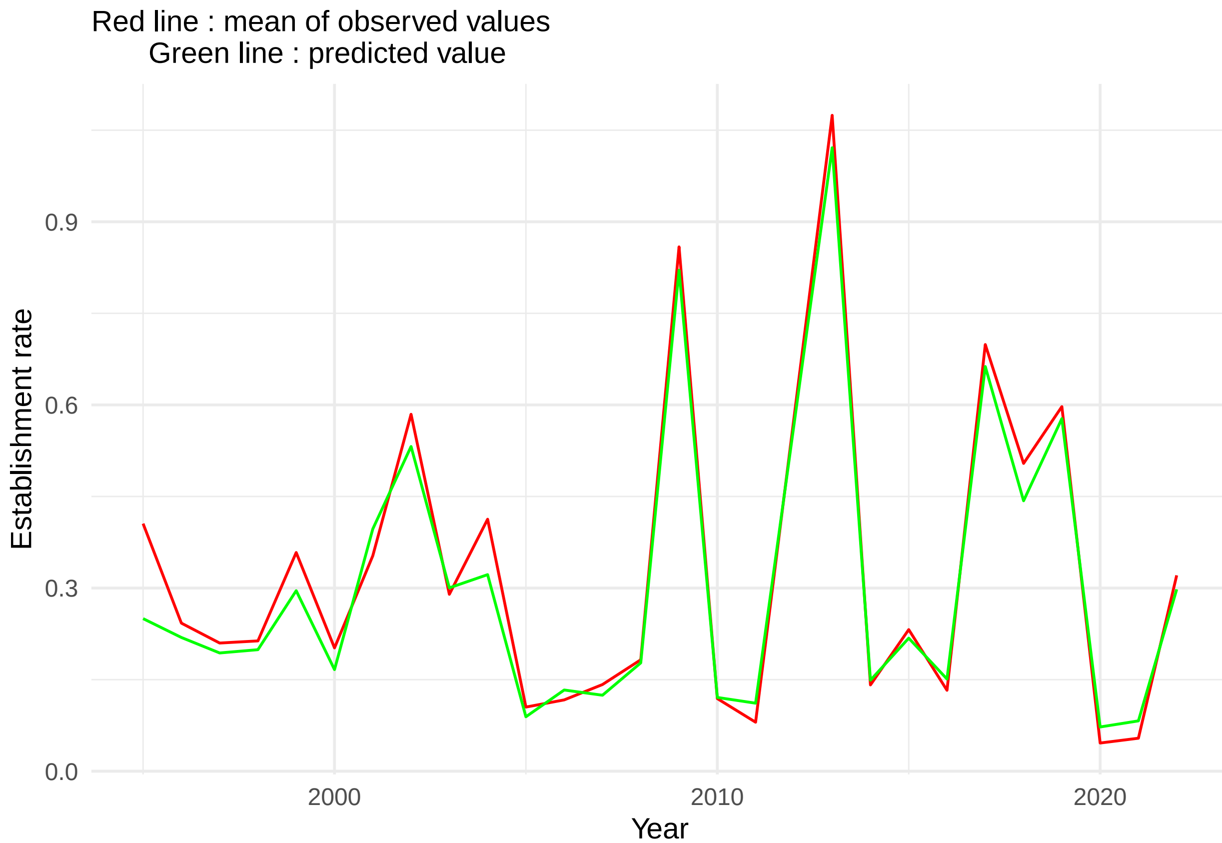
```
cptl$year <- cptl$year+1

Estb <- inner_join(plt,cptl, by=join_by(Quadrat,year,Pop))
Estb$ratio <- Estb$NombrePlantules/Estb$Capitule
```

Number of seedlings observed is dependent on the number of seed heads :

```
Estbglm <- fitme(NombrePlantules ~ 1 + offset(log(Capitule)) + (1|Pop:year),data = Estb,
                 family = Poisson(log))

Estb$pred <- predict(Estbglm, Estb)[,1]
Estb$predratio <- Estb$pred/Estb$Capitule
```

Observed and predicted Values per year

```
Estb %>%
  group_by(year) %>%
  summarise(meanratio = mean(ratio), meanpredratio = mean(predratio)) %>%
  ggplot(aes(x = year, y = meanratio)) +
  geom_line(col="red") +
  geom_line(aes(y=meanpredratio), col="green") +
  labs(y="Establishment rate",
       x="Year",
       subtitle = "Red line : mean of observed values
       Green line : predicted value") +
    theme_minimal()
```



Final name of each model and save to a final object to be imported in the IPM script

```
Seedling_surv
```

```
## formula: Survie ~ bs(Size0Mars, df = 4, degree = 2) + (Size0Mars | year) +
##     1
## Estimation of ranCoefs by ML (p_v approximation of logL).
## Estimation of fixed effects by h-likelihood approximation.
## family: binomial( link = logit )
##  ------------ Fixed effects (beta) ------------
##                                     Estimate Cond. SE t-value
## (Intercept)                          -1.3258   0.2560  -5.179
## bs(Size0Mars, df = 4, degree = 2)1  -0.4160   0.2095  -1.986
## bs(Size0Mars, df = 4, degree = 2)2   0.7485   0.1590   4.709
## bs(Size0Mars, df = 4, degree = 2)3   0.9162   0.8042   1.139
## bs(Size0Mars, df = 4, degree = 2)4   7.1514   3.0511   2.344
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##         --- Random-coefficients Cov matrices:
##  Group        Term     Var.    Corr.
##   year (Intercept)   1.582
##   year   Size0Mars 0.02833 -0.9827
## # of obs: 2842; # of groups: year, 27
##  ------------- Likelihood values  -------------
##                        logLik
##       h-likelihood: -1620.909
## logL       (p_v(h)): -1606.418
```

```r
Plant_surv
```

```
## formula: Survie ~ bs(Age, degree = 3, knots = 6.5) + bs(Size0Mars, df = 3,
##     degree = 2) + (1 | Pop) + (Age | year) + 1
## Estimation of lambda and ranCoefs by ML (p_v approximation of logL).
## Estimation of fixed effects by h-likelihood approximation.
## family: binomial( link = logit )
##  ------------ Fixed effects (beta) ------------
##                                      Estimate Cond. SE t-value
## (Intercept)                            -1.426   0.3201  -4.455
## bs(Age, degree = 3, knots = 6.5)1     1.070   0.3859   2.772
## bs(Age, degree = 3, knots = 6.5)2    -1.802   0.6620  -2.722
## bs(Age, degree = 3, knots = 6.5)3     1.083   0.6949   1.559
## bs(Age, degree = 3, knots = 6.5)4    -1.050   0.3301  -3.180
## bs(Size0Mars, df = 3, degree = 2)1    2.024   0.2895   6.991
## bs(Size0Mars, df = 3, degree = 2)2    4.546   0.3978  11.429
## bs(Size0Mars, df = 3, degree = 2)3    3.385   0.8440   4.010
##  --------------- Random effects ---------------
## Family: gaussian( link = identity )
##          --- Random-coefficients Cov matrices:
##  Group        Term    Var.   Corr.
##   year (Intercept)   1.443
##   year         Age 0.01105 -0.8682
##            --- Variance parameters ('lambda'):
## lambda = var(u) for u ~ Gaussian;
##    Pop  :  0.07594
##              --- Coefficients for log(lambda):
##  Group        Term Estimate Cond.SE
##    Pop (Intercept)   -2.578  0.6928
## # of obs: 2156; # of groups: Pop, 6; year, 26
##  ------------- Likelihood values  -------------
##                         logLik
##        h-likelihood: -1100.556
## logL       (p_v(h)): -1088.777
```

```r
Growth <- Growthglm1
Flowering <- Flowglm
Fecundity <- Cptlglm1
Seedling_size <- Pltglm
Estbrate <- Estbglm

save(Seedling_surv,Plant_surv,Growth,Flowering,Fecundity,Seedling_size,Estbrate,
     file="../IPM/VitalRates.RData")
```