

# Lab7 - Natural selection

Martina Martini s306163

Politecnico di Torino

---

**Assignment** Consider a simulator for natural selection with the following simplified simulation model:

All the individuals belong to  $S$  different species. Let  $s$  be the index of each species,  $s=1\dots S$ . The initial population of  $s$  is equal to  $P(s)$ . The reproduction rate for each individual is  $\lambda$ . The theoretical lifetime  $LF(k)$  of individual  $k$  whose parent is  $d(k)$  is distributed according to the following distribution:  $LF(k) = \text{uniform}(LF(d(k)), LF(d(k)) \cdot (1 + \alpha))$  with probability  $\text{prob\_improve}$  and  $\text{uniform}(0, LF(d(k)))$  with probability  $1 - \text{prob\_improve}$ , where  $\text{prob\_improve}$  is the probability of improvement for a generation and  $\alpha$  is the improvement factor ( $\geq 0$ ). The individuals move randomly in a given region and when individuals of different species meet, they may fight and may not survive. In such a case, the actual lifetime of a individual may be lower than its theoretical lifetime. A died individual cannot reproduce.

---

## 1 Random elements, inputs and main assumptions

---

### 1.1 Introduction

The simulation of a basic natural selection mechanism brings us to pose some questions. Some of them may be the following. Which parameters are more effective in terms of time to the extinction of the species and the average lifetime of the population? In particular, is more impactful the varying of the extension of the area in which they live or of the probability of improvement? How many fights for survival happen in order to live as long as possible? But overall, more generally, what are the parameters more correlated between them and how do they influence each other? To answer all of these questions, a simulation model is implemented as explained in the following sections.

### 1.2 Stochastic elements

To simulate such a natural selection model, some stochastic elements must be considered. Firstly, during the generation of the individual, a binomial distribution related to the singular species assigns a probability that its gender is female or male. Such a probability is given as input and it is different according to the species, in order to be as reliable as possible to the reality. Secondly, at the trigger of the event *born*, the choice of which individual of the population must reproduce itself is random. Another stochastic parameter is related to the strength of an individual: another binomial distribution over the species assigns the boolean variable  $[True, False]$  to the *strong* or *weak* attribute according to the fraction parameter fixed as input. As assigned, the improvement of the newborn depends on the probability *probability\_to\_improve* and, consequently, the lifetime of the child depends on this probability. In fact, if there is an improvement, its

lifetime is uniformly distributed between the lifetime of the father and the lifetime of the father multiplied by  $(1 + \alpha)$ ; instead, if there is not an improvement, its lifetime is uniformly distributed between 0 and the lifetime of the parent. Furthermore, as requested, the inter-birth time (i.e., the time between a *born* event and the following one) is distributed as a Poisson process with rate *reproduction\_rate*. Concerning this point, the reproduction rate is interpreted as a global parameter, equal for each individual independently to the species and fixed as input. Moreover, when a species overcomes the maximum limit of individuals allowed for the specific area in which the species is settled, a random selection of individuals moves toward a randomly chosen destination. Such a destination consists of one of the other areas, originally occupied by the other species. When the selected warriors arrive in that area, the fight for survival begins and, in the end, just a fraction of the winning individuals survive. The same happens for the losers, but in a smaller percentage. To conclude, a random seed is introduced for reproducibility purposes.

### 1.3 Input parameters

For what concerns the input parameters, I assume to set the following fixed variables: the initial population size is equal to 10,  $\alpha$  is always set to 0.2 and the standard reproduction rate is fixed at  $1/365$  days. As mentioned above, in case of a fight for survival the ratio of winning survivors is 0.9, while the one for the losers is 0.2. The simulation model also presents an additional percentage of lifetime assigned to the winners (0.2 times the individual lifetime) and a subtractive one for the losers (0.1 times the individual lifetime). The parameters taken as fixed for each species are listed in

Species	Fraction of strong individuals	Fraction of female individuals	Speed (Km/h)	Initial LF (days)
Wolves	0.9	0.5	56	15 * 365
Eagles	0.8	0.47	120	20 * 365
Deers	0.5	0.47	51	6 * 365
Foxes	0.4	0.46	51	4 * 365
Mouses	0.1	0.46	13	5 * 365

**Table 1:** Input parameters divided by species

the Table 1. Finally, to evaluate the system’s behavior when changing the value of some parameters, I decided to vary one of them and keep fixed the others. When varying, the single input is taken from a list of values: the wood extension is between 100 and 10000 kilometers while the probability of improvement is between 0.1 and 0.9.

## 1.4 Main assumptions

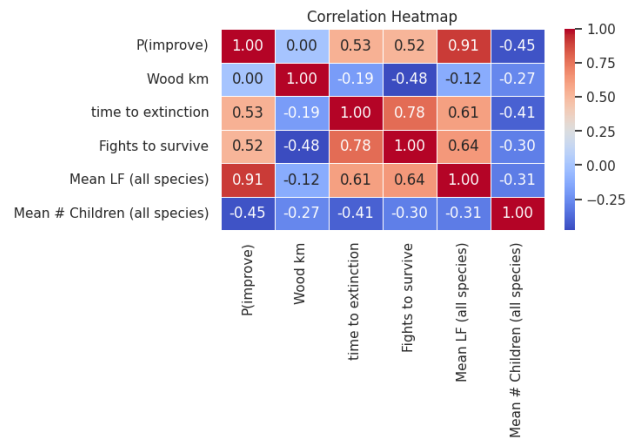
Regarding the assumption I made, some hypotheses are considered in the simulation. First of all, I considered the natural selection of a specific selection of species of individuals. Secondly, I assumed that the parent is just one, so each child has only one father. Concerning the reproduction model, I specify that the mature age of each individual for getting pregnant is not considered in the model. In addition, I considered a continuous extension of the wood in which the species reproduce: it has to be intended as along a single axis (1 dimension) and measured in kilometers. To conclude, I assumed that when a species reaches the overpopulation case, it moves to another area: if such a species wins the fight, it keeps reproducing in that area, without coming back to its original one.

## 1.5 Mobility, fight and survival models

As outlined above, the simulation model presents a quasi-reliable situation that happens in nature. When a species reaches the condition of overpopulation (i.e., the number of individuals in an area overcomes the allowed limit), chooses a section of the entire population to send to war. Such a fraction of individuals randomly decides the destination and moves according to the species’ speed. When the warriors arrive at the new area, they start fighting: the fight is conducted in a fixed timestamp (set to 7 days). At the end of the conflict, just a fraction of individuals from both sides survive. The survival model is, then, built based on the fraction of strong individuals for each species: indeed, the 5 chosen species present a fixed input parameter regarding the percentage of strong individuals. In such a way, the wolves’ and the eagles’ species have more likelihood of survival given the fact that their fraction of strong individuals is higher than the other species’. On the contrary, since the foxes, the deers and the mice have a higher fraction of weak organisms, they have fewer chances to survive. The simulation also modeled the condition for which if a species of series A encounters a species of series B, surely the first one wins: for instance, if the eagles fight against the mice, the eagles will win the war and move to the mice’s area.

## 2 Output metrics, data structures and further analyses

The output metrics used to evaluate the results of the simulation consist of the general trend of the population over time, the flow of individuals towards the areas and the focus on the fight situation: when the curve has a rapid increasing means that the individuals in that area are attacked by another species. At the same time, the original area of the attacking species notices a rapid decrease in individuals. The, depending on who wins, the curve becomes more or less stable (see the output code). Further analyses are added at the end of the simulation and consist in the general analysis of the correlation between all of the variables to evaluate which one influences the most another one. As we can interestingly evaluate from the Figure 1, the proportional or disproportional dependencies are examined in the plots in Figure 2. In conclusion, I can say that the results respect our expectations and reflect the correct behavior of a natural selection model.



**Figure 1:** Correlation heatmap

For clearer evaluations, please refer to the code provided.



Figure 2: Main output metrics