

Applying Ellner approach to leprechaun population dynamic

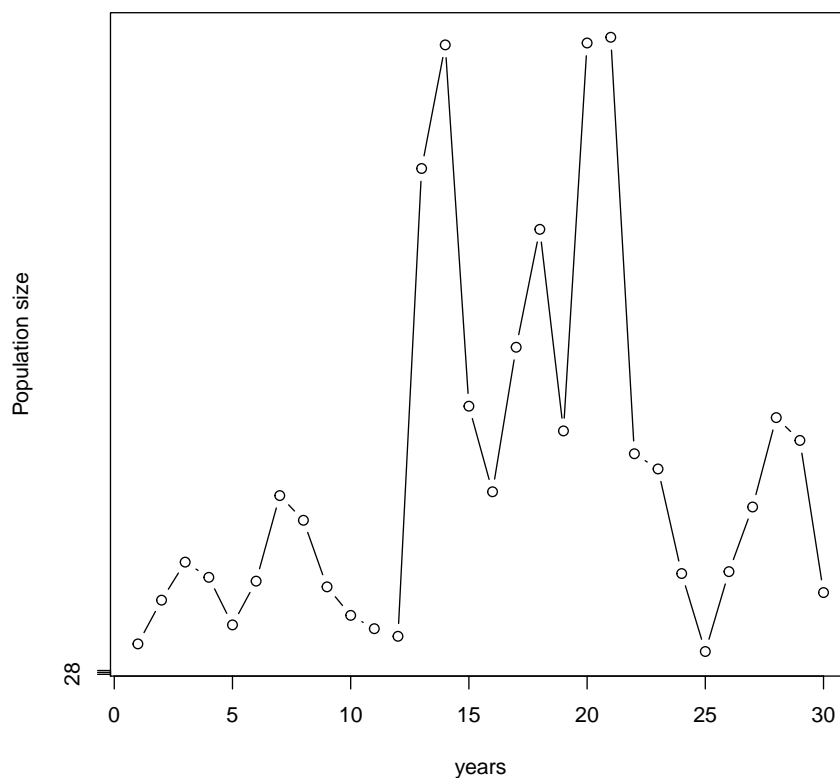
Timothee Bonnet

May 19, 2014

1 A glance at the magical world

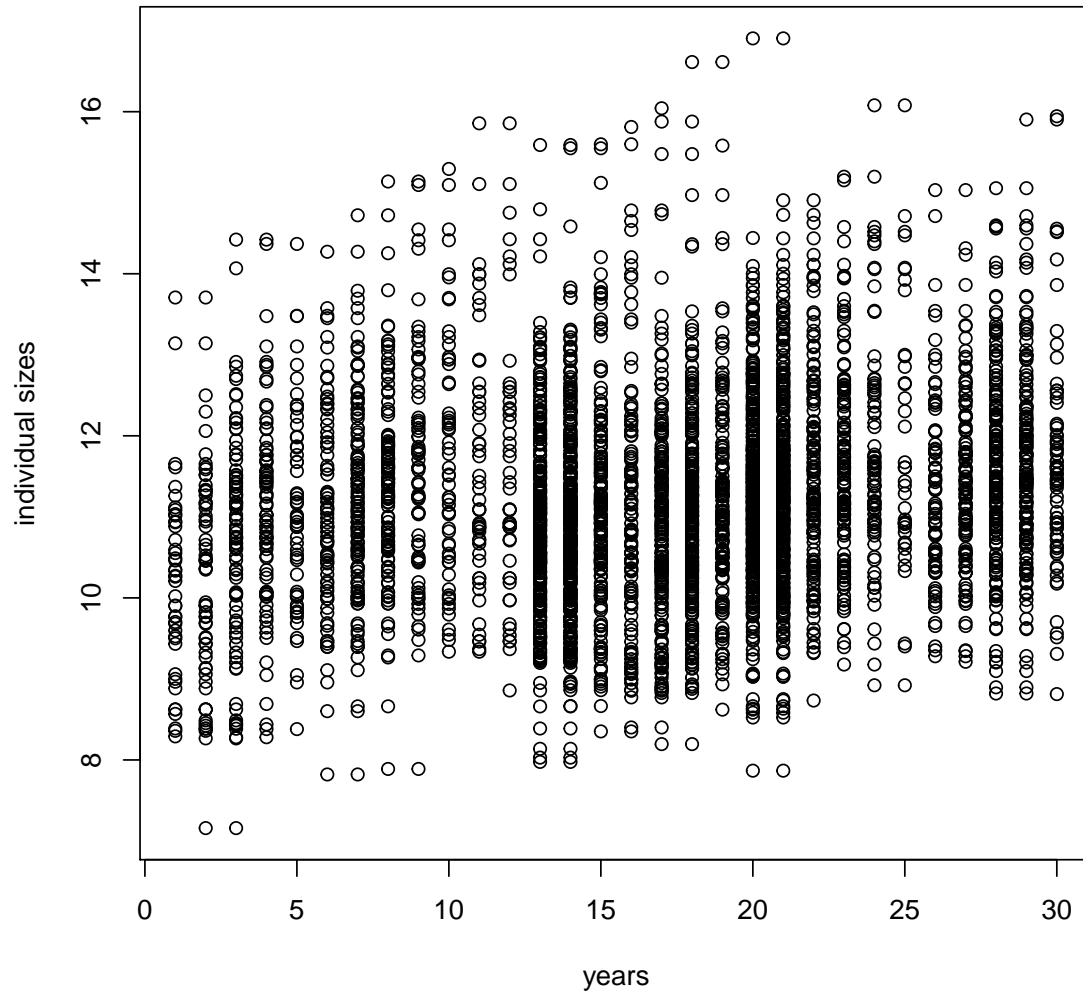
Leprechaun are magical creatures feeding on camembert. In the population of interest competition for camembert is pretty tough and some individual can monopolize most of the ressource on a given year. Camembert ownership affects reproduction and survival probability (it can only decrease it).

Population size changes like that:

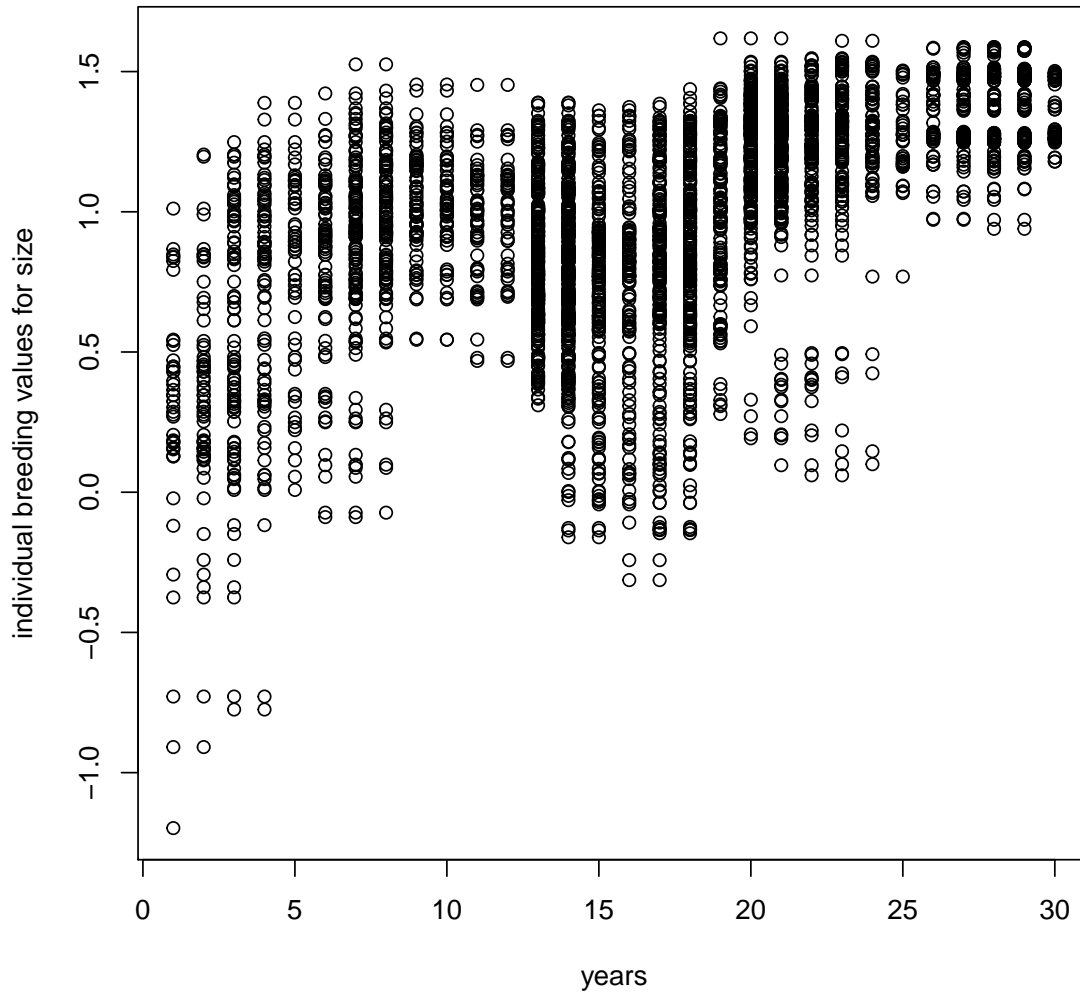


Note the demographic bottleneck on generation 13.

This is the distribution of sizes accross years:



And this is the distribution of breeding values for size at birth accross years.



Note the general decrease in genetic variance due to directional selection on size, but the increase in genetic variance after the population bottleneck. Some people think that it represent a jaguar or a map of Asia, but it is still unclear what we can deduce from this.

2 A decomposition attempt

We will try to decompose the dynamic of juvenile production into its ecological and evolutionary components. *(first I wanted to decompose the population size, until I realized that this is not possible with Ellner approach, because it requires some replication to estimate the partial derivatives of the trait of interest by its components. So I used juvenile production in order to get replication at the level of individual.)* We will consider only adult females in the decomposition of juvenile production. In a first subsection, we will decompose the trait of interest into the effect of size and the effect of camembert ownership. In a second subsection, we will try an extension of Ellner's decomposition method in discrete time to three components, still considering the effect of camembert as previously, but splitting the effect of size into its genetic and plastic components.

At the moment what I am doing is probably wrong, and if it is not, you might not be able to

understand what I am trying to do with the code unless you go back to Ellner's paper. I will try to improve this document soon.

2.1 Phenotype and ressources

This is an attempt to reproduce the method applied in Becks & al. (2012) and decribed in the section *discrete time: comparing endpoints instead of instantaneuous rates* in Ellner & al. (2011). The dynamic of a trait of interest is decomposed into two components, one identified as ecological and one identified as evolutionary. The decomposition is repeated for every pair of successive time points.

2.1.1 the decomposition process

As a reminder this is "Hairston equation":

$$\frac{dX}{dt} = \frac{\partial X}{\partial z} \frac{dz}{dt} + \frac{\partial X}{\partial k} \frac{dk}{dt} \quad (1)$$

Where X is the trait of interest of which we want to decompose the dynamic. z is a trait of the organisms of interest, k is an environmental variable, t is time.

Let us explain each term:

- $\frac{dX}{dt}$ is the temporal dynamic of the trait of interest.
- $\frac{\partial X}{\partial z}$ is the partial effect of z on X . *"what happens to X when z change of 1 unit."*
- $\frac{dz}{dt}$ is the temporal change of z . *"of how many units has z changed with time."*
- thereby $\frac{\partial X}{\partial z} \frac{dz}{dt}$ is the change in X due to the change in z . It can be called the "evolutionary" component of change in X if one assume no phenotypic plasticity.
- $\frac{\partial X}{\partial k}$ is the partial effect of k on X . *"what happens to X when k change of 1 unit."*
- $\frac{dk}{dt}$ is the temporal change of k . *"of how many units has k changed with time."*
- thereby $\frac{\partial X}{\partial k} \frac{dk}{dt}$ is the change in X due to the change in k . It can be called the "ecological" component of change in X .

Koen pointed out that given the use of the equation made by Hairston and Ellner making the assumption that the effet of z on X is linear over the population's z distribution, the second right-side term would be better written $\frac{\partial X}{\partial \bar{z}} \frac{d\bar{z}}{dt}$. In what I have done, I get the first part of the term regressing individual values of X on individual values of z , but I use the second term with the population mean assumption. Therefore I wonder, but I guess this is not correct, I am actually asking for an explanation, whether in my case, the term could be written $\frac{\partial X}{\partial \bar{z}} \frac{d\bar{z}}{dt}$.

In our context, X will be the number of offspring per female on a given year, or mean annual reproductive success (ARS), z will be the animal sizes, k will be the camemebert ownership and t will be time. For each

pair of successive points in time, we want an estimate of $\frac{\partial X}{\partial z} \frac{dz}{dt}$ and one of $\frac{\partial X}{\partial k} \frac{dk}{dt}$, following Becks and Ellner notations. Again, it might be wrong, but a more exact formulation might be

$$\frac{dX}{dt} = \frac{\partial X}{\partial z} \frac{dz}{dt} + \frac{\partial X}{\partial k} \frac{dk}{dt} \quad (2)$$

One problem I might have is that this decomposition does not actually map the decompositions in Becks and Ellner, as X , z and k are all traits observable at the individual levels and not population-level properties. Maybe then the framework is not valid anymore or rather less interesting as the distinction between ecology and evolution is less clear: the environmental parameter is not clearly independent of individual values. However, I could not find a set of parameters with replication in our simulations that would allow me to map exactly Ellner decomposition. I still believe that it is close enough to understand in practice the principals of the decomposition. Because our data are simulated in discrete time, we need to use the discrete time version of Hairston's and Ellner's method. I think that this would also be the case with most empirical studies since having access to instantaneous rates of change seem challenging in practice. The discrete time method keep the the previous equations but use them a bit less directly.

At t we observe $X(z_t, k_t)$ and at $t + 1$ we observe $X(z_{t+1}, k_{t+1})$. Becks et al. following Hairston et al., predict what would be $X(z_t, k_{t+1})$, that is to say X if k would change with z held constant, and $X(z_{t+1}, k_t)$, that is to say X if z would change with k held constant. These predictions and observed X are used to estimate the components of X dynamic, from time t to time $t + 1$, as follow:

$$\frac{\partial X}{\partial z} \frac{dz}{dt} = ((X(z_{t+1}, k_t) - X(z_t, k_t)) + (X(z_{t+1}, k_{t+1}) - X(z_t, k_{t+1}))) / 2 \quad (3)$$

and

$$\frac{\partial X}{\partial k} \frac{dk}{dt} = ((X(z_t, k_{t+1}) - X(z_t, k_t)) + (X(z_{t+1}, k_{t+1}) - X(z_{t+1}, k_t))) / 2 \quad (4)$$

Which is equivalent to a two-way ANOVA explaining X by the two times two crossed levels of z and k . Of course, the difficult point lies in the prediction of the $X(z_t, k_{t+1})$ and the $X(z_{t+1}, k_t)$. They must be based on a good knowledge of the mechanism and on assumptions specific to the system. A possible approach to get them is to fit a linear model $X \sim 1 + z + k$, which gives $\frac{\partial X}{\partial z}$ and one of $\frac{\partial X}{\partial k}$. Which requires some kind of replication of the data, or an independent experiment. From that, one can predict what would be the unobserved X s:

$$X(z_t, k_{t+1}) = \hat{\beta}_1 + \hat{\beta}_z z_t + \hat{\beta}_k k_{t+1} \quad (5)$$

$$X(z_{t+1}, k_t) = \hat{\beta}_1 + \hat{\beta}_z z_{t+1} + \hat{\beta}_k k_t \quad (6)$$

Where β_1 is the intercept and $\hat{\beta}$ is the maximum-likelihood (or equivalently least-squares as we use a linear model) estimate of the parameter β . The ANOVA approach does not let any residual degree of freedom, so there is no information about the accuracy of the estimates. However, the process is repeated numerous times (for each pair of successive time points), so the global relative importance of evolution and ecology is averaged

over many (inaccurate) points.

2.1.2 A leprechaun illustration

We will use only adult females

```
females <- popfile[which(popfile$s == "F" & popfile$age > 0), ]
```

In a first paragraph we will see in details the calculation process for the two first time steps, and in a second paragraph, we will repeat the process for all pairs of successive time steps.

Two first time steps

Let us first detail the method for the first pair of time steps:

```
tempus = 2
mARSO <- glm(ARS ~ 1 + C + z, data = females[which(females$t == (tempus - 1) |
  females$t == tempus), ], family = poisson)
```

This is our linear model $X \sim 1 + z + k$. We use a Poisson distribution with a log link to normalize the ARS distribution. We will back-transform the estimates to estimate the theoretical unobserved values of X .

We extract the intercept, $\hat{\beta}_1$ and the parameter estimates $\hat{\beta}_z$ and $\hat{\beta}_k$,

```
intercept <- mARSO$coefficients[1]
dXdz <- mARSO$coefficients["z"]
dXdK <- mARSO$coefficients["C"]
```

We calculate z_t , k_t , z_{t+1} and k_{t+1} :

```
Zt0 <- mean(females$z[which(females$t == (tempus - 1))])
kt0 <- mean(females$C[which(females$t == (tempus - 1))])
Zt1 <- mean(females$z[which(females$t == tempus)])
kt1 <- mean(females$C[which(females$t == tempus)])
```

And here we get the $X(z_t, k_t)$, $X(z_{t+1}, k_{t+1})$, $X(z_t, k_{t+1})$ and $X(z_t, k_t)$:

```
Xtt <- mean(females$ARS[which(females$t == (tempus - 1))])
# not exactly exp(intercept+Zt0*dXdz+kt0*dXdK) I do not want to account for
# the interaction (??) as it is problematic for XtT and XTt
XTT <- mean(females$ARS[which(females$t == tempus)])
# not exactly exp(intercept+Zt1*dXdz+kt1*dXdK) I do not want to account for
# the interaction (??) as it is problematic for XtT and XTt
XtT <- exp(intercept + Zt0 * dXdz + kt1 * dXdK)
```

```

# this is done with population mean assumption. Would be impossible on an
# individual basis
XtT <- exp(intercept + Zt1 * dXdz + kt0 * dXdk)
# this is done with population mean assumption. Would be impossible on an
# individual basis

```

in the code the lower-case "t" means t and the upper-case "T" means $t + 1$. In the next code chunk these four X values are arranged into a dataframe also containing two columns of dummy variables. The first column contains the X s. The second column indicates the time of the z value causing the X in this row: 0 indicates z_t ; 1 indicates z_{t+1} . The third column indicates the time of the k value causing the X in this row with the same code: 0 indicates z_t ; 1 indicates z_{t+1} .

```

AX <- as.data.frame(matrix(data = c(Xtt, XtT, XtT, XTT, 0, 1, 0, 1, 0, 0, 1,
1), ncol = 3))
# using an anova we get the respective importance of evolution and of
# ecology:

```

The X s in this dataframe are now regressed on these two dummy variables, which give their relative importance and direction.

```

m1X <- lm(V1 ~ V2 + V3, data = AX)
m1X$coefficients["V2"] #evolution

##      V2
## -1.967

m1X$coefficients["V3"] #ecology

##      V3
## -0.5088

# or equivalently:
0.5 * (XTT - XtT + XtT - Xtt) #evolution

## (Intercept)
##      -1.967

0.5 * (XTT - XTt + XtT - Xtt) #ecology

## (Intercept)
##      -0.5088

```

That's it, the "evolution" coefficient is $\frac{\partial X}{\partial z} \frac{d\bar{z}}{dt}$ and the "ecology" coefficient is $\frac{\partial X}{\partial k} \frac{d\bar{k}}{dt}$.

All time steps

These two vectors will save the "evolutionary" and "ecological" contributions change between every pair of time steps for later:

```
evol <- vector(length = max(females$t) - 1)
ecol <- vector(length = max(females$t) - 1)
```

We repeat the exact same procedure as in the previous paragraph, but for every pair of successive time steps:

```
for (tempus in 2:max(females$t))#for every time step
{
  mARSO<-glm(ARS~1+C+z,
             data=females[which(females$t==(tempus-1) | females$t==tempus),],
             family=poisson)

  intercept<-mARSO$coefficients[1]
  dXdz<-mARSO$coefficients["z"]
  dXdk<-mARSO$coefficients["C"]

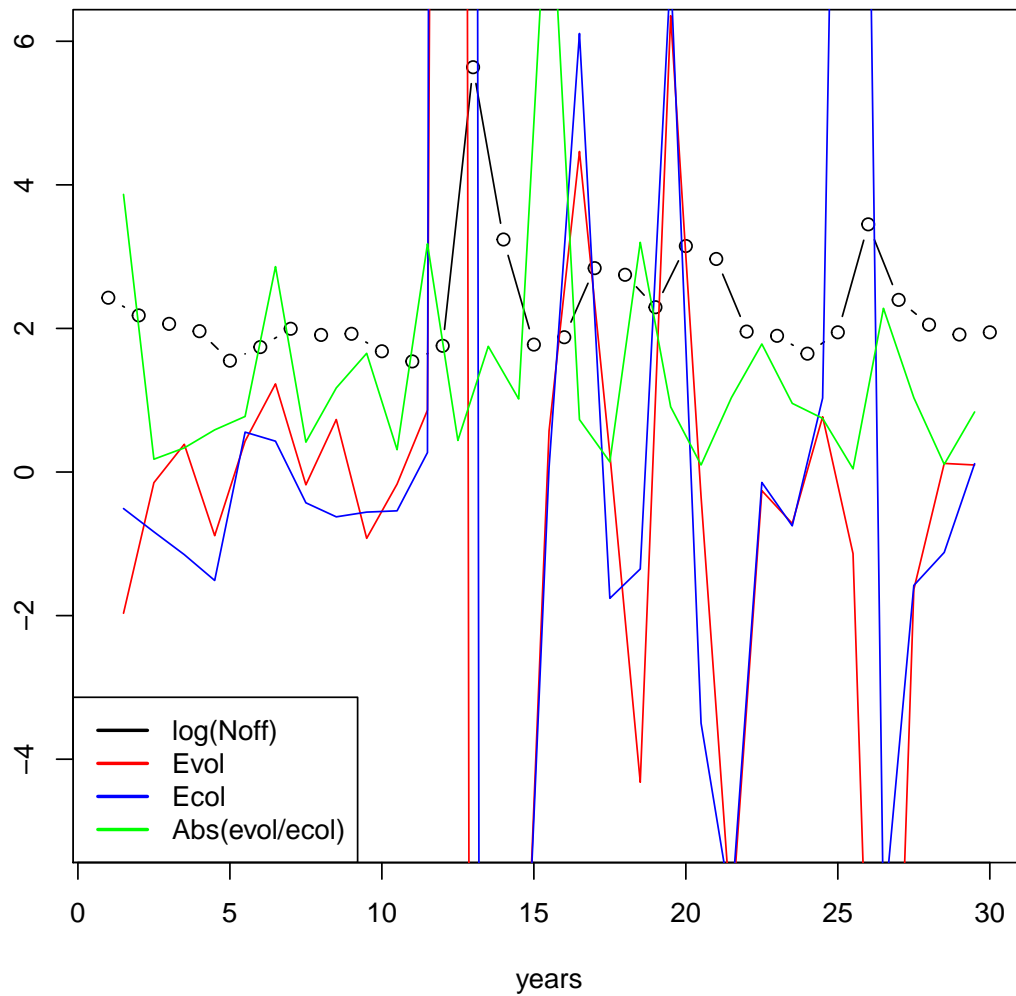
  Zt0<-mean(females$z[which(females$t==(tempus-1))])
  kt0<-mean(females$C[which(females$t==(tempus-1))])
  Zt1<-mean(females$z[which(females$t==tempus)])
  kt1<-mean(females$C[which(females$t==tempus)])

  Xtt<-mean(females$ARS[which(females$t==(tempus-1))])
  XTT<-mean(females$ARS[which(females$t==tempus)])
  XtT<-exp(intercept+Zt0*dXdz+kt1*dXdk)
  XTt<-exp(intercept+Zt1*dXdz+kt0*dXdk)

  AX<-as.data.frame(matrix(data=c(Xtt,XTt,XtT,XTT,0,1,0,1,0,0,1,1),ncol=3))
  #using an anova we get the respective importance of evolution and of ecology:
  m1X<-lm(V1~V2+V3,data=AX)
  evol[tempus-1]<-m1X$coefficients["V2"]
  ecol[tempus-1]<-m1X$coefficients["V3"]
}
```

A reassuring thing is that we can detect the bottleneck episode on generation 13 quite clearly by looking at the components of trait change.(This is not surprising given that what happened is completely crazy and

unrealistic: a single female bore 284 pups while all the other adult females got 0 and died of starvation.)



The relative importance of evolution averaged over all years is:

```
mean(abs(evol)/(abs(evol) + abs(ecol)))  
## [1] 0.4579
```

Yes, this is a beautiful number. I am very happy I managed to get it. Now, does somebody have any idea about what this means?

2.2 Evolution, plasticity and ressources

We now consider phenotypic plasticity for size as a component distinct from evolution. We use breeding values to describe the evolutionary component, which is terribly unrealistic, as the breeding values are not observable. I will not give so many details as the principle is exactly the same as in the previous section, but expanded to three, instead of two, components.

```

evol<-vector(length=max(females$t)-1)
plas<-vector(length=max(females$t)-1)
ecol<-vector(length=max(females$t)-1)
for (tempus in 2:max(females$t))
{
  mARS1<-glm(ARS~1+C+z+bvs,
             data=females[which(females$t==(tempus-1) | females$t==tempus),],
             family=poisson)

  intercept<-mARS0$coefficients[1]
  dXdz<-mARS1$coefficients["z"]
  dXdk<-mARS1$coefficients["C"]
  dXdg<-mARS1$coefficients["bvs"]

  Zt0<-mean(females$z[which(females$t==(tempus-1))])
  kt0<-mean(females$C[which(females$t==(tempus-1))])
  gt0<-mean(females$bvs[which(females$t==(tempus-1))])
  Zt1<-mean(females$z[which(females$t==tempus)])
  kt1<-mean(females$C[which(females$t==tempus)])
  gt1<-mean(females$bvs[which(females$t==tempus)])

  Xttt<-mean(females$ARS[which(females$t==(tempus-1))])
  XTtT<-mean(females$ARS[which(females$t==tempus)])
  XttT<-exp(intercept+Zt0*dXdz+gt0*dXdg+kt1*dXdk)
  XtTt<-exp(intercept+Zt0*dXdz+gt1*dXdg+kt0*dXdk)
  XtTT<-exp(intercept+Zt0*dXdz+gt1*dXdg+kt1*dXdk)
  XTtt<-exp(intercept+Zt1*dXdz+gt0*dXdg+kt0*dXdk)
  XTtT<-exp(intercept+Zt1*dXdz+gt0*dXdg+kt1*dXdk)
  XTtT<-exp(intercept+Zt1*dXdz+gt1*dXdg+kt0*dXdk)

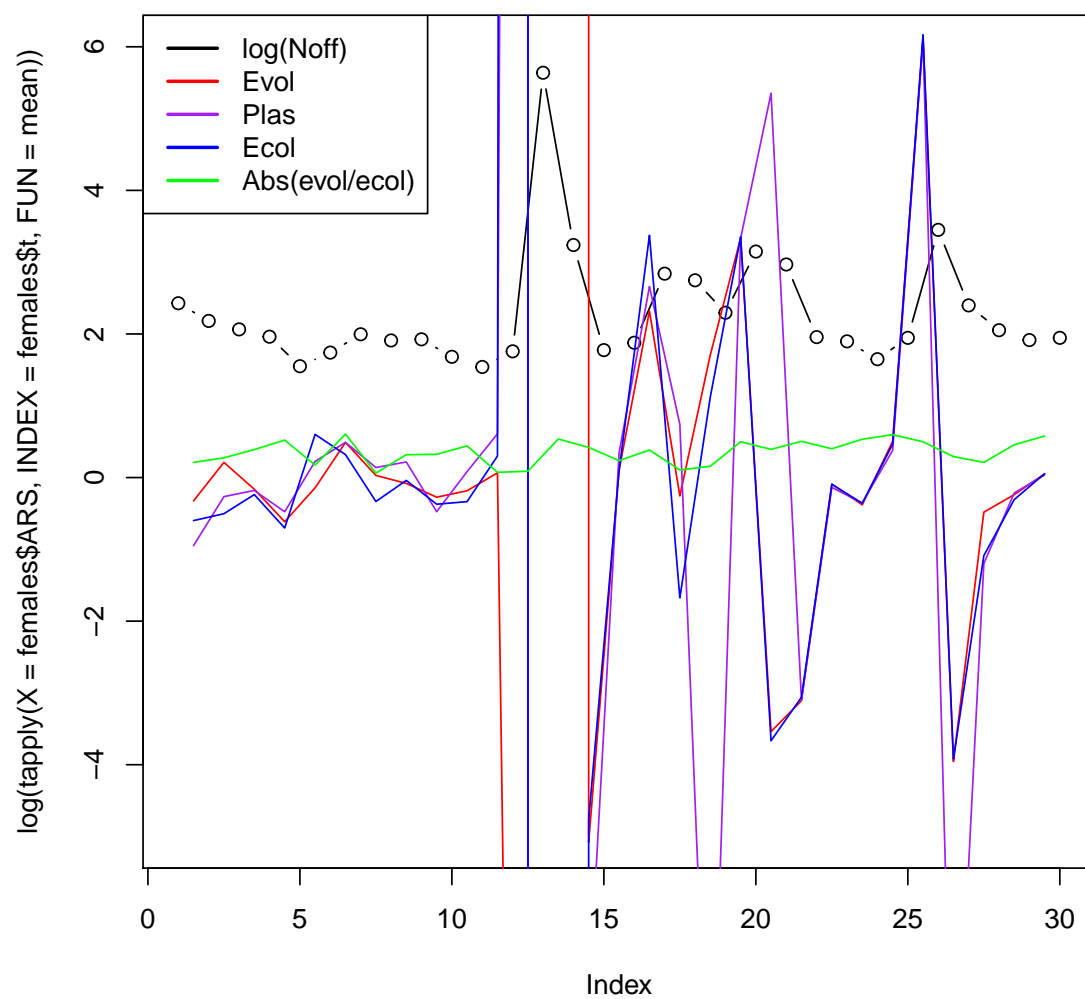
  AX<-data.frame(c(Xttt,XttT,XtTt,XtTT,XTtt,XTtT,XTTt,XTTT),# the Xs
                 c(0,0,0,0,1,1,1,1),# code for the time of mean z
                 c(0,0,1,1,0,0,1,1),# code for the time of mean breeding value
                 c(0,1,0,1,0,1,0,1))# code for the time of mean k

  names(AX)<-c("X","Z","G","K")
  m1X<-lm(X~Z+G+K,data=AX)

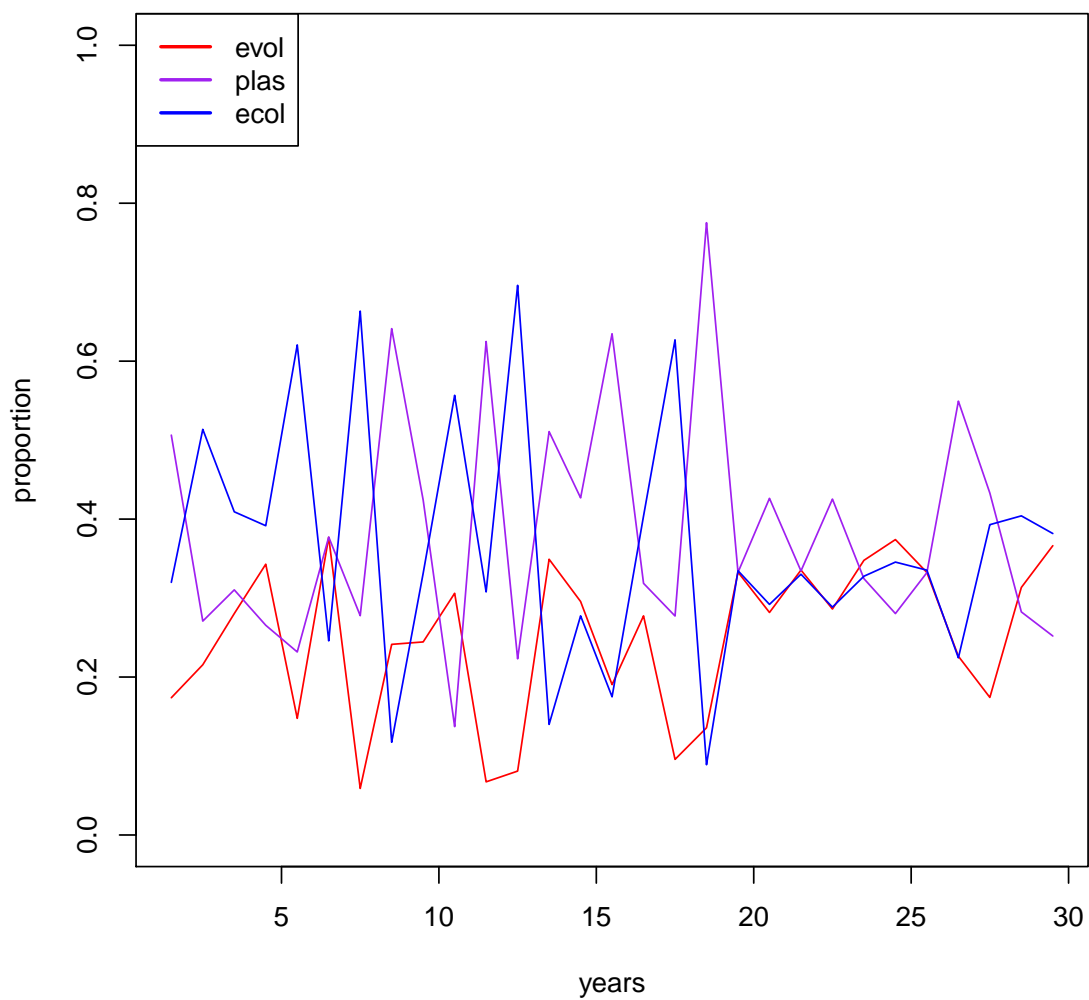
```

```
evol[tempus-1]<-m1X$coefficients["G"]#breeding values
plas[tempus-1]<-m1X$coefficients["Z"]#phenotype
ecol[tempus-1]<-m1X$coefficients["K"]#environment
}
```

The absolute effect of the three components:



The change in relative importance of the components (sum to 1 at any time):



The mean relative importance of evolution is now:

```
mean(abs(evolver)/(abs(evolver) + abs(ecol) + abs(plas)))
```

```
## [1] 0.25
```

This is two time less than without considering size plasticity, but I am not more comfortable with its meaning.