Demographic study of *Centaurea corymbosa* using Integral Projection Models

Loïc Pages

Internship carried out from 13/01/2025 to 20/06/2024

Mentors : Ophélie Ronce, Eric Imbert, François Rousset

Educational institution : Université de Montpellier

Master : IDIL - M2 Quantitative Ecology and Evolution

Internship host structure: Institut des Sciences de l'Evolution de Montpellier (ISEM)


fds

Description of my contribution

During this internship my goal was to complete the implementation of IPM started by Asma Hadjou Belaid during her Ph D thesis. For this, I had her R code as base material, as well as the demographic monitoring dataset for *C. corymbosa*. I began by exploring the selection of vital rate models, which necessitated to learn using the package spaMM to fit generalized models. I then wrote a script to automate the selection process, using the set of possible effect combinations. For the new effect using splines function I had to learn how to use them, and wrote a script for converting estimated parameters for splines in spaMM to piecewise polynomial expression. We also found errors in the dataset that we had to resolve before fitting models. For this part of the work, I read several papers on statistical modeling.

After model selection, I improved the script for calculating the Integral Projection Model, as well as correcting errors in the implementation of selected models. With my advisors, we also re-calculated the discretization equations of the integral model. Finally, with my supervisors, we defined the predicted outputs of the IPM that might be interesting to compare with MPM. We in particular calculated how to compute the same life history traits as in MPM. I finally coded the result calculations and plotted them by using the frequently used package for demographic models popbio, and IPM dedicated package such as IPMpack. With the help of my advisors, I carried out a bibliographical study in order to learn more about the biology of a monocarpic perennial plant like *C. corymbosa*, about GLMM modeling and about integral and matrix demographic models.

I finally wrote the report with the contribution of my advisors.

AI Statement

In my work, I used ChatGPT (https://openai.com/) to help writing, improving and debugging my R scripts.

I also used Deepl (https://www.deepl.com/en/write) to improve sentences for the writing of this report.

Summary

Demographic models are essential for studying population dynamics, conservation biology and eco-evolutionary processes. While Matrix Population Models (MPMs) have been traditionally used for these purposes, MPMs necessitate to assign individuals to discrete classes, even if they are structured by continuous traits, such size for instance. By modeling flexible relationship between vital rates and continuous variable (*e.g.* size, mass), Integral Projection Models (IPMs) avoid creating arbitrary class divisions and may provide refined predictions about the dynamics of structured populations. *Centaurea corymbosa* is an endemic plant species subject to a continuous demography survey since 1994. Previous studies using MPMs concluded that the species is declining and is threatened by climate warming. Here, we present a robust IPM framework including both size- and age-dependent vital rates and incorporating temporal and spatial variability, which can be used for evolutionary or demographic analyses on *C. corymbosa*. We improve the implementation of IPM. First by using a model selection strategy and fitting generalized linear mixed models for vital rates depending on a continuous variable, rosette size, a discrete variable, age and the random effects of year and population. Second, we improved the quality of IPM calculation. Third, we used the IPM to predict population dynamics of *Centaurea corymbosa* and compared its results with previous matrix model predictions, examining the stable age and size structure, the population growth rates and their variations across years. We found significant dependence on size for all vital rates, and on age for many, which show for the first time clear senescence of older plants in *C. corymbosa*. Our model predicts that the species is in decline, but less drastic than predicted by MPM. IPM also predicts less variation between years of recruitment than previous MPMs.

Key words: Structured population; Integral Projection Model; population dynamics; conservation biology, *Centaurea corymbosa*; life history trait

Contents

1 Introduction

Understanding species demography is crucial in both ecology and evolutionary biology (Metcalf and Pavard 2007). Demographic models provide a powerful framework to predict how populations numbers change over time. In evolutionary biology, such models are key to our understanding of the evolution of life-history traits (Rees and Rose 2002; Kentie et al. 2020; Childs et al. 2004), and of senescence, *i.e.* the decline of vital rates such as fecundity or survival with age (Hamilton 1966). In the current context of accelerating biodiversity loss, demographic models have also become essential tools in conservation biology (Morris and Doak 2002), helping to identify threats to natural populations and to define and implement effective management actions (Beissinger and Westphal 1998; Fieberg and Ellner 2001). For instance, evolutionary rescue models combining evolution and demography can help predicting when introducing genetic diversity into small and vulnerable populations can allow them to adapt fast enough to avoid extinction (Kelly and Phillips 2019; Erlichman 2024). Here, we wish to develop such demographic models, with the ultimate aim to assess which interventions could rescue a perennial endemic plant species, *Centaurea corymbosa*, which is threatened by climate warming in the short term (Hadjou Belaid et al. 2018).

Most populations are heterogeneous, containing individuals that make different contributions to population growth depending on some of their attributes, such as their age or size. Stage-structured population models, in which individuals are grouped into discrete classes, allow describing population dynamics while accounting for major sources of individual heterogeneity. Matrix projection models (MPMs), based on the estimation of transition probabilities between distinct stages, are thus widely used (Caswell 2001). A rich mathematical theory allows inferences of parameters of biological interest from these matrices, such as predicting the asymptotic growth rate of the population (λ), its stable stage structure, and the reproductive values of each stage (Caswell 2001). Sensitivity and elasticity analyses explore how perturbations of different vital rates affect the population dynamics (Caswell 2001). Estimating projections matrices in different environmental conditions allows studying the effect of environmental stochasticity on population dynamics due to temporal and spatial variability in vital rates (Caswell 1996). In practice, simulations are often done by randomly drawing entire matrices from some empirical distribution (*e.g.* Fréville et al. 2004; Hadjou Belaid et al. 2018). When incorporating demographic stochasticity and the random fate of individuals in finite populations, MPMs predict extinction probability in a short and long term to conduct population viability analyses (Fieberg and Ellner 2001).

However, matrix projection models have some limitations (described in Ellner and Rees 2006, but see Doak et al. 2021), which could affect the biological realism of the model, the precision of estimation of parameters, and the flexibility of their use. First, the grouping of individuals into discrete classes in matrix models can be arbitrary in some respect, treating all individuals as identical within a class, whereas vital rates vary continuously within classes with variables such as size or mass (Caswell 2001). Adding more stages to describe heterogeneity in greater detail can lead to inaccurate estimates because fewer data will be available for each stage to estimate demographic rates (Ramula, Rees, and Buckley 2009). The latter problem can be particularly acute when vital rates depend on multiple variables, such as both age and size. Indeed, while numerous studies show that demography can be size and age dependent (McGraw 1989; Picó and Retana 2008), very few studies used size- and age- dependent matrix models. In perennial plants in particular, most demographic studies consider that vital rates such as survival and fecundity depend only on plant size and ignore the effect of plant age (for discussion of this issue see Edelfeldt, Bengtsson, and Dahlgren 2019) because of the lack of long term data documenting plant age over their entire lifespan. Integrating variation of life history traits with both size and age sheds light on the evolution of size and age-dependent flowering strategies in monocarpic perennials plants, which flower only once before dying (Metcalf, Rose, and Rees 2003), and also allows describing patterns of senescence in perennial plants, which still constitutes a knowledge gap (Edelfeldt, Bengtsson, and Dahlgren 2019).

Integral Projection Models (IPMs) provide a novel approach to modeling population demography by using continuous relationships between vital rates (e.g., survival, reproduction, growth) and measured quantitative traits, both continuous (*e.g.* size, mass) and discrete (*e.g.* age). Projection matrices are then replaced by projection kernels describing the probability density function of transitions in trait space between censuses. IPMs were mainly introduced by Easterling, Ellner, and Dixon 2000 and Ellner and Rees 2006 and have become very popular. IPMs rely on the flexibility of continuous relationship between vital rates and state variables, easily implemented by regression models (Merow et al. 2014; Doak et al. 2021). This makes IPMs a powerful tool for predicting population dynamics in size- and age-structured populations (Edelfeldt, Bengtsson, and Dahlgren 2019). IPMs facilitate the implementation of environmental stochasticity, when regression models incorporate random effects such as year or population effects (Childs et al. 2004). Like MPMs, IPMs allow us to determine which life history traits are sensitive to these variations and how they in turn impact the population growth rate (Rees and Ellner 2009). Furthermore, because regression models are formulated as continuous functions, IPMs may require few parameters to be fitted, allowing vital rates to be estimated for

all values of the state variables, rather than separately estimating vital rates for many discrete stages. Therefore, it has been claimed that IPMs outperform matrix models when the dataset is limited (Ramula, Rees, and Buckley 2009, but see Doak et al. 2021).

While they are formally based on a continuous description of life history variation, in practice the numerical computation of integrals in IPMs generally relies on approximations subdividing the continuous state space in discrete narrow classes and on the manipulation of very large matrices. It makes IPMs implementation not that different from MPMs (Ellner and Rees 2006; Doak et al. 2021). Choices about the discretization of the continuous state variables (number of classes, and their range) or methods to approximate vital rates within classes and transitions rates between them can affect the accuracy of IPMs (Doak et al. 2021). Common issues include avoiding eviction (removing individuals because they fall out of the modeled range of sizes; Williams, Miller, and Ellner 2012). Furthermore, representing the variation in vital rates by simple continuous functions of state variables could in some instance be a limitation of IPMs, with weak empirical justification for the choice of these simple functional forms, where MPMs do not constrain the variation of vital rates across stages (Doak et al. 2021). The use of spline regressions can allow considering more flexible forms for the variation of vital rates with age and size than the linear or quadratic functions that are often assumed by default (Edelfeldt, Bengtsson, and Dahlgren 2019).

Adjusting flexible regression models with several state variables and random effects in IPMs thus entails defining a model selection process to choose the best model among many options to predict population dynamics. As explained in Tredennick et al. 2021, the model selection criteria changes if the study focuses on exploration, inference or prediction. Aikake Information Criterion (AIC) was originally designed to approximate the capacity of a model to predict data not used to fit the model, when such external data is not available (Tredennick et al. 2021). AIC seeks to optimize prediction accuracy by trading-off between bias and variance in estimates, penalizing model complexity. Models selected based on their AIC score however tend to be over-fitted, frequently including spurious covariates that make minor contribution to predictions (Tredennick et al. 2021). Alternative criteria to AIC includes Schwarz Information criteria (commonly named BIC, but see Delattre, Lavielle, and Poursat 2014 for discussions on how to use BIC for mixed effects models). One goal of this study is to implement the best regression models defining vital rates in the study species, *C. corymbosa*.

Centaurea corymbosa (Asteraceae) is an endemic plant from the Massif de la Clape (Narbonne, South of France). Only 6 populations are present on a 3 km^2 area. Individuals occur on rocky habitats where there is very little soil and competition is low. *C. corymbosa* is a

perennial monocarpic plant: individuals flower, on average, after 4.9 years (ranging from a one to 11 years) and then die after reproducing. The species is endangered according to the IUCN red-list, it is protected both in France and Europe. Its life-history traits (monocarpic, self-incompatibility, with a limited seed dispersal) increase its sensitivity to disturbance, as the species is a very poor colonizer. This makes the persistence of the 6 extant populations particularly critical for the species survival (Olivieri et al. 2016). The demography of *C. corymbosa* has been first studied by Fréville et al. 2004 and Hadjou Belaid et al. 2018 with matrix models, building on a long-term demographic survey started in 1994. These matrix models distinguish three stages: seedlings, rosettes and flowering plants. They predict that all *C. corymbosa* populations are declining and threatened by climate change in the short term: Hadjou Belaid et al. 2018 used stochastic matrix projection models to show that an increased frequency of warmer and dryer years increases population extinction risk for *C. corymbosa*. Under the worst scenario, they predicted that the species could be extinct in as soon as 20 years. Fréville et al. 2004 and Hadjou Belaid et al. 2018 predict that the population growth rates vary significantly over time, but less so across the 6 populations. The population growth rate is more sensitive to variation in rosette survival than variation in fecundity or seedling survival, and as rosette survival varies a lot across years, the former explains most of the between year variation in growth rates (Fréville et al. 2004). Vital rates may however vary substantially within stages due to variation in the size or age of plants.

A first attempt at developing a size and age structured IPM for *C. corymbosa* was made during the PhD of Hadjou Belaid 2018, who used the model to predict the optimal size and age at flowering in different populations and years. The goal of the present study is to build on this work and improve previous IPM for *C. corymbosa*, addressing several issues that have been discussed in the comparison of MPMs and IPMs (Doak et al. 2021). (i) First, we improved the quality of the estimates by using new generalized linear mixed models for vital rates. We allowed more flexible types of functional forms for the effect of age and size variables (polynomial and spline) on vital rates, in particular with the aim to detect possible senescence with age and complex effects of size. We used the entire data set to estimate these functions and included random effects such as year and population to allow relationship between vital rates and continuous variables to vary or not across space and time. We used a model selection strategy based on AIC. (ii) Second, we improved the quality of numerical integration by using cumulative distribution for appropriate vital rate functions, instead of using midpoint method for all functions (Doak et al. 2021). We varied the number and range of intervals used in the discretization to test whether this choice affected the precision of our predictions. (iii) Third, we

used the fitted IPM to predict population dynamics of *Centaurea corymbosa* and compared its results with previous matrix model predictions, examining the stable age and size structure, the populations asymptotic growth rate, its variation across years and population and its elasticity in response to variation in transition rates.

2 Material & Methods

2.1 Demographic data - Data collection

An individual-based demographic monitoring has been set up in 1994 with the installation of 48 permanent quadrats (with 4 additional quadrats in 2020) distributed across the 6 populations (in descending order of size: Enferret 1 and 2, Auzils, Portes, Peyral, Cruzade) of *C. corymbosa* in the La Clape massif, near Narbonne, Southern France. Following the current recommendation for protected species, precise GPS coordinates are not given. Every 3 month (September, December, March and June), presence/absence of each plant and its rosette size are recorded and new seedlings are added to the survey. Plant size is defined as the average of two measures of the diameter of the rosette to the nearest 0.5 cm. In June, during the flowering period, reproductive status is noted, and the number of capitula produced by flowering plants is counted (except between 2003 and 2011 when this measure was not taken). The age of all plants first seen as a seedling is recorded. In total, the data set contains 5249 yearly observations on 2931 distinct individuals.

For the modeling of changes in population numbers across years, and consistently with previous matrix projection models on *C. corymbosa*, we consider the status of the plants in June of each year. However, the plant size measured in March is a better predictor of probability of flowering than the plant size in June (the best model using size in June has $AIC = 1069$ while the best one using size in March has $AIC = 840$). This can be explained by drought stress at the beginning of Summer, which affects rosette size and makes its measurement in June a less reliable indicator of a plant below-ground resources. Rosette size also shrinks with flowering, which is problematic when assessing which size triggered flowering. Size in March of plants still alive in June was thus used for modeling purposes for all vital rates. The age of plants ranges from 1 year for seedlings to 13 years. However, the plants aged 8 years or older were grouped into the last age class (8^+) because of the low number of individuals that exceed this age (23 individuals).

2.2 Integral Projection Models

The Integral Projection Model is based on a kernel function $K(B, A)$, which describes all individual transitions between state A at time t to state B at time $t+1$. The general equation for IPM is given by:

$$\boldsymbol{n(B, t+1) = \int_{\Delta} K(B, A) n(A, t) dx} \tag{1}$$

with Δ the set of all possible states, and $n(A, t)$ the density of individuals in state A at time t in the population.

In this study, the state of an individual is defined by both its age (from 1 to 8^+ years) and size (from 0.5 to 15 cm for seedlings and 25 cm for rosettes). Thus, the kernel function defines all possible transitions for an individual of size x and age a at time t to an individual of size y and age a' at time $t+1$. These transitions correspond either to the growth and survival of the same individual, which then reaches age $a+1$, or to the birth of a new individual, which is then of age 1; all other transitions are null. We can split the kernel into two sub-kernels, a survival/growth kernel and a reproduction kernel, P and F, such as $K = P + F$. The non-zero terms in each sub-kernels are precisely defined as:

- $P(x, a, y, a+1)$: the probability that an individual of age a and size x at time t becomes an individual of age $a+1$ and size y at time $t+1$ (Eq. 3);
- $F(x, a, y, 1)$: the density of seedlings (age 1) of size y at time $t+1$ born to a mother plant of age a and size x at time t (Eq. 4).

We predict the density $n_a(y, t+1)$ at time $t+1$ (*i.e.* in June) of individuals of size y for each age a from the densities at time t (*i.e.* in June the previous year):

$$\begin{cases} n_1(y, t+1) = \sum_{a=1}^{8} \int_{\Omega} n_a(x, t) F(x, a, y, 1) dx \\ n_a(y, t+1) = \int_{\Omega} n_{a-1}(x, t) P(x, a-1, y, a) dx, & \text{if } a \in [2, 7] \\ n_a(y, t+1) = \int_{\Omega} [n_{a-1}(x, t) P(x, a-1, y, a) + n_a(x, t) P(x, a, y, a+1)] dx, & \text{if } a = 8^+ \end{cases} \tag{2}$$

with Ω the set of all possible sizes.

We now express the kernels as a function of underlying vital rates. We wish to estimate how these vital rates change with size and age of the plant, but also across populations and years. However, for the sake of readability of the equations, we here omit in our notation the dependency on year and population. Because the species is monocarpic and most plants die after flowering, individuals that are observed to be alive in June the following year have not

flowered in previous June. Similarly, growth can be observed only for individuals that have survived from June at time t to June at time $t+1$. The sub-kernels are thus given by:

$$P(x, a, y, a+1) = [1 - f(x, a)]s(x, a)g(x, y, a) \tag{3}$$

$$F(x, a, y, 1) = f(x, a)c(x, a)\ \xi\ \omega(y) \tag{4}$$

with $f(x, a)$ the probability of flowering for an individual of size x and age a in a given population and year; $s(x, a)$ the probability of survival for a non-flowering individual of size x and age a in a given population and year; $g(x, y, a)$ the probability density that, in a given population and year, a surviving and non-flowering individual of size x and age a grows to size y the next year; $c(x, a)$ the number of capitula produced by a flowering individual of size x and age a in a given population and year; ξ the establishment rate in a given population and year, measuring the density of surviving seedlings in June at time $t+1$ per capitulum produced in the population at time t; $\omega(y)$ the probability density of seedling size y in a given population and year, for those seedlings that are still alive in June; seedling size is measured in March.

2.2.1 Estimating vital rate functions

2.2.1.1 Model selection

The vital rate functions to estimate from the data are the survival probability, flowering probability, plant growth distribution, fecundity, establishment rate, and seedling size distribution.

To identify the functions that better fit our data, we fitted alternative models combining different fixed and random effects with generalized linear mixed models (GLMM) using the package spaMM - version 4.5.35 (Rousset and Ferdy 2014) in R - version 4.4.2 (R Core Team 2023).

For each vital rate function, we determined the distribution in which observations are assumed to be drawn, depending on the nature of the response variable. We modeled the fixed effect of size and age on vital rates using either polynomial functions up to degree 4 or spline functions. Splines are piecewise polynomial functions and are defined by the degree of the polynomials and the coordinate of one or more knots (which are the coordinates at which the polynomial form changes). Population, year and individual (for traits measured several times during the life of an individual) were modeled as random effects. Because variance between years or between populations in vital rates may vary with the age or size of an individual, we included random variation in the intercept of the regression model but also in the regression slope of the linear effect of age and size. Table 3 in Supplementary Section 5 lists all alternative

fixed and random effect forms considered. We combined these alternatives to generate models and evaluate which combinations best predict variation in the vital rate functions.

Models were compared using the Akaike Information Criteria (AIC): the model with the lowest AIC was selected to be included in the IPM. As a test of robustness with regard to our choice of model selection criterion, we also computed Bayesian Information Criteria (BIC) for each model. We compared IPM predictions based on vital rate models selected by AIC versus those selected by BIC

2.2.1.2 Vital rate functions

Flowering probability. A binomial model with a logit link function was fitted. Because of convergence issues, we had to remove splines and individual random effects from possible models for this trait.

Survival probability for non flowering individuals. Due to their ecological differences, seedling (1-year-old) and rosette (>1 year) survival were modeled separately using binomial logit models, excluding individuals that flowered at time t. Individual random effects were included only for rosettes, as seedling survival was measured once per individual.

Growth. The growth function describes the distribution of possible new sizes y at $t+1$ of an individual of age a and with current size x, conditional on its survival. The distribution of future size was assumed to be Gaussian with a mean value and a variance that both depend on current size and age. We therefore fitted a regression model for future size of every surviving individual with a residual variance that also depended on its age and size. The log-transformed residual standard error was assumed to vary with the log-transformed age and size of the growing plant.

Number of capitula. The log-transformed number of capitula per flowering plant was modeled as a Gaussian distribution, fitted on a sub-sample of years in the survey since data were unavailable from 2003 to 2011.

Establishment rate. We neglect seed dormancy, consistently with previous modeling in *C. corymbosa* (Fréville et al. 2004). The establishment rate is calculated for each quadrat in the survey as: (number of seedlings at $t+1$)/(number of capitula at t). Quadrats with no observed flowering plant at time t were excluded. In years where capitula were not counted on flowering plants (2003-2011), we used observed plant size of each flowering plant to predict its number of capitula with our fitted best model for capitula number, setting the random effects of year to 0, and used this prediction to compute the establishment rate during these years. As the establishment rate is defined at the scale of the quadrat and not of individual plants, the compared models only included random effects of population, year and their interaction.

Seedling size. In the quadrats, seedlings cannot be assigned to a specific mother-plant, thus we assumed that the size of a seedling is independent of the size or age of its mother. After examining the empirical distribution of seedling size across the entire data set, we modeled it with a Gamma distribution, with random effects of year, population, and their interaction.

2.2.2 Implementation of the Integral Projection Model

The integrals in Eq. 1 cannot be analytically solved, so we discretized the transition kernel to numerically evaluate the integrals using approximations. We combine the use of the mid-point rule and of cumulative distribution functions to avoid eviction issues and improve the precision of our approximation.

For each age class a, we divided the continuous size range $[x_{min,a}, x_{max,a}]$ into k sub-intervals of width $h_a = (x_{max,a} - x_{min,a})/k$. By default, we used $k = 50$ size intervals, but also varied that number to test the precision of our approximation. We set different possible size ranges for seedlings and rosettes: for all ages, $x_{min,a} = 0.5$, which is the precision of our measures; but $x_{max,1} = 15$ for seedlings ($a = 1$) and $x_{max,a} = 25$ for rosettes ($a > 1$), which exceed the maximal observed plant size in the data set. Thus, for each interval i, corresponding to $[x_{i,a}^{inf}, x_{i,a}^{sup}]$, where $x_{i,a}^{sup} = 0.5 + ih_a$ and $x_{i,a}^{inf} = 0.5 + (i-1)h_a$, we define m_i the midpoint size of the i^{th} interval by: $m_i = (x_{i,a}^{sup} - x_{i,a}^{inf})/2$.

The number of individuals of age a and in the size interval i at t is then given by:

$$N_a(i,t) = \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} n_a(x,t)dx \simeq h_a n_a(m_i,t)$$

For $a \in [2,7]$, we for instance have:

$$N_a(i,t+1) = \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \int_{\Omega} P(x,a-1,y,a)n_{a-1}(x,t)dx\ dy \simeq \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \sum_j P(m_j,a-1,y,a)N_{a-1}(j,t)dy$$

Using the same type of approximation (see S.M. 5), we obtain the future number of individuals of each age and in each size interval by using a projection matrix $\tilde{\mathbf{K}}$, such that: $\mathbf{N}(t+1) = \tilde{\mathbf{K}}\mathbf{N}(t) = (\tilde{\mathbf{F}} + \tilde{\mathbf{P}})\mathbf{N}(t)$. The matrix $\tilde{\mathbf{K}}$ is a 8 by 8 super Leslie matrix, with size-structured matrices on the first block of lines and below the diagonal. More precisely, we have:

$$\begin{cases} N_1(i,t+1) = \sum_{a=1}^{8} \sum_j \tilde{F}(j,a,i,1)N_a(j,t) \\ N_a(i,t+1) = \sum_j \tilde{P}(j,a-1,i,a)N_{a-1}(j,t), & \text{if } a \in [2,7] \\ N_a(i,t+1) = \sum_j [\tilde{P}(j,a-1,i,a)N_{a-1} + \tilde{P}(j,a,i,a)N_a], & \text{if } a = 8^+ \end{cases} \quad (5)$$

with

$$\tilde{P}(j, a-1, i, a) = s(m_j, a-1)[G(x_{i,a}^{sup}, m_j, a-1) - G(x_{i,a}^{inf}, m_j, a-1)] \quad (6)$$

$$\tilde{F}(j, a, i, 1) = c(m_j, a) f(m_j, a) \xi \, [\mathcal{W}(x_{i,a}^{sup}) - \mathcal{W}(x_{i,a}^{inf})] \quad (7)$$

where $G(x_{i,a}^{sup}, m_j, a-1) = \int_{-\infty}^{x_{i,a}^{sup}} g(y, m_j, a-1)dy$ is the cumulative distribution function (CDF) of size after growth and $\mathcal{W}(x_{i,a}^{sup}) = \int_{-\infty}^{x_{i,a}^{sup}} \omega(y)dy$ is the CDF of seedling size. All other elements of $\tilde{\mathbf{P}}$ and $\tilde{\mathbf{F}}$ are null. The eviction problem (individuals that are out of the range size), is handled in the predicted size at $t+1$ by taking for the calculation of the differences of the CDF: $x_{i,a}^{inf} = -\infty$ instead of 0.5 for the first interval, and $x_{i,a}^{sup} = +\infty$ instead of 15 or 25 for the last interval.

2.3 Analyses

2.3.1 Asymptotic analyses

After discretization, the Integral Projection Model has the same form as a matrix population model and quantities of interest can be computed directly from the matrix by using the package popbio - version 2.8 (Stubben and Milligan 2007). We manipulated different matrices. First, we use our model to compute a matrix for each pair of population and year (6×28 kernels). Second, we computed an average matrix per year (28 kernels); this was done by averaging all predicted age- and size-specific vital rates for that year across all 6 populations and assembling a matrix with these vital rates as in Eq. 3. Finally, we obtained a single matrix by averaging all predicted age- and size-specific vital rates across years and populations.

We computed the asymptotic growth rate and the stable size and age distribution, respectively as the largest eigenvalue λ and associated right eigenvector $\mathbf{W}$ of this average matrix. We compared, for each age, the observed sizes distribution across the entire data set to the predicted stable size distribution. By pooling all sizes in each age, we calculated the stable age distribution, which was also compared to the empirical age distribution. The distribution of size at flowering is given by the probability for a flowering plant to have size in interval i:

$$Z(i) = \frac{\sum_a f(m_i, a) W_a(i)}{\sum_a \sum_j f(m_j, a) W_a(j)}$$

with and f the flowering function, and $W_a(i)$ gives the proportion of the population in size interval i and with age a. The mean size at flowering is thus given by: $\overline{x_f} = \sum_i m_i Z(i)$ Similarly, we computed the distribution of age at flowering and the mean age at flowering.

Finally, we performed an elasticity analysis to assess which transitions are the most important for variations of the asymptotic population growth rate. Elasticities measure the proportional change in the asymptotic growth rate λ caused by a proportional change in one element in the matrix $\tilde{\mathbf{K}}$ and can be computed as:

$$\mathbf{E} = \frac{\tilde{\mathbf{K}}}{\lambda} * \frac{\mathbf{V}\mathbf{W}^T}{\mathbf{V}^T\mathbf{W}}$$

with λ the asymptotic growth rate, $\mathbf{V}$ the vector of reproductive values (dominant left eigenvector),$\mathbf{W}$ the vector of stable size and age distribution (dominant right eigenvector) and we note ^{T} the transpose and $*$ the element-by-element product (Griffith 2017).

For each pair of population and year, we also computed the asymptotic growth rate from the matrix estimated for that population and year, to examine how it varies in space and time. To describe the effect of environmental stochasticity on the dynamics of the different populations, we calculated the geometric mean of these asymptotic growth rates across all years in each population. Finally we computed the asymptotic growth rate for each of the 28 years using the average matrix across populations.

2.3.2 Comparison with MPM

We checked the consistency of predictions from our newly developed Integral Projection Models with those from previous matrix models in *C. corymbosa* and examined how they differ. As described in Fréville et al. 2004, MPMs have been defined with three stages: seedlings (individuals $<$ 1 year old), rosettes (vegetative plants older than 1 year), and flowering plants. For each population and each year, the projection matrix representing the life-cycle is given by:

$$A = \begin{pmatrix} 0 & 0 & fs_0 \\ s_1(1-\alpha_1) & s_2(1-\alpha_2) & 0 \\ s_1\alpha_1 & s_2\alpha_2 & 0 \end{pmatrix}$$

with s_1 the survival rate of seedlings to the rosette stage, α_1 the probability that a seedling will flower the next year conditional on its survival,s_2 the survival rate of rosettes, α_2 the probability that a rosette will flower conditional on its survival, fs_0 the number of new seedlings per flowering plant. We calculated the same life history traits from our projection IPM kernel, averaging over the age and size distribution in each stage (the calculations are given by Eq. 8, 9, 10, 11 in Section 5 Supplementary Materials).

3 Results

3.1 Selected models for vital rates

We present the model with the lowest AIC for each vital rate. Detailed equations are summarized in Table 2 in S.M.5.

The survival probability for seedlings increases quickly with their size, especially when seedlings are still small. The variation between years has a strong effect on the survival of small seedlings, but these yearly variations are buffered for larger seedlings (Fig. 1a). Seedling survival does not vary across populations (Table 2).

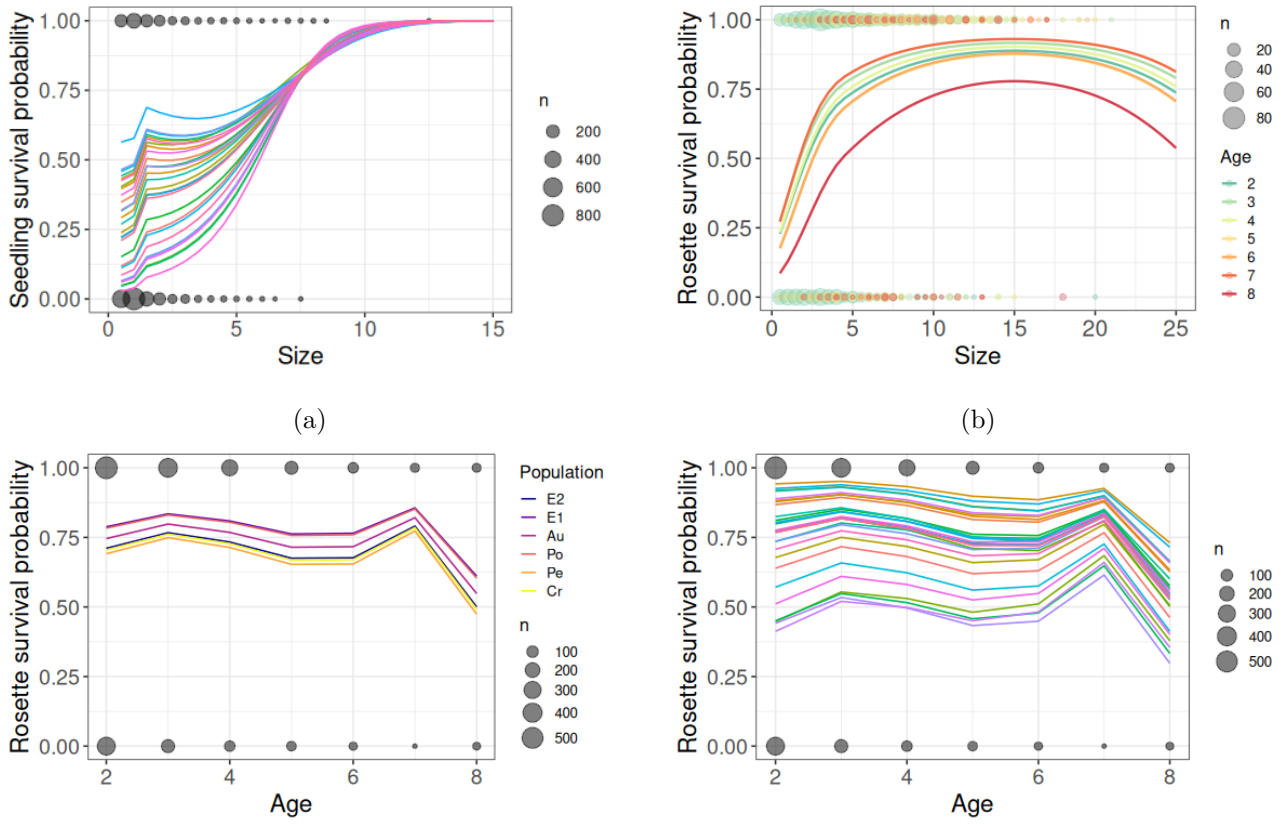
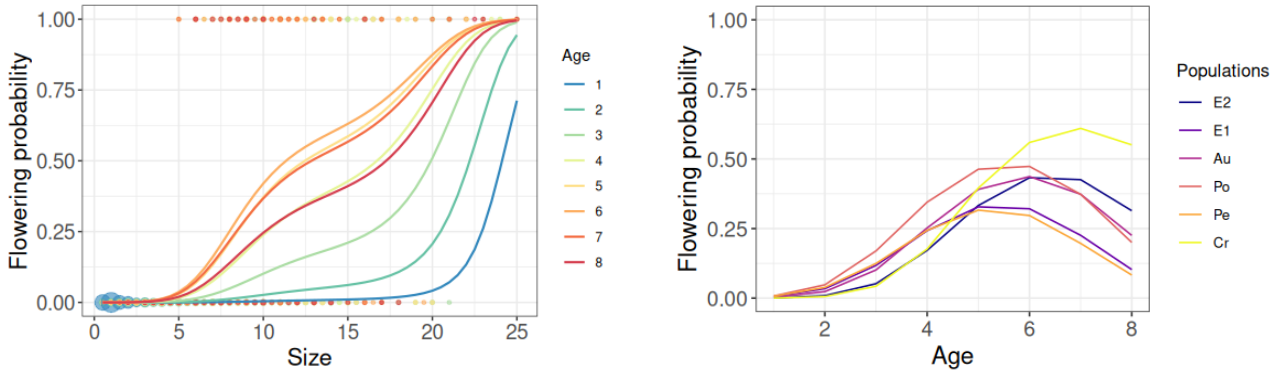


Fig. 1: (a) Survival probability for seedling (age 1) as a function of plant size. Each line corresponds to a year. The dots are the observed data. (b,c,d) Survival probability of rosette: (b) as a function of size, (c,d) as a function of age (size = 5); (b) each line corresponds to an age class; (c) each line correspond to a population, which are listed in descending order of size; (d) each line correspond to a year. The dots correspond to the observed data.

Rosette survival depends both on the plant size and age. The model with the lowest AIC predicts that survival increases quickly with plant size, plateaus and eventually declines for

greater sizes (Fig. 1b). The rarity of plants reaching these large sizes suggests caution about the robustness of this prediction. For a given size, rosette survival tends to decrease slowly with age until plants reach the age of 7 years, where survival increases before dropping at age 8 and older (Fig. 1b,c,d). The survival of young rosettes varies widely across years, but these fluctuations are buffered for older rosettes (Fig. 1d and Table 2 in S.M.5). In addition, rosette survival varies slightly across populations, with a better survival whatever the rosette age or size in the populations Portes and Enferet 1 than in Enferet 2, Cruzade and Peyral. When averaging across size and age, the average survival probability for rosettes ($\overline{s_2} = 0.65$) is higher than for seedlings ($\overline{s_1} = 0.32$).

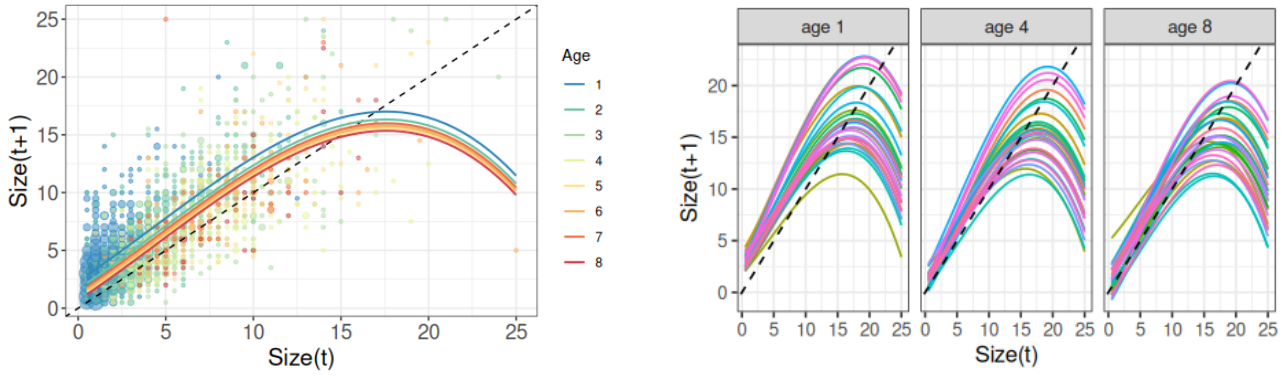


(a) (b)

Fig. 2: Flowering probability as a function of (a) size (mean over all populations), (b) age (size = 10), the populations are listed in descending order of size. The dots are observed data.

Flowering probability increases fast with increasing plant size, the size above which flowering is very likely also depends on plant age (see Fig. 2a). Consistent with empirical observations, seedlings almost never reach sizes that would allow them to flower. The probability of flowering at a given size first increases with age, but decreases later in life (Fig. 2). The distribution of age at flowering varies across populations (Fig. 2b). The age maximizing flowering probability for a given size varies from 5 years to 7 years.

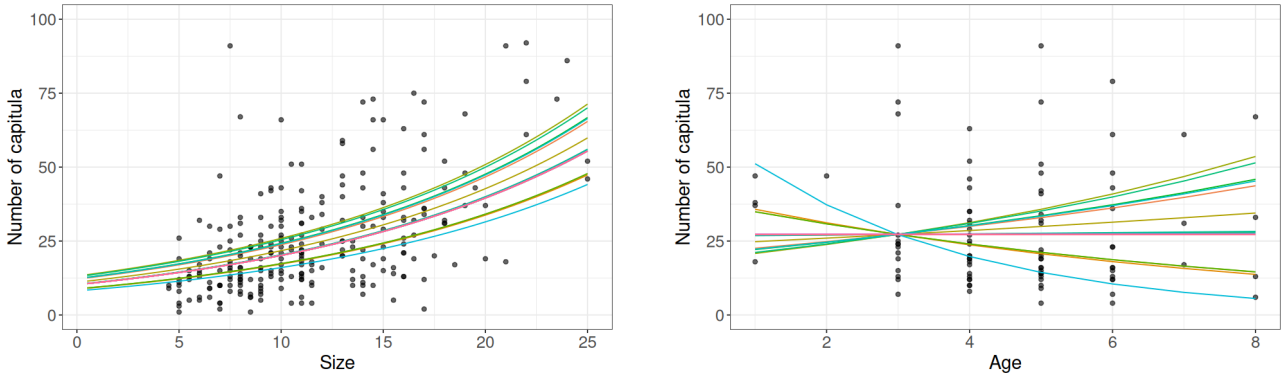
Growth of small plants is fast but slow down as the plant increases in size (Fig.3a,b). Plants larger than 15 cm seem more often to shrink than increase in size (future size is on average smaller than current size in Figures 3). Growth is also faster for young plants than older plants. Plant growth varies a lot between years, with higher variance for larger and younger plants (Fig. 8b in S.M.5). Similarly, the residual variance in growth increases with size, but decreases with age. We detected no permanent individual effect on growth within populations, and minor variation between populations, essentially due to faster growth in Cruzade (Fig.8a



(a) (b)

Fig. 3: (a),(b) Size at year $t+1$ as a function of the size at year t. The dashed line is $\text{Size}(t+1) = \text{Size}(t)$; (a) Mean over all populations and years. The dots represent observed data; (b) Mean over all populations, each line corresponds to a year.

in S.M.5), other populations being very similar.

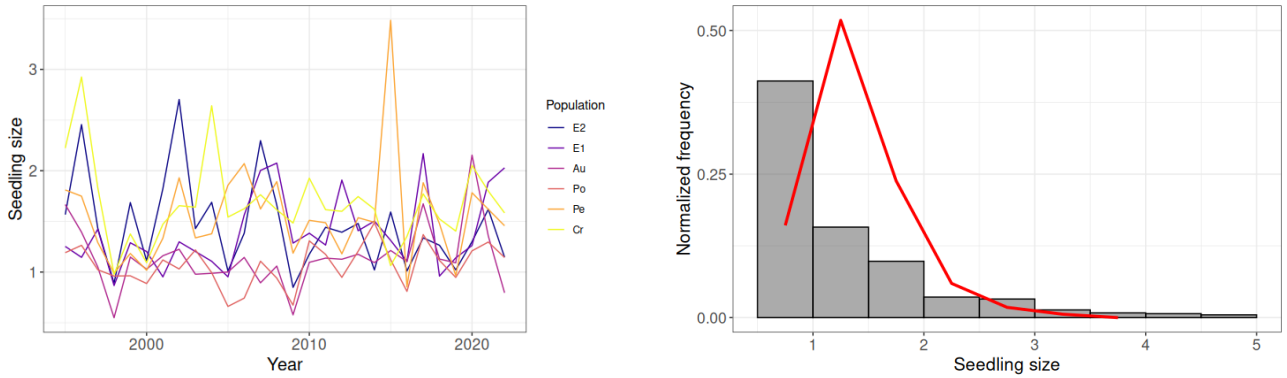


(a) (b)

Fig. 4: Number of capitula produced by a flowering plant (a) as a function of size, (b) as a function of age. The dots correspond to observed data and each line represents a year.

The number of capitula produced by a flowering individual increases exponentially with size (Fig. 4). The sensitivity of fecundity to variation of environmental conditions across years depends on the age of the plant: the fecundity of relatively young flowering plants (around 3 years) is constant across years while the fecundity of older plants is more variable.

Seedling size varies widely across populations and years, the ranking of populations changing across years (Fig. 5a). The estimated mean seedling size is 1.38 ($\pm$0.4) cm, which corresponds to the empirical mean seedling size. However, seedlings of size 0.5 are underestimated, maybe due to the log-Gamma distribution (Fig. 5b).



(a) (b)

Fig. 5: (a) Variations of mean seedling size according to year and population (in descending order of size). (b) Normalized histogram of seedling size distribution: observed in grey, estimated in red.

3.2 Demographic analyses

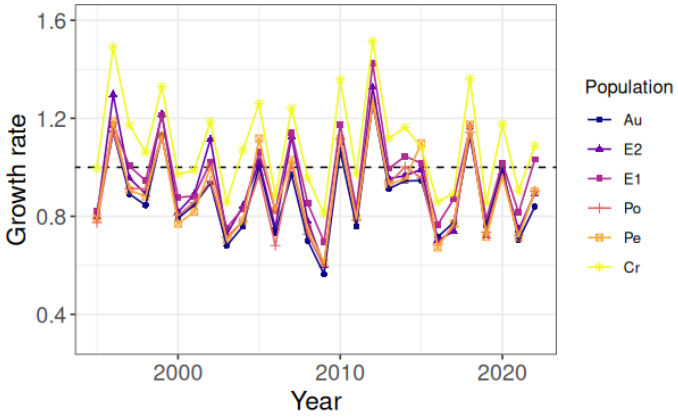


Fig. 6: Asymptotic growth rate variation across years for all populations (listed in descending order of size). Dashed line: $\lambda = 1$.

Tab. 1: Geometric mean of the asymptotic growth rates across years for the 6 populations of *C. corymbosa*

Population	λ
Enferet 2	0.892
Enferet 1	0.951
Auzils	0.850
Portes	0.889
Peyral	0.867
Cruzade	1.058

First, the asymptotic growth rate for each population and year is frequently below 1 (Fig. 6), suggesting common decline. Using the average kernel, the asymptotic growth rate $\lambda_a = 0.969$ is below one. Variation of the asymptotic growth rate across years further depresses their geometric mean, which is lower than 1 in all populations, except for the Cruzade population (Table 1). Averaging the kernel over populations, the geometric mean growth rate is $\lambda_s = 0.928$. We thus confirm the ongoing decline of the species over the past few decades ($\lambda < 1$).

We predict the stable size distribution for each age class from the average kernel: our IPM captures the initial skewed seedling size distribution, but for later ages it predicts a Gaussian-like form, unlike the persistently skewed empirical distribution with many small and few very

large plants (Fig. 9a). Our IPM does not capture well the shape of the size distribution in older plants and seem to overestimate plant size, but it predicts well the age distribution (Fig. 9b).

The model captures relatively well the distribution of size at flowering, even though it fails to predict the empirical fat tail of plants flowering at very large sizes (Fig. 9c). The predicted mean size at flowering (12.63 cm) and mean age at flowering (5.35 years old) are close to the observed ones (11.69 (±4.56) cm and 4.93 (±1.57) years old).

The elasticity matrix shows the disturbances in transitions between age and size classes that have the most impact on the asymptotic population growth rate. From Fig. 10 (in S.M. 5), we see that the elasticity to disturbance in growth and survival of rosettes declines with increasing age of the rosettes. Disturbance in the fecundity kernel has most impact on plants of intermediate age (ages 4, 5 and 6), which corresponds to ages around the mean age at flowering.

3.3 Robustness of predictions

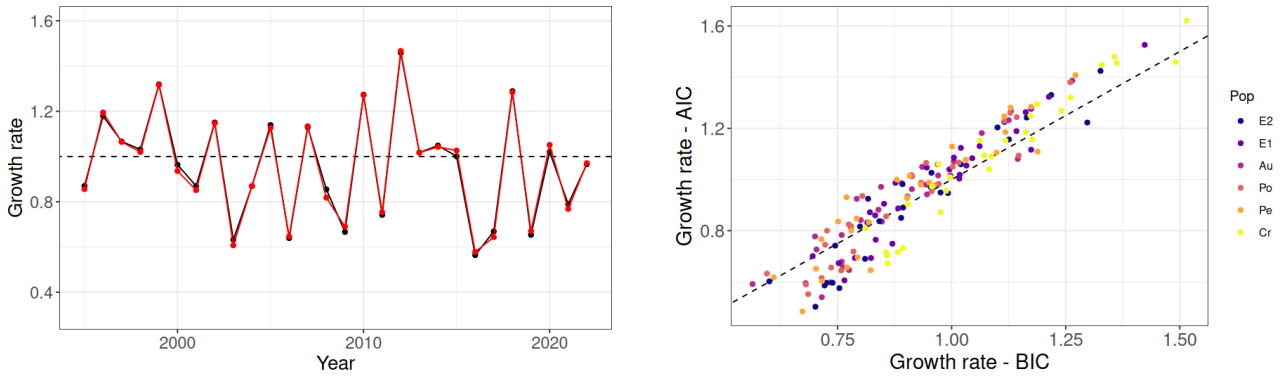


Fig. 7: (a) Variation of predicted asymptotic growth rate for IPM with AIC selected models (red) and BIC selected models (black) (Dashed line: $\lambda = 1$). (b) Predicted asymptotic growth rate from AIC and BIC for different populations and years (Dashed line: $x = y$), the populations are listed in descending order of size.

We varied the number of intervals when discretizing the kernel (Fig. 13 in S.M.5) and found that the predicted asymptotic growth rate was already stable for 20 to 30 intervals. Using 50 intervals, the continuous kernel approximation is considered sufficiently precise.

Predictions of asymptotic growth rates using the AIC and those using the BIC as a model selection criterion are globally similar (Fig. 7), even though the selected models for vital rates are different. The good and bad years are the same, but predictions from BIC and AIC selected models diverge more when predicting the dynamics at the level of populations.

3.4 Comparison with Matrix Population Models

IPM and MPM have very similar predictions about the variation of seedling survival across years (see Fig. 12a in S.M.5). Both also predict that the probability for a seedling to survive and flower is very low ($< 10^{-5}$). Predictions about rosette survival are also similar between IPM and MPM with the same variations across years (Fig. 12b in S.M.5). However, predicted rosette flowering probability differs between IPM and MPM. Although IPM captures similar yearly variations globally, it predicts on average a lower rate of flowering than MPM. Conversely, IPM estimates on average more new seedlings per flowering plant than MPM and little variation across years contrarily to MPM where estimates of recruitment shows very large peaks in 2016 and 2018.

The asymptotic growth rate for each year obtained with the IPM is greater that those obtained with MPM (Fig. 11 in S.M. 5). However, the effect of years seems to be similar between the two models, and they capture approximately the same good and bad years for the species dynamic. IPM predictions however are more optimistic than matrix result with a predicted asymptotic growth rate for MPM (0.87) versus (0.969) for IPM.

4 Discussion

4.1 Life history traits variation

4.1.1 Size and age dependencies

A major question for modeling monocarpic perennial species demography is how life history traits vary with size and age (Caswell 2001; Metcalf, Rose, and Rees 2003). Age and size are expected to covary and models considering the distinct effect of age in addition to the effect of variation is size remain rare because they require long-term studies with information on the age of many individuals to be fitted. Our IPM based on 28 years of demographic survey sheds new light on this question as previous work with MPM on *C. corymbosa*, neither considered the variation of plant size and age within stages, nor modeled their separate effects on life history traits. We found that all vital rates depend on size, but many of them also depend on age, which has a distinct effect on life history variation. This is reminiscent of findings obtained by Edelfeldt, Bengtsson, and Dahlgren 2019 who found, also using a 28-years demographic survey in a perennial plant, that some vital rates depended only on size, some only on age and some showed dependency on both.

Variation in size remains the main driver of life history variation in *C. corymbosa*. Seedling survival, flowering probability and fecundity increase with size, which is consistent with many past studies on perennial plants (Kuss et al. 2008; Ellner and Rees 2006; Easterling, Ellner, and Dixon 2000). However, rosette survival has a more complex pattern of variation with size. It increases with size and reach a plateau around 10 to 15 cm, then decreases. This could be interpreted as a cost to maintain a large vegetative rosette. However, the uncertainty about the estimation of survival for very large plants is high because of the low number of observations in that size range used to fit the regression model. Models with AIC close to the best model vary in their predictions about the decline in rosette survival for supra-optimal sizes (results not shown). This uncertainty could lead to badly estimating survival probability if we extrapolate beyond the observed size range. We also found that plant growth slowed down with increasing size, with frequent shrinkage for very large plants, consistently with other studies (Salguero-Gómez and Casper 2010).

We found that rosette survival, flowering and growth are also dependent on the age of plants, beyond the effect of plant size. We in particular found evidence for senescence on several vital rates in this monocarpic species, but with distinct patterns for different life history traits. For the same size, plants grow faster when they are younger. Rosette survival declines slowly with age, until 7 years old for which we observed a better survival rate, which is difficult to understand biologically and could be due to overfitting. Survival however clearly drops after 8 years. Similarly, flowering probability also declines around age 7, while it increases with age for a given size before that age. Even though fecundity is mostly explained by plant size, with no main effect of age, the sensitivity of reproduction to yearly variation in environmental conditions increases in old age classes, suggesting less capacity to buffer the effect of environmental stress. Such general decline in performance in older plants that have not reproduced yet is clearly maladaptive. Note that, because of the monocarpic life cycle of *C. corymbosa*, old plants necessarily belong to a category of plants that have failed to grow fast to a sufficient size to flower. Their poor overall performance may thus reflect some other source of heterogeneity in the population affecting vital rates. Yet, our model selection failed to retain permanent individual effects for any vital rates, so we lack evidence supporting this hypothesis.

4.1.2 Spatial and time variations

Previous work on *C. corymbosa* using MPMs has concluded that variation among populations for life-history traits and asymptotic growth rate are low, and unrelated to population size or genetic diversity (Fréville et al. 2004; Hadjou Belaid et al. 2018). Our IPM largely confirm

these conclusions as variation between populations is absent for seedling survival, fecundity, establishment rate and minor for rosette survival. By analyzing size-variation we however uncover some variation among populations that had not been described by MPM: in particular, seedling size, one of the best-estimated parameters, fluctuates greatly between populations; for example the Auzils and Portes populations have, in average, smaller seedlings than others. But this population effect is not consistent over time for the other populations. Flowering probability at a given size and age varies across populations, which changes the age maximizing flowering probability. Whether this divergence of flowering strategies across populations correspond to a local adaptation should be investigated in the future, by comparing it to the age at flowering that maximizes the population growth rate in each population (Hadjou Belaid 2018).

Consistently with previous predictions of MPMs in *C. corymbosa*, we found large differences among years in life history traits and asymptotic growth rates, which is likely due to variation in climatic conditions (Fréville et al. 2004; Hadjou Belaid et al. 2018). Surprisingly, the establishment rate did not vary across years while one would expect that both the number of viable seeds per capitulum and the early survival of very young seedling from germination in the Fall to June would be very sensitive to climatic conditions. Because the establishment rate is estimated at the quadrat level and not at the individual level, we may lack power to detect year effects beyond the stochastic variation of our estimate across quadrats due to low number of quadrats with flowering plants and low number of flowering plants per quadrat. A second explanation could be that the establishment rate is a composite parameter with seed fertilization and viability depending on climatic conditions during Summer while germination and early survival depend on environmental conditions in the following Fall, Winter and Spring (Hadjou Belaid et al. 2018). Negative covariation between seed germination and later survival could dampen the fluctuations of the establishment rate across years. In particular, it would be interesting to examine the effect of density-dependent mortality at this early stage. Mean seedling size, seedling survival, rosette survival, plant growth and production of capitula for a given size all showed significant temporal variations across years, consistently with the impact of climatic conditions on the species (Hadjou Belaid et al. 2018). Hadjou Belaid et al. 2018 in particular found that Summer temperature affected all life history traits (fecundity, survival of just emerged seedlings, seedling and rosette survival and probability of flowering), while the number of wet days either in the Fall and Winter or in Summer strongly affected the survival of very young plants (just emerged to on-year old seedlings). Interestingly, our analysis shows that once size variation has been taken into account, there is no variation across years of flowering probability, while MPM models that did not consider size variation within the rosette stage

found that the proportion of flowering rosettes fluctuated across years. These fluctuations may then entirely be driven by variation across years in the size of plants.

4.2 Demography

As previously shown by Fréville et al. 2004; Hadjou Belaid 2018, our IPM predicts that *C. corymbosa* populations are declining. However, this decline seems to be less drastic than predicted by MPMs. Our model predicts similar plant survival rates than MPM, but a lower flowering rate and a greater recruitment rate, the latter being the most likely explanation for greater asymptotic growth rates predicted by the IPM. In one hand, the higher fecundity predicted by the IPM may be in part inflated by the overestimation of plant size with increasing rosette age in our predictions that we have documented. This however does not seem to affect much the prediction of rosette survival in the IPM, which agrees well with estimates in the MPM. For larger plants, rosette survival may be less sensitive to size overestimation than fecundity is. On the other hand, by using data on capitula numbers, the IPM may better estimate fecundity and suffer less from the strong stochasticity and frequent missing data that affect the estimate of recruitment rate at the quadrat level used in MPMs.

Temporal fluctuations in population growth rate depress the stochastic geometric growth rate of the populations, aggravating their decline. Consistently with previous studies (Fréville et al. 2004; Hadjou Belaid et al. 2018), we also predict temporal fluctuations in the populations growth rate with the IPM, but less intense than with MPM. Environmental stochasticity is hard to separate from demographic stochasticity and measurement error in these past studies, because projection matrices are estimated using the data for a single year and population, with a limited number of quadrats and individuals, without a model selection strategy. Conversely, the regression models used in our IPM were fitted on the entire data set and year random effects estimated with a specific error model, so our estimation of environmental stochasticity should be more precise than in MPMs.

In MPMs, disturbance of rosette survival has the highest and flowering probability the lowest impact on population dynamics of *C. corymbosa*, as measured by demographic elasticities (Hadjou Belaid et al. 2018). Survival of large plants is indeed the most important life history traits for a monocarpic plant (Cotto et al. 2017). If we pool all rosette ages (from 2 to 8), our elasticity analysis using the IPM confirms the importance of rosette survival and growth compared to fecundity and seedling survival. However, by distinguishing rosettes of different size and age, we can deepen our understanding of critical transitions in the life cycle of the species. In particular, survival and growth of younger plants (seedlings and young rosettes)

have a higher contribution to the population dynamics than that of older plants, consistently with senescence evolution models predicting declining elasticity with age (Hamilton 1966).

4.3 Model improvements and limits

Using AIC, we selected more complex functions for vital rates than those in the previous IPM applied on *Centaurea corymbosa*, which could improve predictions by relaxing constraints on the variation of vital rates. However, biological interpretation might be difficult with these complex relationships and AIC tends to overfit the data (Tredennick et al. 2021). As a test of robustness, we used an alternative model selection criterion: BIC selects simpler models, particularly with fewer spline functions (results not shown). However, the computation of BIC does not properly account for random effects (but see Delattre, Lavielle, and Poursat 2014 for discussions on how to use BIC for mixed effects models). Despite these differences, the model selection procedure does not impact deeply the estimation of population growth rates. As with all model selection processes, we made arbitrary choices regarding the explanatory variables to limit the possibilities (*e.g.* for splines) to avoid having too many models to fit. Furthermore, we chose not to consider interaction between fixed effects (size x age), though some studies suggest that this type of interaction can be significant (Baden et al. 2021).

To have a better calculation of the discretization of the IPM kernel, we used Cumulative Distribution Function (CDF) method for appropriate vital rates such as the growth function and seedling size distribution(Doak et al. 2021), but we still rely on approximation based on midpoint method for integrals involving fecundity, flowering and survival. Also, we tested different number of intervals for discretization, and we confirm Doak et al. 2021 results who showed that 20 to 30 classes are sufficient for asymptotic IPM to converge.

4.4 Conclusion and perspectives

This study on *Centaurea corymbosa* using a novel approach, Integral Projection Models, confirms previous predictions about the demography of this endangered species, but also sheds a different light on these dynamics, by examining variation between vital rates and both size and age. Validation of these new predictions with data not used to fit the models is now necessary to evaluate if these new models predict better the persistence prospects of the species. Data on the annual count of flowering plants (outside the quadrats) provide such an opportunity.

In the future, this model could be used for other types of analyses, such as predicting the optimal flowering size strategy in each population and how it may change with climate warming (Hadjou Belaid 2018), or evaluating the critical level of genetic variance that could rescue the

populations from extinction (Erlichman 2024), which could facilitate the design of effective conservation actions. In order to improve our understanding of the sensitivity of vital rates to global change, it is necessary to incorporate climatic parameters, such as temperature and precipitation as covariates in our functions for vital rates (Hadjou Belaid et al. 2018). Analyzing the sensitivity and elasticity of the population growth rate to disturbance in parameters of the vital rate functions rather than in kernel transition rates should also improve our understanding of the exact life history traits that are critical for the species persistence (Griffith 2017).

References

Baden, H. Maria et al. (2021). "The effects of age on the demography of a perennial plant depend on interactions with size and environment". *Journal of Ecology* 109.2, pp. 1068–1077.

Beissinger, Steven R. and M. Ian Westphal (1998). "On the Use of Demographic Models of Population Viability in Endangered Species Management". *The Journal of Wildlife Management* 62.3, pp. 821–841.

Caswell, Hal (July 1, 1996). "Analysis of life table response experiments II. Alternative parameterizations for size- and stage-structured models". *Ecological Modelling* 88.1, pp. 73–82.

— (2001). "Matrix population models: Construction, analysis, and interpretation". Sunderland, Mass. : Sinauer Associates.

Childs, Dylan Z. et al. (Feb. 22, 2004). "Evolution of size–dependent flowering in a variable environment: construction and analysis of a stochastic integral projection model". *Proceedings of the Royal Society of London. Series B: Biological Sciences* 271.1537, pp. 425–434.

Cotto, Olivier et al. (May 5, 2017). "A dynamic eco-evolutionary model predicts slow response of alpine plants to climate warming". *Nature Communications* 8.1, p. 15399.

Delattre, Maud, Marc Lavielle, and Marie-Anne Poursat (Jan. 2014). "A note on BIC in mixed-effects models". *Electronic Journal of Statistics* 8.1, pp. 456–475.

Doak, Daniel F. et al. (May 2021). "A critical comparison of integral projection and matrix projection models for demographic analysis". *Ecological Monographs* 91.2, e01447.

Easterling, Michael R., Stephen P. Ellner, and Philip M. Dixon (2000). "Size-Specific Sensitivity: Applying a New Structured Population Model". *Ecology* 81.3, pp. 694–708.

Edelfeldt, Stina, Karin Bengtsson, and Johan P. Dahlgren (2019). "Demographic senescence and effects on population dynamics of a perennial plant". *Ecology* 100.8, e02742.

Ellner, Stephen P. and Mark Rees (Mar. 2006). "Integral Projection Models for Species with Complex Demography". *The American Naturalist* 167.3, pp. 410–428.

Erlichman, Adèle (Apr. 22, 2024). "Développement de nouveaux modèles éco-évolutifs pour évaluer les impacts démographiques des flux de gènes assistés dans le contexte des changements climatiques". PhD thesis. Université de Montpellier.

Fieberg, John and Stephen P. Ellner (2001). "Stochastic matrix models for conservation and management: a comparative review of methods". *Ecology Letters* 4.3, pp. 244–266.

Fréville, Hélène et al. (2004). "Spatial and Temporal Demographic Variability in the Endemic Plant Species Centaurea Corymbosa (asteraceae)". *Ecology* 85.3, pp. 694–703.

Griffith, Alden B. (Dec. 2017). "Perturbation approaches for integral projection models". *Oikos* 126.12, pp. 1675–1686.

Hadjou Belaid, Asma (2018). "Démographie et réponses adaptatives des populations végétales aux changements environnementaux".

Hadjou Belaid, Asma et al. (July 1, 2018). "Predicting population viability of the narrow endemic Mediterranean plant *Centaurea corymbosa* under climate change". *Biological Conservation* 223, pp. 19–33.

Hamilton, W. D. (Sept. 1, 1966). "The moulding of senescence by natural selection". *Journal of Theoretical Biology* 12.1, pp. 12–45.

Kelly, Ella and Ben Phillips (2019). "How many and when? Optimising targeted gene flow for a step change in the environment". *Ecology Letters* 22.3, pp. 447–457.

Kentie, Rosemarie et al. (2020). "Life-history strategy varies with the strength of competition in a food-limited ungulate population". *Ecology Letters* 23.5, pp. 811–820.

Kuss, Patrick et al. (2008). "Evolutionary demography of long-lived monocarpic perennials: a time-lagged integral projection model". *Journal of Ecology* 96.4, pp. 821–832.

McGraw, James B. (1989). "Effects of Age and Size on Life Histories and Population Growth of Rhododendron Maximum Shoots". *American Journal of Botany* 76.1, pp. 113–123.

Merow, Cory et al. (2014). "Advancing population ecology with integral projection models: a practical guide". *Methods in Ecology and Evolution* 5.2. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/2041-210X.12146, pp. 99–110.

Metcalf, C. Jessica E. and Samuel Pavard (Apr. 2007). "Why evolutionary biologists should be demographers". *Trends in Ecology & Evolution* 22.4, pp. 205–212.

Metcalf, C. Jessica E., Karen E. Rose, and Mark Rees (Sept. 1, 2003). "Evolutionary demography of monocarpic perennials". *Trends in Ecology & Evolution* 18.9, pp. 471–480.

Morris, William F. and Daniel F. Doak (2002). *Quantitative conservation biology : theory and practice of population viability analysis*. Sinauer Associates.

Olivieri, Isabelle et al. (Jan. 1, 2016). "Why evolution matters for species conservation: perspectives from three case studies of plant metapopulations". *Evolutionary Applications* 9.1, pp. 196–211.

Picó, F. X. and J. Retana (2008). "Age-specific, density-dependent and environment-based mortality of a short-lived perennial herb". *Plant Biology* 10.3, pp. 374–381.

R Core Team (2023). "R: A Language and Environment for Statistical Computing". URL: https://www.R-project.org/.

Ramula, Satu, Mark Rees, and Yvonne M. Buckley (2009). "Integral projection models perform better for small demographic data sets than matrix population models: a case study of two perennial herbs". *Journal of Applied Ecology* 46.5, pp. 1048–1053.

Rees, Mark and Stephen P. Ellner (2009). "Integral projection models for populations in temporally varying environments". *Ecological Monographs* 79.4, pp. 575–594.

Rees, Mark and Karen E. Rose (July 22, 2002). "Evolution of flowering strategies in Oenothera glazioviana: an integral projection model approach". *Proceedings of the Royal Society of London. Series B: Biological Sciences* 269.1499, pp. 1509–1515.

Rousset, François and Jean-Baptiste Ferdy (2014). "Testing environmental and genetic effects in the presence of spatial autocorrelation". *Ecography* 37.8, pp. 781–790.

Salguero-Gómez, Roberto and Brenda B. Casper (2010). "Keeping plant shrinkage in the demographic loop". *Journal of Ecology* 98.2, pp. 312–323.

Stubben, Chris J and Brook G Milligan (2007). *Estimating and Analyzing Demographic Models Using the popbio Package in R.*

Tredennick, Andrew T. et al. (2021). "A practical guide to selecting models for exploration, inference, and prediction in ecology". *Ecology* 102.6, e03336.

Williams, Jennifer L., Tom E. X. Miller, and Stephen P. Ellner (2012). "Avoiding unintentional eviction from integral projection models". *Ecology* 93.9, pp. 2008–2014.

5 Supplementary materials

Implementation of IPM

For $a \in [2, 7]$:

$$N_a(i, t+1) = \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \int_{\Omega} P(x, a-1, y, a) n_{a-1}(x, t) dx \; dy \simeq \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \sum_{j} P(m_j, a-1, y, a) N_{a-1}(j, t) dy$$

For $a = 8$:

$$N_a(i, t+1) = \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \int_{\Omega} [P(x, a-1, y, a) n_{a-1}(x, t)] + [P(x, a, y, a) n_a(x, t)] dx \; dy$$

$$\simeq \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \sum_{j} [P(m_j, a-1, y, a) N_{a-1}(j, t)] + [P(m_j, a, y, a) N_a(j, t)] dy$$

For $a = 1$:

$$N_1(i, t+1) = \sum_{a} \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \int_{\Omega} F(x, a, y, 1) n_a(x, t) dx \; dy \simeq \sum_{a} \int_{x_{i,a}^{inf}}^{x_{i,a}^{sup}} \sum_{j} F(m_j, a, y, 1) N_a(j, t) dy$$

Tab. 2: Fitted models selected by AIC for each vital rate. x represents the size, a the age, and the random effects are noted $year$, pop, $year : pop$ for population effect nested in year effect, and ε for residual variance.

Survival probability for seedling

$$\text{logit}(s_1(x)) = \begin{cases} year_0 - 1.66(x-0.5) + (3.99 + year_1)(x-0.5)^2 & \text{if } 0.5 \leq x \leq 1 \text{ cm} \\ year_0 + 0.166 + (2.33 + year_1)(x-1) - 2.3(x-1)^2 & \text{if } 1 \leq x \leq 1.5 \text{ cm} \\ year_0 + 0.756 + (0.0292 + year_1)(x-1.5) + 0.0502(x-1.5)^2 & \text{if } 1.5 \leq x \leq 15 \text{ cm} \end{cases}$$

$Var(year_0) = 1.582$; $Var(year_1) = 0.02833$; $Cov(year_0, year_1) = -0.04404271$

Survival probability for rosettes

$$\text{logit}(s_2(x)) = pop_0 + year_0 + spline_1(x) + spline_2(a)$$

$$spline_1(x) = \begin{cases} 1.01(x-0.5) - 0.0958(x-0.5)^2 & \text{if } x \leq 2 \\ 2.52 + 0.246(x-4.5) - 0.0117(x-4.5)^2 & \text{if } x \geq 4.5 \end{cases}$$

$$spline_2(a) = \begin{cases} (0.713 + year_1)(a-2) - 0.478(a-2)^2 + 0.0709(a-2)^3 & \text{if } a \leq 6.5 \\ (0.723 + year_1)(a-6.5) + 0.48(a-6.5)^2 - 0.952(a-6.5)^3 & \text{if } a \geq 6.5 \end{cases}$$

$Var(pop_0) = 0.076$; $Var(year_0) = 1.443$; $Var(year_1) = 0.011$; $Cov(year_0, year_1) = -0.014$

Flowering probability

$$logit(f(x,a)) = -10.21 + pop_0 + 226.93x - 76.57x^2 + 39.27x^3 + (133.48 + pop_1)a - 54.57a^2$$

$Var(pop_0) = 2.252$; $Var(pop_1) = 0.09228$; $Cov(pop_0, pop_1) = -0.2041362$

Plant growth

$$g(y,x,a) = \frac{1}{\sigma(x,a)\sqrt{2\pi}} \exp\left(-\frac{(x-\mu(x,a))^2}{2\sigma(x,a)^2}\right)$$

$$\mu(x,a) = pop_0 + year_0 + poly(x) + spline(a)$$

$$poly(x) = 6.8681 + (175.3943 + year_1)x - 27.1381x^2 - 14.0231x^3$$

$$spline(a) = \begin{cases} (-0.793 + year_2)(a-1) + 0.109(a-1)^2 & \text{if } a \leq 6.5 \\ -1.08 + (0.401 + year_2)(a-6.5) - 0.527(a-6.5)^2 & \text{if } a \geq 6.5 \end{cases}$$

$$\sigma(x,a) = \exp(1.01 + 0.55\log(x) - 0.12\log(a))$$

$Var(pop_0) = 0.331$; $Var(year_0) = 0.3225$; $Var(year_1) = 0.0362$; $Cov(year_0, year_1) = 0.00468$;
$Var(year_2) = 0.0579$; $Cov(year_0, year_2) = -0.0108$; $Cov(year_1, year_2) = -0.00135$

Fecundity (number of capitulas)

$$\log(c(x,a)) = 2.32734 + year_0 + 0.06752x + year_1 a + \varepsilon$$

$Var(year_0) = 0.2127$; $Var(year_1) = 0.0238$; $Cov(year_0, year_1) = -0.00506226$; $Var(\varepsilon) = 0.282$

Seedling size distribution

$$\omega(y) \sim \text{Gamma}(\log(0.2873 + pop_0 + year_0 + pop : year_0 + \varepsilon))$$

$Var(year_0) = 0.0249$; $Var(pop_0) = 0.0248$; $Var(pop : year_0) = 0.0781$; $Var(\varepsilon) = 0.0213$

Establishment rate

$$\xi = 0.338 + \varepsilon$$

$Var(\varepsilon)$ =0.132

Tab. 3: Possible forms for each fixed and random effects.

Fixed effect	Possible forms
Size (x)	polynomial function of degree 1, 2, 3 or 4 splines function of degree 2 with 1 or 2 knots, or of degree 3 with 2 knots
Age (a)	polynomial function of degree 1, 2 or 3 splines function of degree 2 with 1 knot at coordinate 6.5 or of degree 3 with 1 or 2 knots at coordinates 1.5 and 6.5
Year ($year$)	effect on intercept, plus on size, age or both
Population (Pop)	effect on intercept plus, on size, age or both
Interaction year population	effect on intercept
Individual	effect on intercept

Calculation of transition probabilities in MPMs

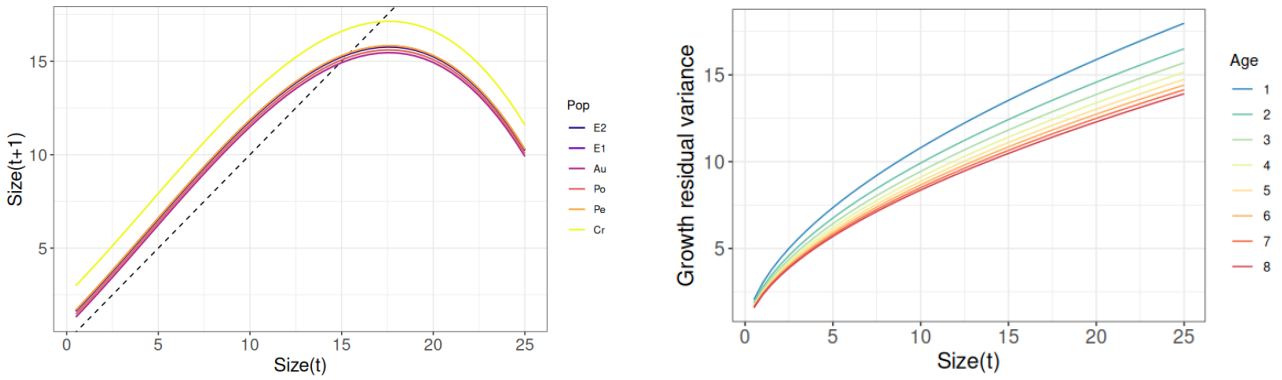
$$s_1(1-\alpha_1) = \frac{\sum_i \sum_j W_{1,j} \tilde{P}(j,a,i,1)(1-f(i,1))}{\sum_i W_{1,i}(1-f(i,1))} \tag{8}$$

$$s_2(1-\alpha_2) = \frac{\sum_{a=2}^{7} \sum_i \sum_j W_{a,j} \tilde{P}(j,a-1,i,a)(1-f(j,a))}{\sum_{a=2}^{8} \sum_i W_{a,i}(1-f(i,a))} \tag{9}$$

$$s_2\alpha_2 = \frac{\sum_{a=2}^{7} \sum_i \sum_j W_{a,j} \tilde{P}(j,a-1,i,a) f(j,a)}{\sum_{a=2}^{8} \sum_i W_{a,i}(1-f(i,a))} \tag{10}$$

$$fs_0 = \frac{\sum_{a=1}^{8} \sum_i \sum_j W_{a,j} \tilde{F}(j,a-1,i,a)}{\sum_{a=2}^{8} \sum_i W_{a,i} f(i,a)} \tag{11}$$

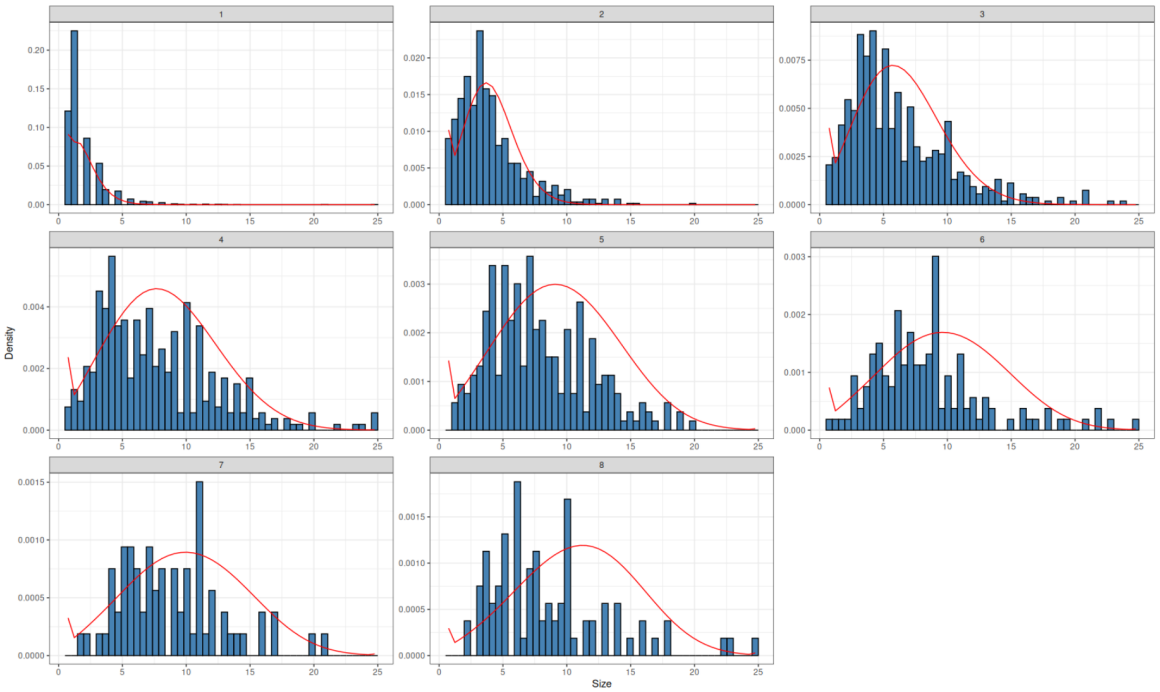
Selected models for vital rates



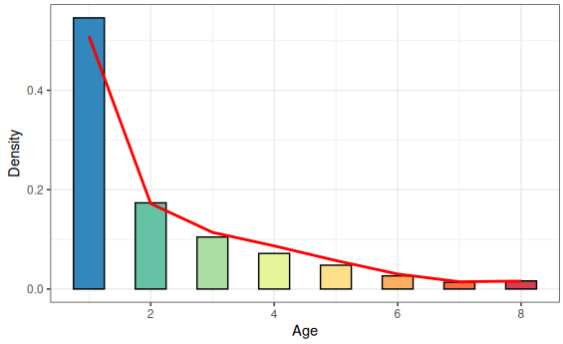
(a) (b)

Fig. 8: (a)Size at year $t+1$ as a function of the size at year t. The dashed line is $\text{Size}(t+1) = \text{Size}(t)$, mean over all ages and years, each line corresponds to a population, the populations are listed in descending order of size; (b) Residual variance of the growth model as a function of the size. Each line correspond to an age class

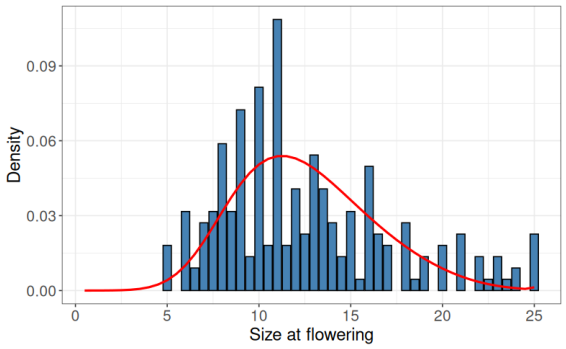
Demographic analyses



(a)



(b)



(c)

Fig. 9: The red curves are (a) Stable size distribution for each age, (b) stable age distribution, (c) stable flowering distribution. The bars are observed (a) size distribution, (b) age distribution, (c) flowering size distribution.

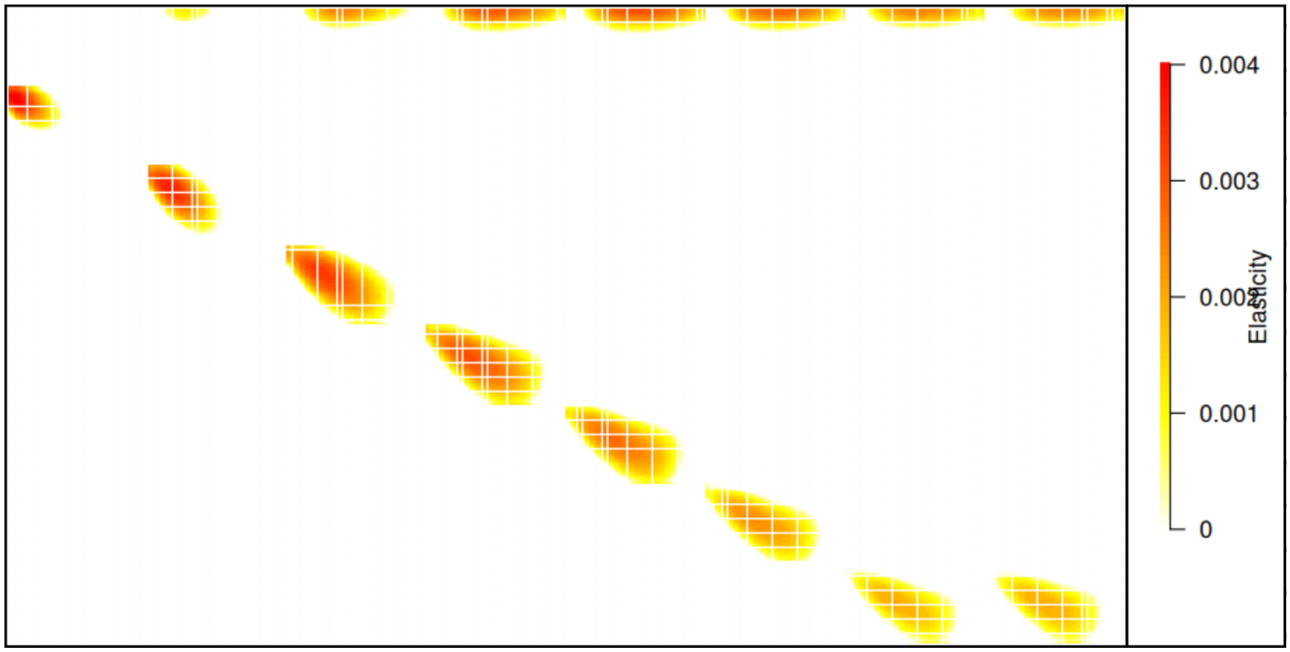


Fig. 10: Elasticity matrix.

Comparison with Matrix Population Models

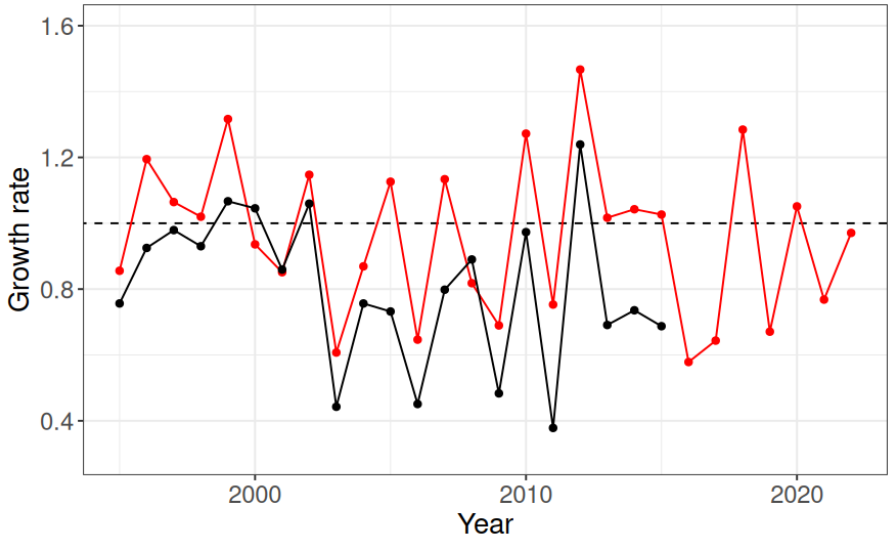


Fig. 11: Variation of growth rate for IPM (in red) and MPM (in black). Dashed line: $\lambda = 1$

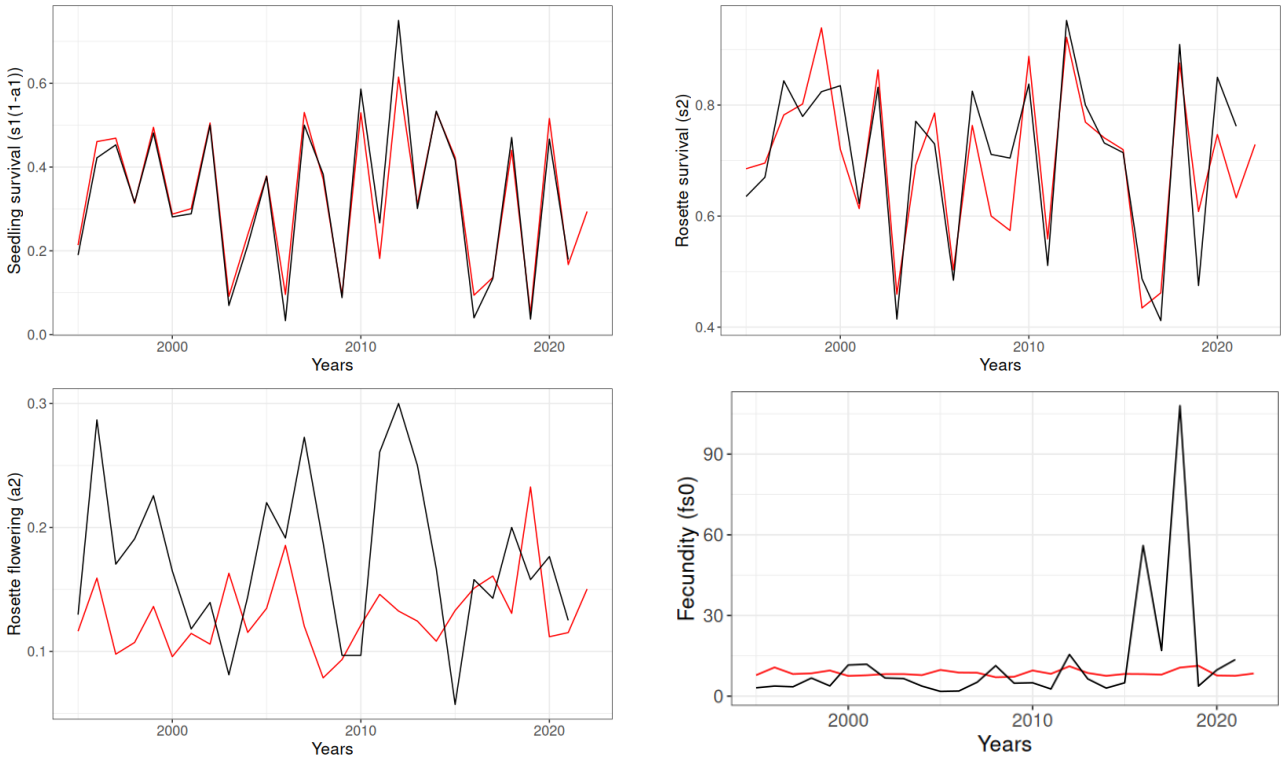


Fig. 12: Variation of life history traits across years, estimated with IPM in red and with MPM in black.

Robustness of predictions

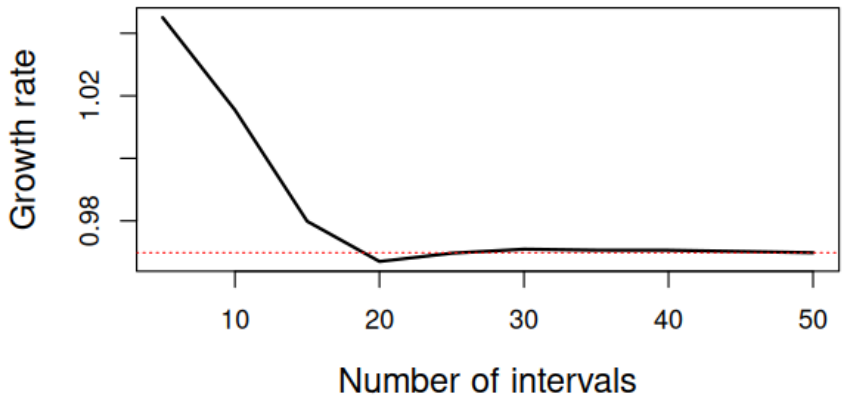


Fig. 13: Effect of number of intervals for IPM discretization on estimated growth rate