PNAS

Main Manuscript for

Social copying drives a tipping point for non-linear population collapse

Daniel Oro[1,2*], Lluís Alsedà[3,4,5], Alan Hastings[2,6], Meritxell Genovart[1], Josep Sardanyés[4,5]

[1]Theoretical and Computational Ecology Laboratory, Centre d'Estudis Avançats de Blanes (CEAB-CSIC), Cala Sant Francesc 14, 17300 Girona, Spain.

[2]Department of Environmental Science and Policy, University of California, Davis, CA 95616, USA.

[3]Departament de Matemàtiques, Universitat Autònoma de Barcelona, 08193 Bellaterra, Spain.

[4]Centre de Recerca Matemàtica, Edifici C, 08193 Campus de Bellaterra, Barcelona, Spain.

[5]Barcelona Graduate School of Mathematics (BGSMath), Edifici C, 08193 Campus de Bellaterra, Barcelona, Spain.

[6]Santa Fe Institute, Santa Fe, NM 87501, USA.

* Correspondence to: Daniel Oro

Email: d.oro@csic.es

Author Contributions: D.O. conceived, designed and conducted the population and other fieldwork data; D.O. suggested the main ideas and concepts and took part in all the analyses. All authors conceived and designed the mathematical model. Ll.A. and J.S. analyzed the mathematical models. Ll.A. programmed the optimization algorithms. All authors analyzed the data, wrote the article and gave final approval for submission.

Competing Interest Statement: The authors declare no competing interests.

Classification: Paste the major and minor classification here. Dual classifications are permitted, but cannot be within the same major classification.

Keywords: tipping points, runaway dispersal, non-linear population dynamics, social behavior, feedback.

This PDF file includes:

Main Text
Figures 1 to 4

Abstract

Sudden changes in populations are ubiquitous in ecological systems, especially under perturbations. The agents of global change may increase the frequency and severity of anthropogenic perturbations, but complex responses of populations hamper our understanding of their dynamics and resilience. Furthermore, the long-term environmental and demographic data required to study those sudden changes are rare. Fitting dynamical models with an artificial intelligence algorithm to population fluctuations over 40 years in a social bird reveals that feedbacks in dispersal after a cumulative perturbation drives a population collapse. The collapse is well described by a non-linear function mimicking social copying, whereby dispersal made by a few individuals induces others to leave the patch in a behavioral cascade for decision-making to disperse. Once a threshold for deterioration of the quality of the patch is crossed, there is a tipping point for a social response of runaway dispersal corresponding to social copying feedback. Finally, dispersal decreases with population density likely due to the long times spent in a quasi-extinction state as observed in many populations of social animals after occupying a patch for extended periods. In providing the first evidence of copying for the emergence of feedbacks in dispersal in a social organism, our results suggest a broader impact of self-organized collective dispersal in complex population dynamics. This has also implications for the theoretical study of population and metapopulation non-linear dynamics, including population extinction, and the management of endangered and harvested populations of social animals subjected to behavioral feedback loops.

Significance Statement

Among the complex dynamics arising in all living systems, sudden population collapses are one of the most fascinating. Understanding the mechanisms that may cause these collapses is fundamental to the conceptual study of population dynamics. We fit dynamical models to population fluctuations over 40 years in a social bird that showed an unexpected collapse after a perturbation press that progressively eroded environmental conditions at the world's most suitable breeding patch. We demonstrated that collapse was explained by density-dependence feedbacks related to the simple behavior of social copying for dispersal to other patches. The significance of our study lies in showing that environmental stochastic perturbations may trigger a tipping point by runaway dispersal driving populations to a new state of quasi-extinction.

Main Text

Introduction

Understanding abrupt declines in the responses of populations to environmental perturbations is crucial for the theoretical study of population extinction and for managing harvested and endangered species, especially under the impacts of global change (1–3). Under the conceptual framework of the logistic model, the transition from a population level near carrying capacity to collapse in populations subjected to perturbations should occur through a negative exponential decay, i.e. a density-independent process (4–6). However, the logistic model has several limitations, since it assumes both a linear association between density and growth rate and a lack of time lag in the response of individuals to changes in density (5). Time-lagged responses may generate transient phenomena, which can explain abrupt regime shifts that are not directly associated with environmental changes (7, 8).

Dramatic sudden collapses in populations may be especially likely to occur in social organisms. Social groups are complex systems in which the number of interactions within a group is not additive, but grows in a factorial manner resulting in important behavioral feedback loops, such as those arising for information gathering, social copying and group cohesion. These feedback loops

are autocatalytic: the more individuals perform a certain behavior, the more other individuals decide to follow suit. The density of conspecifics drives the individual decision-making for staying or dispersing in a patch when the trade-off between the benefits of foraging and the risks of predation is considered (9). This is linked with collective behavioral loops that trigger informational cascades used in the location of resources and threats (10). Collective behavior may influence population responses such as tipping points, especially when perturbations drive negative growth rates (11–13). Collective behavior may influence complex population responses such as tipping points for regime shifts, boom-bust dynamics and chaotic dynamics, all of them having potential consequences for population extinction (11–14). A process linking behavior to a demographic trait with consequences for populations is decision-making for dispersal (14). How dispersal is influenced by copying behavior in social animals, including humans, remains unexplored, but empirical and theoretical studies show that copying is favored under perturbations (4, 10). Less is even known about the consequences of that copying in dispersal and other behavioral feedbacks for the occurrence of tipping points and the collapse of populations. The main reason for this is that theoretical population dynamic models including a mathematical expression of dispersal, such as coupled logistic equations, mostly consider dispersal as simple passive diffusion and not as a density-dependent, non-linear process (15, 16) (Fig. 1). Determining how populations of social organisms may collapse and diagnosing what are the behavioral mechanisms involved is hampered by the challenges of collecting long-term data on both population fluctuations and environmental changes, including perturbations (17). Perturbations may have different temporal features, such as stochastic pulses and presses, and these dynamics are crucial to understanding the impact and responses they may generate on populations.

An unprecedented opportunity to investigate the effects of social feedbacks on population non-linear collapses is afforded by 40 years of population and environmental data of the colonial long-lived Audouin's gull (*Ichthyaetus audouinii*) breeding seasonally at Punta de la Banya (Ebro Delta, western Mediterranean) (Fig. 2). Despite being a species with a slow evolutionary life history, the patch held almost 50% of the total world population after only 6 years since colonization, due to its habitat suitability allowing for extremely high rates of survival, fertility and immigration from other patches (18, 19). These rates were enhanced by low interference competition with other sympatric species of the same ecological guild and by the absence of terrestrial predators (20, 21). Population continued to increase through a rapid transition ending in a dynamic equilibrium state, which occurred when predators invaded the patch (Fig. 2A). Predator densities did not vary over the years, but the perturbation was pressing over the breeding seasons eroding habitat quality and causing not only a slight decrease of adult survival, but mainly a high dispersal to other occupied patches and especially to empty patches for colonization (22–24) (Fig. 2B, *SI Appendix,* section S1). Since we accurately monitored the metapopulation at a large spatial scale, we quantify dispersal processes of emigration and immigration among the different patches, including extinction and colonization events over the years (Fig. 2B). Just after peaking at a maximum density and holding 73% of the total world population, the population underwent a steady decline leading to collapse, when it held only 3% of that total world population (Fig. 2A). In summary, we tested the hypothesis that the arrival of predators triggered a tipping point by runaway dispersal driven by social copying behavior that caused a population collapse.

To test our hypothesis, we first built a population model that described the local population dynamics as a single-patch system and considered immigration and dispersal of individuals within the metapopulation, which was not explicitly considered. The model took also into account other ecological processes, such as intra-specific competition for resources and density-independent death rates. In previous studies, we accurately estimated death rates by capture-recapture modeling of +30,000 banding birds and +28,000 resights over 32 years encompassing 69 local patches of the metapopulation (+90% of the total world population)(21, 23). Metapopulation dynamics and most dispersal processes in terms of density of individuals were governed by the dynamics of the study patch, since it held a large percentage of the total world population (Fig. 2A)(23). Importantly, for modelling the dispersal out of the study patch in our model, we included two different forms of dispersal to assess which one played a major role in explaining the population dynamics observed since the arrival of predators. One dispersal form was density-independent and

varied with a constant rate, whereas the other dispersal form was density-dependent and mimicked the behavior of social copying (4). To assess how population density influenced dispersal, we used an Elliot function, which allowed us to obtain several different shapes of density-dependence. The function typically increases at decreasing population values (i.e. dispersal would increase with decreasing population density), but it allows some other forms, such as sigmoidal shapes, sudden non-linear changes with density and density-independent dispersal (*SI Appendix,* section S2, Figs. S5.1 and S5.2). For instance, non-linear shapes with a drastic increase of dispersal beyond a threshold value of population density would fit with a tipping point for runaway dispersal by social copying (4, 25). Finally, we developed a new optimization technique to fit population models to our real data on stochastic population dynamics using artificial intelligence and genetic algorithms (*SI Appendix,* section S6).

Results and Discussion

The geometric mean of the population growth rate expressed as x_{t+1} / x_t, being x_t the size of the population at time t since colonization (t_0 = 1981) was high and it showed large fluctuations (mean = 1.106; SD = 0.835), mostly due to the influence of large rates of both immigration and dispersal (*SI Appendix,* Fig. S1.1). Population growth rate estimated by the model fitting our data during the initial phase without carnivores suggested an exponential growth (growth rate in linear-logarithmic scale = 0.285, SE = 0.033, correlation coefficient r^2 = 0.910, N = 17, *SI Appendix,* Fig. S4.1). Furthermore, the model clearly showed a logistic behavior with a growth rate = 0.349 year^{-1} and carrying capacity = 18,882 birds (least squares |LS| = 2,593, N = 17) (*SI Appendix,* sections S3 and S4)(Fig. 3). During this initial phase, the population was in a transient state and it did not reach the predicted carrying capacity even though the patch already hold a large proportion of total world population, and this was likely due to the arrival of predators (Fig. 3B) (*SI Appendix,* Fig. S4.1). This was confirmed when we fitted field data with the logistic model from 1981 to 2004 to test whether the model without considering the impact of predators on the dynamics performed well, but the fitting was worse (|LS| = 5,639, N = 24) (*SI Appendix,* lines 572–577, Fig. 7.2).

After the onset of perturbation, demographic noise and density-dependent dispersal increased, and the population showed a slight decrease (Fig. 2A, *SI Appendix,* Fig. S7.1). This decrease could not be attributed to a deterioration in environmental conditions, since all environmental parameters influencing resource availability and predator (both their densities and killing rates) did not significantly change during this phase (*SI Appendix,* Figs. S1.1 and S1.2, sections S5 and S7)(4, 23). The increase in demographic noise likely caused the worst fit of the model compared to the initial phase (correlation coefficient r^2 = 0.531, N = 8, *SI Appendix,* Fig. S7.3). Despite the low amount of population data for this phase, the parameters estimated for the two types of dispersal (linear and positive density-dependent) also indicated that the dominant dispersal was due to social copying (*SI Appendix,* lines 587–593).

For the final phase of population collapse, the model fitted the population non-linear decline with high accuracy (correlation coefficient r^2 = 0.977, |LS| = 2,427, N = 12). The model considered an inverse density-dependent dispersal mimicking social copying for dispersal (4) (Fig. 4). Contrarily to what is predicted by simple theory, models considering negative exponential growth towards extinction had a worse fit, similar to models with density-independent and positive density-dependent dispersal (26)(*SI Appendix*, section S6). The set of parameters that better fitted real data for describing the inverse density-dependence dispersal function showed that during the first years of the collapse phase, dispersal increased only slightly with a decrease in population density (*SI Appendix*, section S6.3). Interestingly, dispersal accelerated once a tipping point of population density was trespassed (Fig. 4B). After attaining its maximum rate, dispersal decreased with population density (Fig. 4B). This contrasts with the simple linear dispersal process assumed in simpler models.

We used a novel mathematical technique fitting a population model to real data for exploring which type of dispersal may explain the collapse of a local population. Our capacity to study how perturbations may cause population collapses is increasing in recent decades, thanks to the greater

availability of long-term data on both populations and environmental variability (27–29). Furthermore, population models are progressively considering the complex nature of population dynamics by incorporating simultaneously density-dependent mechanisms and the different types of stochasticity (mainly environmental and demographic)(30–33). Nevertheless, disentangling the role of endogenous (e.g. density-dependence) and exogenous (density-independent, environmental perturbations) drivers on population declines remains a challenge, especially when environmental changes are not accurately monitored (34). Further challenges are modeling explicitly dispersal processes to understand local and metapopulation dynamics and to assess the role of density-dependence on those processes, which may be particularly crucial during phases of colonization and declines (i.e. immigration and emigration respectively)(18, 35). Here, we built a mechanistic population model to explain the population dynamics since colonization and its collapse in a social bird at a patch that went from holding 73% of total world population to only 3% in just 10 years. Results showed that the collapse was caused by a non-linear density-dependent dispersal that mimicked behavior of social copying. The generality of copying in social animals is supported by the bulk of empirical evidences showing copying in human behavior when decision-making is at play and the framework of the theory of conformity (36). Other than primates, social copying has been found in a broad range of animals, even in simple organisms, and it may have important consequences for the individual, e.g. for mate choice and thus for its fitness (37, 38). In an ecological context, theoretical models exploring how individuals weigh private and social information to make decisions showed that following the majority option is enhanced by deteriorating environmental conditions, e.g. the presence of a predator (10).

Beyond the behavioral consequences that social copying may have for individual fitness, very little attention has been devoted to the potential effects that copying may have for generating non-linear density-dependence dispersal and for population dynamics (4, 39). In colonial long-lived species with spatially-structured populations, individuals dispersing may use inadvertent social information available at each patch (in the form of conspecific density and the fertility of conspecifics) to assess its suitability and make the decision about where to settle each reproductive season (40–42). Thus, social copying would influence the decision-making in the trade-off between staying in the patch and dispersing to other patches, which may influence dispersal at the level of the group and the dynamics of local populations and metapopulations (e.g. periodic dynamics, boom-bust, rescue effects, source-sink dynamics). Furthermore, some individuals of the population (e.g. specific ages, sexes) may be more prone to disperse and generate demographic heterogeneity among patches, then increasing stochasticity and spatial variability in extinction risk (24, 31).

The first non-linear response that the study population showed related to social processes occurred after the onset of perturbation. Population fluctuations increased mainly due to an unexpected sharp and short increase in population to reach the maximum density, and this corresponded to what resilience theory calls demographic compensation (43). Compensation was likely caused by high local recruitment and a decrease in inter-specific competition favored by a pulse of higher food availability recorded at the patch (Fig. 2A, fig. S1.2) (21, 43, 44). The transient from the arrival of predators to the beginning of the phase of collapse lasted nearly a generation time of the species. That this long time elapsed under the perturbation press regime was consistent with predictions of the transient theory for fast-slow dynamics: while population growth after colonization of new patches can be extremely fast (4), it only occurs once individuals gather enough information about the suitability of alternative patches for what is called informed dispersal (45, 46). This behavioral process of prospecting to gather information may be slow in empty patches due to the lack of public information, which is used as a cue for assessing the quality of occupied patches (fig. S1.2) (4, 22, 47, 48). When predators first invaded La Banya, only a few small patches were occupied and their population growth was constrained by competition for breeding sites, thus dispersal to these patches, where social information was available, was challenging (Fig. 2B)(22, 23).

At this point, it seems clear that population dynamics of social organisms in deteriorating environments may be more influenced by transient phenomena and density-dependent mechanisms than solitary, territorial species (7). Empirical data on declining populations of non-

social organisms suggest in fact that their dynamics should follow a density-independent decrease towards extinction (49, 50). The existence of social behavioral feedback, from very simple (e.g. social copying) to more complex (e.g. cooperation, hierarchies), may increase stochasticity for local population processes like abrupt extinctions and the appearance of tipping points for collapses (4, 12, 51, 52). Nevertheless, disentangling the different processes that, alone or in interaction, may drive population non-linear declines, and the mechanisms behind those processes that may cause the tipping point to appear, remain a challenge (53, 54). Out of mesocosmos experimental studies, interpreting complex local population dynamics in empirical studies needs accurate monitoring of both metapopulation dynamics and environmental stochasticity. The exceptional fitting of our population model during the phase of collapse is an example of how an environmental stochastic perturbation (i.e. the arrival of predators) triggered a tipping point by runaway dispersal, i.e. a density-dependent accelerating dispersal caused by positive behavioral feedback that drove the population to a new state of quasi-extinction. This type of tipping, which occurs by the progressive loss of the resilience of the current population state, can potentially be anticipated using the tools of early warning signals (55).

Conclusions

The critical significance of our results is that social copying, a simple behavioral mechanism operating in all social organisms, may generate non-linear population collapses by the occurrence of tipping points in dispersal processes. These dynamics fit very well with what has been observed in populations of social animals subjected to press perturbations, such as superabundant species subjected to culling and humans during warfare (4). Under these circumstances, there is an initial reluctance to leave the patch due to the large availability of information, the force of social cohesion and the evolutionary advantages of being philopatric against being disperser, which is riskier in terms of fitness prospects (22, 56). In humans, this is reinforced by a sunk-cost effect, which prevents people from abandoning ways of living, cultures and beliefs (57). However, once a threshold for deterioration of the quality of the patch is trespassed, there is a tipping point for a social response of runaway dispersal corresponding to social copying feedback. Finally, dispersal decreases with population density likely due to the long queues of quasi-extinction state observed in many populations of social animals after occupying a patch for extended periods (4). Our study emphasizes the value of combining long-term demographic and environmental data with mathematical modelling to uncover the behavioral mechanisms driving non-linear responses of populations of social organisms under perturbations. Our results suggest a broader impact of self-organized collective dispersal in complex population dynamics, such as tipping points and transient phenomena. This has also implications for the theoretical study of population and metapopulation non-linear dynamics, including population extinction and resilience, predator-prey dynamics and the management of endangered and harvested populations of social animals subjected to behavioral feedback loops.

Materials and Methods

Study system and species. Demographic population monitoring of Audouin's gulls (*Ichthyaetus audouinii*) at Punta de la Banya (Ebro river Delta: 40°34010:8900N, 0°39034:2800E) started in 1981, when the patch was colonized, and it has been continuously performed until present during four decades (1981-2021). La Banya is a 2500ha sandy peninsula formed by sandy dunes covered by halophilous vegetation. The site, once effectively protected in the early 80s, held extremely suitable environmental conditions for breeding ground-nesting waterbirds (i.e. large availability of food and reproductive habitat free of predators). Audouin's gulls are long-lived social birds with a bet-hedging life history. They have evolved to cope with ephemeral habitats typical of Mediterranean marshes (58). Consequently, Audouin's gulls have nomadic behavior between breeding seasons: when patch conditions change and worsen, then individuals are more prone to disperse mainly to other occupied sites, but in recent years and after the arrival of invasive carnivores in the late 90s, colonization rate of new patches has largely increased. Ground-nesting

gulls have not developed evolutionary defenses to cope with terrestrial predators like carnivores, and this is why they select breeding patches isolated and protected against carnivore invasions. The long-term monitoring at the metapopulation scale allowed us a good knowledge of the system in terms of population fluctuations (Fig. 2B, *SI Appendix*, Fig. S1.1), interference competition with species of the same ecological guild, demographic parameters (e.g. survival, fertility, recruitment, dispersal), and environmental stochasticity, including the invasion of predators, the occurrence of other perturbations and fluctuations in food availability over the years (*SI Appendix*, lines 46–67) (18, 20, 21, 23).

Fieldwork methods. Results regarding the methods used to monitor demographic parameters and population size were published previously (21), and exhaustive details concerning the fieldwork protocols and sites can be found therein and in the *SI Appendix* (lines 69–115). Briefly, population density, which corresponded to the number of breeding females, was estimated since colonization in 1981 by counting the nests containing eggs just before hatching, when biases due to the variance in laying dates were minimal. Errors at counting nests were quantified and were small and constant over the years (<5%). The same procedure was followed to estimate population density of the main competitor of Audouin's gull, the larger yellow-legged gull (*Larus michahellis*). The density of carnivores (mostly foxes *Vulpes vulpes*) and the number of killings were estimated every year since their arrival in 1997 thanks to the tracks in the sand and the intense monitoring of roosting and breeding dens. Other impacts affecting demographic parameters, such as mortality by fishery bycatch, were also duly monitored. Finally, we used the statistics of landings of trawling boats in the harbors close to the study site. Landings are highly correlated with the amounts of fish discarded, which represent up to 70% of biomass ingested by the gulls and greatly influence their fertility. We showed that the population collapse was not caused by a deterioration of environmental conditions but by the stochastic fluctuations of carnivore densities and their predation rates (*SI Appendix*, Fig. S1.2, lines 117–137).

Mathematical dynamical model. Due to the very low number of predators (22), we did not consider predator-prey dynamics explicitly as a possible ecological process causing population decline by direct mortality. To analyze population dynamics from colonization through collapse, we used instead an ordinary differential equation logistic population model and a density-dependent migration term (*SI Appendix*, lines 139–156). The model describes the population dynamics in a single-patch system but taking into account immigration and dispersal of individuals:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = (\vartheta + \omega)x(t)\left(1 - \frac{x(t)}{K}\right) - \varepsilon x(t) - [\rho x(t) + \lambda D(x(t))]$$

where $x(t)$ denotes population density at time t, ϑ is the intrinsic reproduction rate, ω is the immigration rate from other patches, and ε is the annual death rate estimated from field data (0.11 year^{-1}, estimated from long-term monitoring (1988-2015) using capture-recapture modeling (21)). The model introduces intra-specific competition for resources constrained by a logistic function with carrying capacity K, and importantly, it included two distinct dispersal terms: ρ is the linear (exponential) dispersal rate to other patches, and function $D(x(t))$ introduces a positive density-dependent term of dispersal by social copying of individuals leaving the patch, with dispersal rate λ.

We applied the model to fit empirical population fluctuations separately for three phases in which we divided the dynamics of the population (*SI Appendix*, lines 156–170): first, from colonization to the arrival of predators (initial phase: 1981-1997). We expected the initial phase to be characterized by a logistic growth due to the absence of predators and the lack of competition (i.e. density-dependence for food resources). Here, we did not consider dispersal triggered by the perturbation since predators were not present in the patch during this period, and this phase was dominated by high immigration rates from outside (18). For this initial phase, we computed the potential function as a method to visualize the stability of the dynamics at carrying capacity K (*SI Appendix*, lines

190–194, Fig. S3.2). For fitting the population model, we used a standard *trust region method* and the Levenberg-Marquardt algorithm (*SI Appendix*, lines 207–217). Second, we analyzed the phase since the establishment of predators to the maximum density (onset of perturbation phase: 1998-2005). Here, we incorporated the social copying dispersal triggered by the arrival of predators. During this phase, there was a decrease in population density with a sudden increase in the last years to attain the maximum densities in the years 2005-2006 (*SI Appendix*, Fig. S1.3). Due to the short number of years, we estimated the trend in population density fitting the model under different assumptions (*SI Appendix*, lines 558–593). For instance, we did not expect that the logistic model provided a good fitting for this phase, due to the presence of carnivores. We tested whether there was positive density-dependent term of dispersal by social copying for the function $D(x(t))$ of our model by using different values of a parameter in the so-called Elliot function (see below) that drove the sharpness of the jump and the smoothness of $D(x(t))$ (*SI Appendix*, lines 257–275, Figs. S5.1, S5.2 and S5.3). Finally, we will consider the phase encompassing from the maximum density attained to patch collapse (collapse phase: 2006-2017) (Fig. 2, *SI Appendix*, Fig. S1.3, lines 347–556). For the collapse phase, we included a linear dispersal and dispersal by social copying into the model to explore which type of dispersal fitted with the real data:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \varphi x(t) - \beta x(t)^2 - \lambda D(x(t), \mu, \sigma, \delta)$$

where φ represents the neat population growth rate including linear (exponential) dispersal, β is the intrinsic growth rate over the carrying capacity, and $\lambda, \mu, \sigma, \delta$ are the parameters concerning the dispersal rate by social copying: λ is the dispersal rate, μ reflects the tendency of dispersal for small population sizes, σ defines the sharpness and smoothness of the dispersal function, and δ models the transition between small and large population sizes (*SI Appendix*, lines 225–228).

By fitting the population model during this phase (*SI Appendix*, lines 347–386, Table 6.1), we tested whether the presence of predators, once the population started to sharply decline, triggered a density-dependent dispersal, i.e. the less individuals at the patch, the faster the dispersal due to social copying (Fig. 1). This behavioral response was modeled again with an Elliot function, for which we searched for the model parameters best fitting our population data using artificial intelligence and genetic algorithms (*SI Appendix*, lines 420–538).

Acknowledgments

The authors thank the many fieldworkers who have contributed to data collection over the years, particularly Albert Bertolero and the technical staff of the Ebro Delta Natural Park. Data collection has been funded by the Spanish Ministry of Science, the Spanish AEI, THE Fulbright Commission and European Union FEDER (last grants: CGL2017-85210-P (to D.O.), PID2021-122893NB-C21 (to D.O.), Salvador de Madariaga-Fulbright (to D.O.), RYC-2017-22243 (to J.S.) and RTI-2018-098322-B-I00 (to J.S.)).

References

1. A. J. Underwood, The analysis of stress in natural populations. *Biol. J. Linn. Soc.* **37**, 51–78 (1989).

2. M. S. Boyce, C. V. Haridas, C. T. Lee, the N. S. D. W. Group, Demography in an increasingly variable world. *Trends Ecol. Evol.* **21**, 141–148 (2006).

3. M. G. Burgess, *et al.*, Range contraction enables harvesting to extinction. *Proc. Natl. Acad. Sci.* **114**, 3945–3950 (2017).

4. D. Oro, *Perturbation, Behavioural Feedbacks, and Population Dynamics in Social Animals* (Oxford University Press, 2020).

5. F. B. Hanson, H. C. Tuckwell, Logistic growth with random density independent disasters. *Theor. Popul. Biol.* **19**, 1–18 (1981).

6. B.-E. Sæther, S. Engen, Pattern of variation in avian population growth rates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **357**, 1185–1195 (2002).

7. A. Hastings, *et al.*, Transient phenomena in ecology. *Science* **361**, eaat6412 (2018).

8. R. Lande, Risks of Population Extinction from Demographic and Environmental Stochasticity and Random Catastrophes. *Am. Nat.* **142**, 911–927 (1993).

9. M. A. Gil, M. L. Baskett, S. B. Munch, A. M. Hein, Fast behavioral feedbacks make ecosystems sensitive to pace and not just magnitude of anthropogenic environmental change. *Proc. Natl. Acad. Sci.* **117**, 25580–25589 (2020).

10. A. Pérez-Escudero, G. G. de Polavieja, Adversity magnifies the importance of social information in decision-making. *J. R. Soc. Interface* **14**, 20170748 (2017).

11. L. Avilés, P. Abbot, A. D. Cutter, Population Ecology, Nonlinear Dynamics, and Social Evolution. I. Associations among Nonrelatives. *Am. Nat.* **159**, 115–127 (2002).

12. L. Avilés, Cooperation and non-linear dynamics: An ecological perspective on the evolution of sociality. *Evol. Ecol. Res.* **1**, 459–477 (1999).

13. V. Dakos, A. Hastings, Editorial: special issue on regime shifts and tipping points in ecology. *Theor. Ecol.* **6**, 253–254 (2013).

14. S. Arganda, A. Pérez-Escudero, G. G. de Polavieja, A common rule for decision making in animal collectives across species. *Proc. Natl. Acad. Sci.* **109**, 20508–20513 (2012).

15. A. Hastings, Complex Interactions Between Dispersal and Dynamics: Lessons From Coupled Logistic Equations. *Ecology* **74**, 1362–1372 (1993).

16. J. M. Travis, D. J. Murrell, C. Dytham, The evolution of density-dependent dispersal. *Proc. R. Soc. Lond. SeriesB* **266**, 1837–1842 (1999).

17. Z. Ratajczak, *et al.*, Abrupt Change in Ecological Systems: Inference and Diagnosis. *Trends Ecol. Evol.* **33**, 513–526 (2018).

18. D. Oro, G. D. Ruxton, The formation and growth of seabird colonies: Audouin's gull as a case study. *J. Anim. Ecol.* **70**, 527–535 (2001).

19. G. Tavecchia, R. Pradel, M. Genovart, D. Oro, Density-dependent parameters and demographic equilibrium in open populations. *Oikos* **116**, 1481–1492 (2007).

20. D. Oro, *et al.*, Interference competition in a threatened seabird community: A paradox for a successful conservation. *Biol. Conserv.* **142**, 1830–1835 (2009).

21. M. Genovart, D. Oro, S. Tenan, Immature survival, fertility, and density dependence drive global population dynamics in a long-lived species. *Ecology* **99**, 2823–2832 (2018).

22. A. Payo-Payo, *et al.*, Colonisation in social species: the importance of breeding experience for dispersal in overcoming information barriers. *Sci. Rep.* **7**, 42866 (2017).

23. A. Fernández-Chacón, *et al.*, When to stay, when to disperse and where to go: survival and dispersal patterns in a spatially structured seabird population. *Ecography* **36**, 1117–1126 (2013).

24. A. Payo-Payo, *et al.*, Predator arrival elicits differential dispersal, change in age structure and reproductive performance in a prey population. *Sci. Rep.* **8**, 1971 (2018).

25. E. H. van Nes, *et al.*, What Do You Mean, 'Tipping Point'? *Trends Ecol. Evol.* **31**, 902–904 (2016).

26. O. Ovaskainen, B. Meerson, Stochastic models of population extinction. *Trends Ecol. Evol.* **25**, 643–652 (2010).

27. N. Reid, *et al.*, Regime shift tipping point in hare population collapse associated with climatic and agricultural change during the very early 20th century. *Glob. Change Biol.* **27**, 3732–3740 (2021).

28. J. A. Hutchings, J. D. Reynolds, Marine Fish Population Collapses: Consequences for Recovery and Extinction Risk. *BioScience* **54**, 297–309 (2004).

29. J. F. McLaughlin, J. J. Hellmann, C. L. Boggs, P. R. Ehrlich, The route to extinction: population dynamics of a threatened butterfly. *Oecologia* **132**, 538–548 (2002).

30. P. Turchin, A. D. Taylor, Complex Dynamics in Ecological Time Series. *Ecology* **73**, 289–305 (1992).

31. B. A. Melbourne, A. Hastings, Extinction risk depends strongly on factors contributing to stochasticity. *Nature* **454**, 100–103 (2008).

32. W. W. Murdoch, Population regulation in theory and practice. *Ecology* **75**, 271–287 (1994).

33. Z. Zeng, R. M. Nowierski, M. L. Taper, B. Dennis, W. P. Kemp, Complex population dynamics in the real world: modeling the influence of the time-varying parameters and time lags. *Ecology* **79**, 2193–2209 (1998).

34. T. J. Hefley, M. B. Hooten, J. M. Drake, R. E. Russell, D. P. Walsh, When can the cause of a population decline be determined? *Ecol. Lett.* **19**, 1353–1362 (2016).

35. S. D. Bona, *et al.*, Spatio-temporal dynamics of density-dependent dispersal during a population colonisation. *Ecol. Lett.* **22**, 634–644 (2019).

36. B. D. Bernheim, A Theory of Conformity. *J. Polit. Econ.* **102**, 841–877 (1994).

37. E. Danchin, *et al.*, Cultural flies: Conformist social learning in fruitflies predicts long-lasting mate-choice traditions. *Science* **362**, 1025–1030 (2018).

38. A. D. Davies, Z. Lewis, L. R. Dougherty, A meta-analysis of factors influencing the strength of mate-choice copying in animals. *Behav. Ecol.* **31**, 1279–1290 (2020).

39. M. Sapage, S. A. M. Varela, H. Kokko, Social learning by mate-choice copying increases dispersal and reduces local adaptation. *Funct. Ecol.* **35**, 705–716 (2021).

40. D. Parejo, D. Oro, E. Danchin, Testing habitat copying in breeding habitat selection in a species adapted to variable environments. *Ibis* **148**, 146–154 (2006).

41. D. Serrano, J. L. Tella, Dispersal within a spatially structured population of lesser kestrels: the role of spatial isolation and conspecific attraction. *J. Anim. Ecol.* **72**, 400–410 (2003).

42. E. Danchin, T. Boulinier, M. Massot, Conspecific reproductive success and breeding habitat selection: implications for the study of coloniality. *Ecology* **79**, 2415–2428 (1998).

43. P. Capdevila, I. Stott, M. Beger, R. Salguero-Gómez, Towards a Comparative Framework of Demographic Resilience. *Trends Ecol. Evol.* **35**, 776–786 (2020).

44. P. A. Abrams, When does greater mortality increase population size? The long history and diverse mechanisms underlying the hydra effect. *Ecol. Lett.* **12**, 462–474 (2009).

45. J. Clobert, J.-F. L. Galliard, J. Cote, S. Meylan, M. Massot, Informed dispersal, heterogeneity in animal dispersal syndromes and the dynamics of spatially structured populations. *Ecol. Lett.* **12**, 197–209 (2009).

46. M. M. Delgado, K. A. Bartoń, D. Bonte, J. M. J. Travis, Prospecting and dispersal: their eco-evolutionary dynamics and implications for population patterns. *Proc. R. Soc. B Biol. Sci.* **281**, 20132851 (2014).

47. M. A. Gil, A. M. Hein, O. Spiegel, M. L. Baskett, A. Sih, Social Information Links Individual Behavior to Population and Community Dynamics. *Trends Ecol. Evol.* **33**, 535–548 (2018).

48. D. Oro, J. Bécares, F. Bartumeus, J. M. Arcos, High frequency of prospecting for informed dispersal and colonisation in a social species at large spatial scale. *Oecologia* **197**, 395–409 (2021).

49. B.-E. Sæther, *et al.*, Time to Extinction of Bird Populations. *Ecology* **86**, 693–700 (2005).

50. B. C. Lister, A. Garcia, Climate-driven declines in arthropod abundance restructure a rainforest food web. *Proc. Natl. Acad. Sci.* **115**, E10397–E10406 (2018).

51. B. C. Lister, Information, behaviour and population dynamics. *Oikos* **123**, 1431–1438 (2014).

52. A. R. Bhowmick, B. Saha, J. Chattopadhyay, S. Ray, S. Bhattacharya, Cooperation in species: Interplay of population regulation and extinction through global population dynamics database. *Ecol. Model.* **312**, 150–165 (2015).

53. C. N. Anderson, *et al.*, Why fishing magnifies fluctuations in fish abundance. *Nature* **452**, 835 (2008).

54. V. Dakos, *et al.*, Ecosystem tipping points in an evolving world. *Nat. Ecol. Evol.* **3**, 355 (2019).

55. L. Dai, D. Vorselen, K. S. Korolev, J. Gore, Generic Indicators for Loss of Resilience Before a Tipping Point Leading to Population Collapse. *Science* **336**, 1175–1177 (2012).

56. M. Genovart, *et al.*, Decrease in social cohesion in a colonial seabird under a perturbation regime. *Sci. Rep.* **10**, 18720 (2020).

57. M. Scheffer, Anticipating societal collapse; Hints from the Stone Age. *Proc. Natl. Acad. Sci.* **113**, 10733–10735 (2016).

58. A. Martínez-Abraín, D. Oro, M. G. Forero, D. Conesa, Modeling temporal and spatial colony-site dynamics in a long-lived seabird. *Popul. Ecol.* **45**, 133–139 (2003).

Figures and Tables

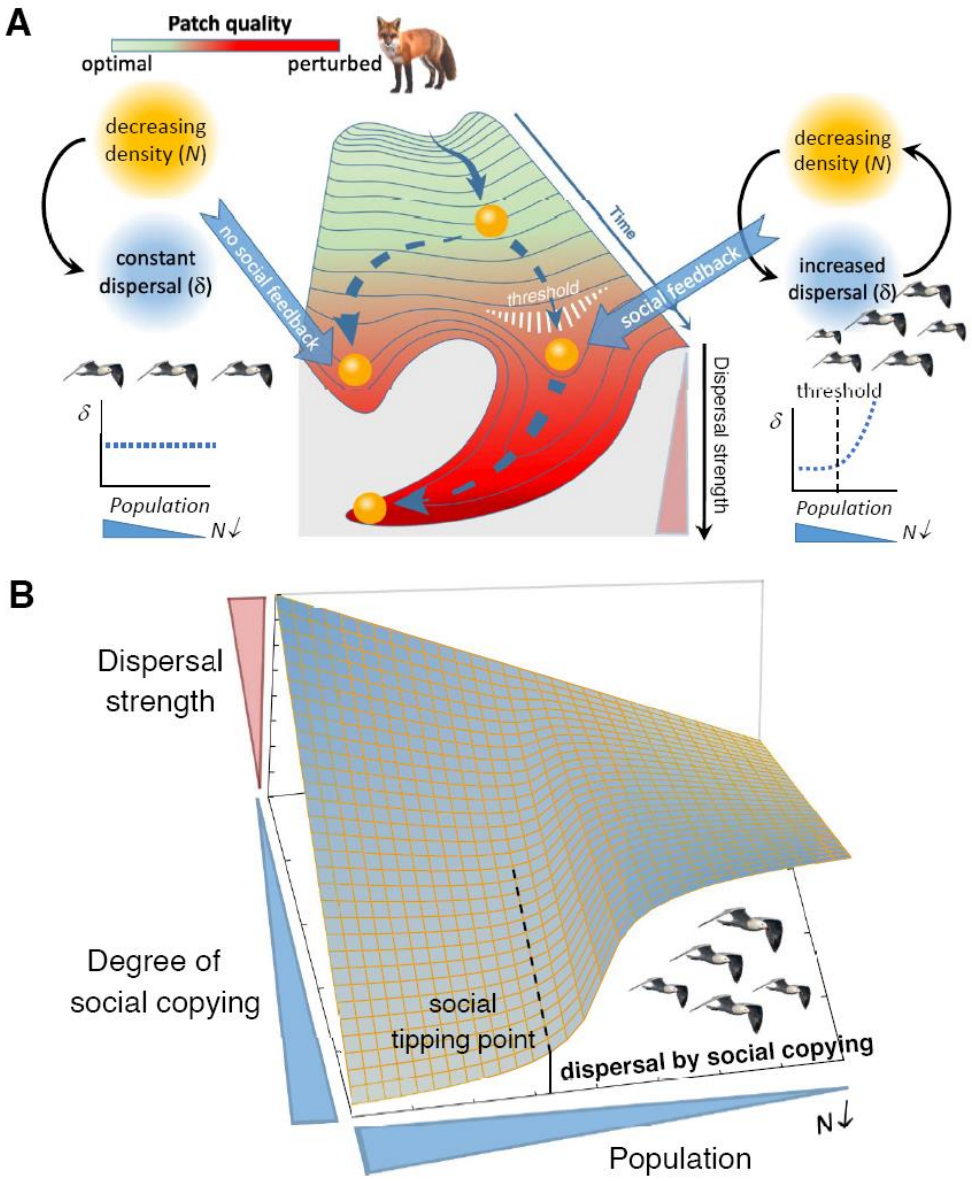


Figure 1. (*A*) Schematic diagram of dispersal represented as a dynamic landscape. Depending on the patch quality (here represented as a perturbed patch with the presence of predators), individuals may undergo different dispersal dynamics (represented by δ) e.g., density-independent or dispersal by social copying. For social copying, a low population threshold may be found, accelerating such a dispersal. (*B*) Dispersal strength as a function of population size and social copying feedback. Here, we tested whether dispersal made by a few individuals induces others to leave the patch in a behavioral cascade once a tipping point is exceeded. This social copying for dispersal should be mimicked by an inverse density-dependent function, where dispersal increases at a decreasing population.

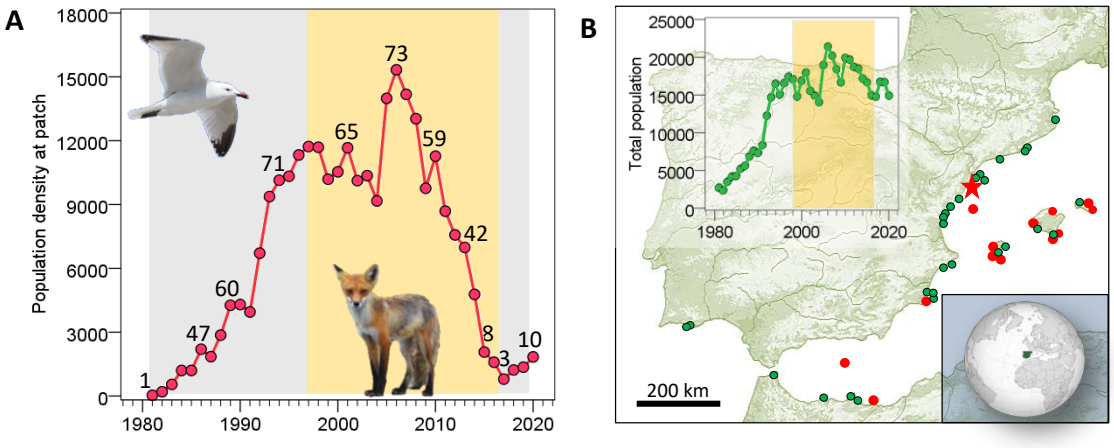


Figure 2. (*A*) Population dynamics of gulls breeding at La Banya since colonization (1981-2020), expressed as density of reproducing females. The colored area shows the duration of the perturbed regime caused by the arrival of carnivores. In the last years, population showed a slow recovery, once predators were removed for conservation purposes. Numbers over the line are percentages of total world population breeding here. (*B*) Spatial distribution of breeding patches, where 90% of the total world population breeds. Red start shows La Banya, red dots are the patches occupied before the perturbed regime, and green dots are the patches colonized since then. (*Inset*) Population dynamics of the whole metapopulation, which holds ca. 45 breeding patches. Fluctuations in metapopulation density were greatly influenced by the dynamics of La Banya, but they did not show the abrupt decline since dispersed individuals were redistributed over the colonized patches.

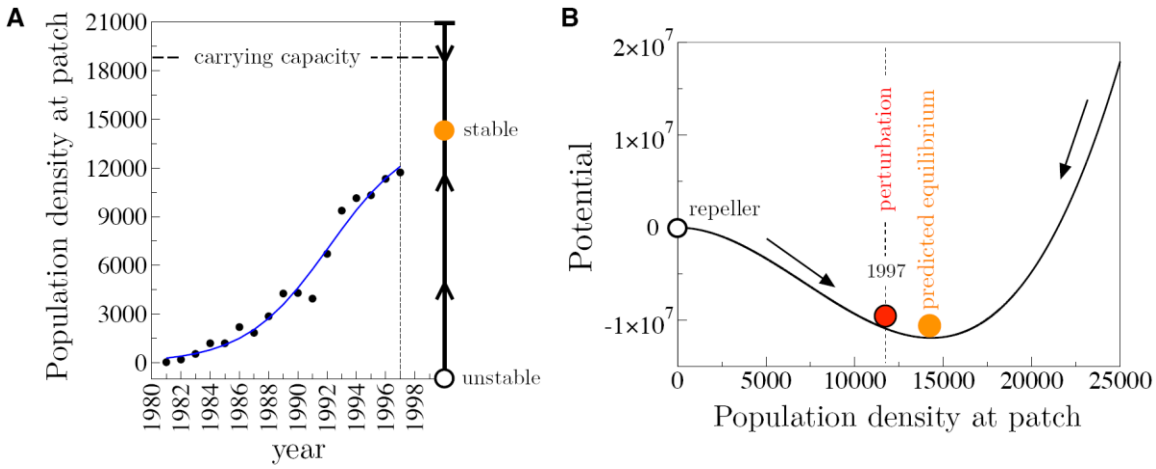


Figure 3. (*A*) Population dynamics from colonization in 1981 to the arrival of predators in 1997 and the fit of the logistic model (blue line, *SI Appendix,* section S3), which shows an exponential increase in density. The fitted parameters suggest that the population was approaching the expected equilibrium value when predators were established at the patch. The estimated carrying capacity is shown with a dashed line, together with the location of the equilibrium points (stable in blue, unstable in white) predicted by the estimated parameters (*SI Appendix*, figs. S3.1, S3.2, supplementary section S4). (*B*) Potential function of the logistic model showing the predicted equilibrium (orange marble), which was not achieved due to the perturbation starting in 1997 (red marble).

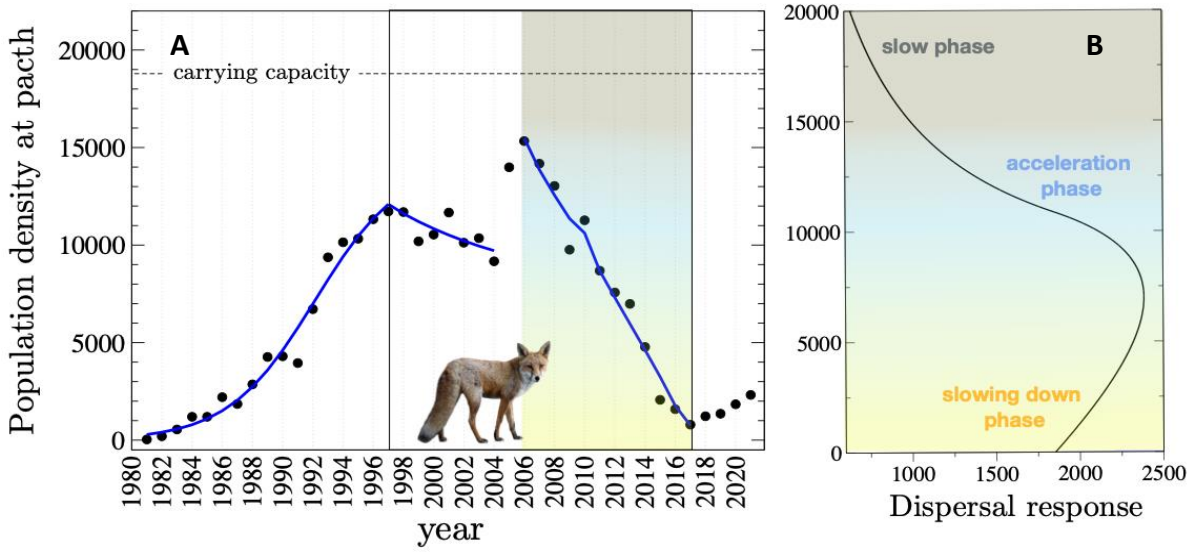


Figure 4. (*A*) The whole population dynamics since colonization to collapse (solid dots) and the fit of the model (blue line) for the three different phases. First, the population showed an exponential growth before the arrival of predators in 1997. Later, the field data showed a first slight decline after this arrival and a very abrupt decline from 2006 to 2017, when predators were removed. (*B*) Dispersal of individuals varied non-linearly with population density. Dispersal over the phase of abrupt decline showed three periods: an initial slow period (in grey), an acceleration with a social tipping point (in blue), and a slowing down phase (in orange).

Social copying drives non-linear population collapse in a colonial bird

Supplementary Information

Daniel Oro[1], Lluís Alsedà[2,5], Alan Hastings[3], Meritxell Genovart[4], and Josep Sardanyés[5]

[1]Theoretical and Computational Ecology Laboratory, Centre d'Estudis Avançats de Blanes (CEAB-CSIC), Cala Sant Francesc 14, 17300 Girona, Spain
[2]Departament de Matemàtiques, Edifici C, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès (Barcelona), Spain
[3]Department of Environmental Science and Policy, University of California, Davis, CA 95616, USA
[4]Theoretical and Computational Ecology Laboratory, Centre d'Estudis Avançats de Blanes (CEAB-CSIC), Cala Sant Francesc 14, 17300 Girona, Spain
[5]Centre de Recerca Matemàtica, Campus de Bellaterra, Edifici C, Universitat Autònoma de Barcelona, 08193 Cerdanyola del Vallès (Barcelona), Spain

July 31, 2022

Contents

1 Ecological data: Audouin's gull study population **3**
1.1 Study site and species . . . 3
1.2 Fieldwork methods and environmental data . . . 3
1.3 Population trends . . . 4

2 Mathematical modeling with dispersal by social copying **7**

3 Dynamics before the perturbation **9**
3.1 Model dynamics . . . 9

4 Model fitting and parameters estimation: Initial phase 1981–1997 **11**
4.1 On the positivity of structural parameters: analytical proof . . . 11
4.2 Estimation of the structural population parameters . . . 13

5 Dynamics after the perturbation: dispersal by social copying **15**
5.1 Modelling dispersal by social copying . . . 16
5.1.1 Properties of the dispersal function $\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{\mu}, \boldsymbol{\sigma}, \boldsymbol{\delta})$ . . . 16

6 Model fitting and parameters estimation: Collapse phase 2006–2017 **25**
6.1 A first approach to fit the collapse phase: a Sparse Anisotropic Grid Search . . . 25
6.2 Analytic and heuristic estimates of a compact domain that contains the optimum . . . 27
6.3 Fitting the collapse phase using artificial intelligence: Genetic Algorithms . . . 30
6.3.1 Population parameters and stopping criteria . . . 31
6.3.2 Individuals . . . 31
6.3.3 Selection with replacement . . . 33
6.3.4 Random initial population . . . 33
6.3.5 Mutation . . . 33
6.3.6 Crossover . . . 34
6.3.7 The Set of Superior variants of the Genetic Algorithm and its execution flow . . . 34
6.3.8 The results . . . 35

7 A change in the tendency of gulls' population increase at the onset of perturbation **37**

Supplementary Section 1

Ecological data: Audouin's gull study population

1.1 Study site and species

Long-term population monitoring of Audouin's gulls (*Ichthyaetus audouinii*) at Punta de la Banya (Ebro river Delta: 40°34′10.89″N, 0°39′34.28″E) started in 1981 when the patch was colonized and it has been continuously carried out until present during almost four decades (1981–2021). La Banya is a 2500ha sandy peninsula covered by halophilous vegetation and connected with the rest of the Ebro Delta by a 6km long narrow barren bar. The patch is at the mouth of the Ebro River and the continental shelf here is wide, which drives a high marine productivity and the overlap of high densities of both marine top predators and human fisheries [1, 2]. Audouin's gulls are long-lived social birds with a bet-hedging life history. They have evolved to cope with ephemeral habitats typical of Mediterranean marshes [3]. As a consequence, they have nomadic behaviour between breeding seasons: when patch conditions change and worsen then individuals are more prone to disperse mainly to other occupied sites [4], but in recent years and after the arrival of invasive carnivores, colonization rate of new patches has largely increased [5]. At La Banya, Audouin's gulls breed in sympatry with nine other species of the same ecological guild, including other gulls and terns. Dynamics and structure of this community is driven by competition governed by body size, with Audouin's gulls among the largest (i.e. dominant) species [6]. Demographic parameters of the species (by age and sex) have been estimated in a bunch of studies, including survival, recruitment curves, fertility and dispersal. Gulls breeding at La Banya show high adult survival (0.898 (SE : 0.01) mean adult survival probability), fast recruitment (birds start to breed when 3y old and most birds recruit before 5y old), and relative low fertility (average 0.471 fledglings per breeding female (SD : 0.287)) [4, 7, 8, 9, 10]. Audouin's gulls are specialized predators to feed on small pelagics at night, but they have learnt to exploit fish discarded by trawlers, which provides up to 70% of their diet by biomass. Despite the noise caused by the presence of carnivores and some extreme climatic events, trawling discards explains 24% of the variance in fertility over the years (phase 1991–2017).

1.2 Fieldwork methods and environmental data

At La Banya, gulls (both Audouin's gulls and their main competitor yellow-legged gulls) build their nest in clumped groups (what is called a sub-colony) [11]. Censuses are performed by teams of 2–15 people depending on the size of the sub-colony. Those people are organized in parallel band strips of 2-3 meters of width and each person counts nests with eggs at the right part of her band, with the spatial limit imposed by the person counting at her right. People moved in parallel to avoid double nest counts and missing nests. Censuses are carried out during the late incubation phase before hatching to avoid biases due to individuals still incorporating to the reproductive season. Additional fieldwork details are explained elsewhere and census errors were estimated and considered small and constant over the years ($< 5\%$) [12]. Several biotic and abiotic drivers can influence population fluctuations at the study patch. However, previous studies show that local biotic drivers explain better these fluctuations than global oceanographic indexes, such as the North Atlantic Oscillation index NAO [9]. Among these biotic drivers, interference competition with the dominant yellow-legged gull and predation and disturbance by invasive carnivores (mainly foxes) are the main factors affecting all crucial demographic parameters, namely adult survival, fertility and dispersal (both immigration and dispersal at spatial mesoscale). The main difference between these two drivers is that yellow-legged gulls are competitors with a long shared evolutionary history and long-term stability occurs when the two

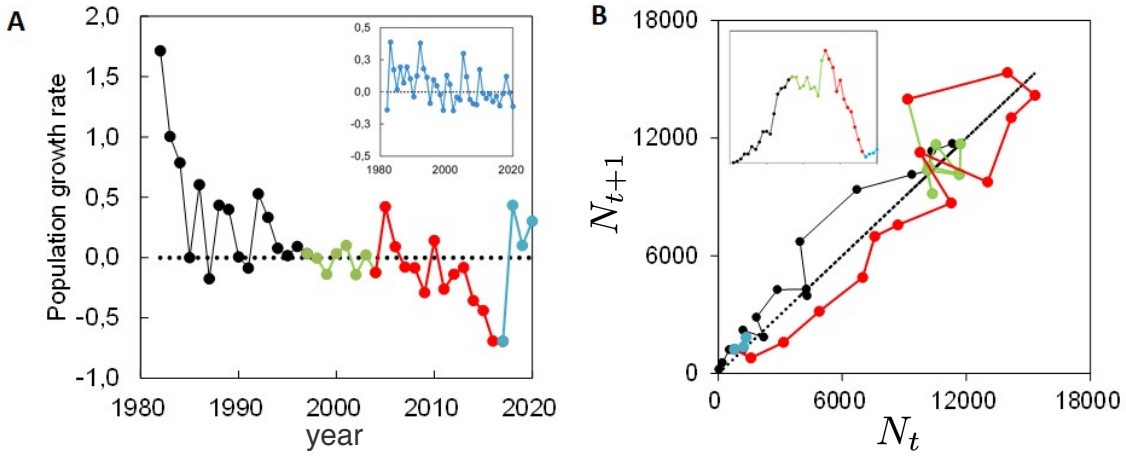


Figure S1.1: (A) Population growth rate since colonization of La Banya by Audouin's gulls in 1981; the dashed line shows no population growth. The inner panel shows population growth rate for the whole metapopulation (90% of total world population); black, green and red colors show the phases of exponential initial growth, dynamic stability, and non-linear decline respectively. (B) Ricker function of population density N at time t versus time $t+1$, with dashed line showing stability; colours as in panel (A); the inner panel shows how population density of Audouin's gulls varied at La Banya since colonization to 2020.

species occur in a specific patch. On the contrary, gulls have not developed evolutionary defenses to cope with terrestrial predators like carnivores, and this is why they select for breeding patches isolated and protected against the invasions of the predators.

Population density of yellow-legged gulls and the number of carnivores present at La Banya have been estimated over the years (Figure **S1.2** A,B respectively), and gull carcasses and tracks in the sand allowed us to estimate yearly predation rates that varied with the individual predator and its foraging preferences [5, 11, 13]. Other biotic factor is food availability, and a proxy to assess its temporal variability is through the statistics of landings of trawlers in the harbors close to the study site, which are highly correlated with the amounts of fish discarded [2, 8]. To account for the strength of density-dependence, this proxy of the changes in food availability was transformed as food per capita by dividing by the sum of the densities of Audouin's and yellow-legged gulls, the two more abundant and dominant species in the community. This density-dependence index explains much of the variance in fertility (see above and Figure **S1.2** C) and juvenile survival, whereas it did not correlate with changes in recruitment and adult survival [8, 9]. Food per capita decreased as population density approached the carrying capacity during the mid 90's and also because trawler catches per unit effort have decreased in recent decades due to overharvesting of fish stocks (Figure **S1.2** D). Adult survival, which is the vital rate with largest elasticity for the population dynamics of the gulls, changes with bycatch mortality at longline fisheries and by carnivore predation [10, 14].

Previous studies have shown that bycatch is relatively constant over the years [15], whereas carnivore density may vary with breeding season, although values were always low (median number of adult carnivores since their first arrival equaled two with range between zero and five) [11]. Predation rate increased with the density of carnivores, but some noise for this association occurred due to individual carnivore preferences for gull predation (Figure **S1.2** B). However, these predation rates did not significantly affect adult survival [4, 10], whereas they increased dispersal probabilities to other patches (either occupied or empty) [5, 11]. The number of colonized patches increased non-linearly since 2006 (Figure **S1.2** E), and metapopulation density followed parallel population dynamics with that at La Banya, except for the last years, when the slope at the former was slower than the slope to patch extinction at the later (Figure **S1.2** F). In summary, we did not record a decrease of food availability in absolute and per capita values (i.e. accounting for density-dependence), nor a decrease of local survival by carnivore predation or an increase of competition with the dominant yellow-legged gulls. Thus, these variables cannot explain the decline of population density of Audouin's gulls to patch collapse at La Banya since 2006, which should respond to an increase of dispersal to other patches, previously recorded using marked individuals and their field monitoring along most of the whole western Mediterranean metapopulation [5, 10, 11].

1.3 Population trends

The geometric mean of the population growth rate expressed as $\ln(N_{t+1}/N_t)$, being N_t the size of the population at time t, of Audouin's gulls since colonization directly estimated from the field data in 1981 to

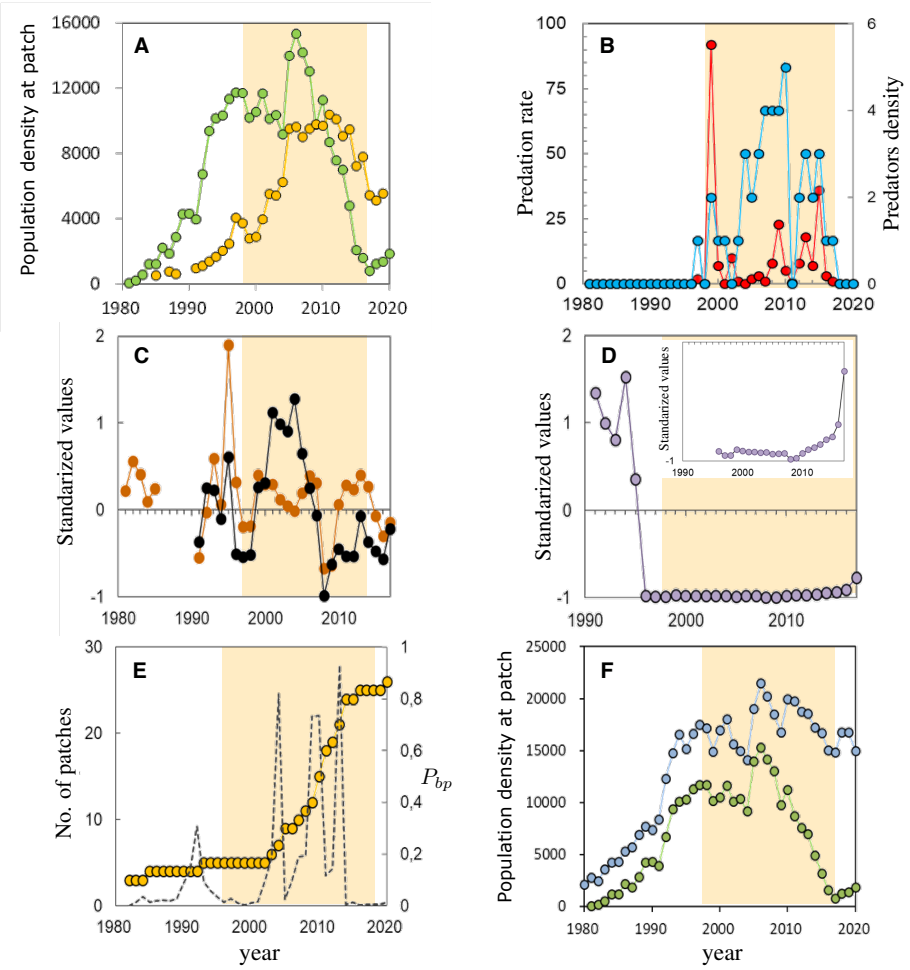


Figure S1.2: (A) Population density of yellow-legged (yellow circles) and Audouin's gulls (green circles), at la Banya since 1981, when the later species colonized the patch. (B) Variability of predation rate by carnivores (as number of corpses found preyed at the patch, red dots) and carnivore density (as number of adult carnivore present, blue dots). (C) Variability of trawling discards, as a proxy of food availability for gulls (maroon dots) and fertility (as mean number of chicks per breeding female, black dots) at La Banya. (D) Variability of food availability per capita during 1991–2017 and for the phase of population of Audouin's gull attaining the carrying capacity (inner panel). (E) Accumulated number of breeding patches occupied in the western Mediterranean and southern Portugal since 1981 (circles) and Bayesian probability of detecting a breaking point for this time series (dashed line). (F) Population density of Audouin's gulls at La Banya (green circles) from colonization in 1981 to 2020 and metapopulation density for this species including all patches in the western Mediterranean and southern Portugal (blue circles).

2017 was 0.086 (Figure **S1.1**). Annual gulls' mortality rate computed from long-term monitoring data (1988-2015) using capture-recapture modeling was estimated to be 0.11 year^{-1}[10]. Four well-defined periods can be distinguished when looking at the time series of breeding gulls, which are shown with vertical dashed lines in the time series of Figure **S1.3**. First, an *initial phase* with exponential population growth partly explained by high immigration rates from the outside [12]. Following this period, the population kept growing until 1997, when predators entered into the patch. The second period, labelled *onset of perturbation phase*, spans from 1998 to 2006. Figure **S1.2** B displays the density of predators (blue dots) during the whole period of study, showing the presence of few predator individuals between 1997 and 2017, with a maximum of 5 individuals in 2010 and absence of carnivores from 2017 to 2021. The predation rate is also shown in Figure **S1.2**B with red dots. Here, predation rate is the percentage of corpses predated by the carnivores with respect to the total number of corpses found each year (see [5, 11] for details). The population of gulls suffered a large increase around 2005–2006 due to an increase in food availability per capita (see also Figure **S1.2** C,D). Since 2006, gull's population started a sustained and sharp decline until the patch held only 3% of total world population in 2017 (see main text), coinciding with the absence of predators in the patch (Figure **S1.2** B). This period from 2006 to 2017 will be denoted as *collapse phase*. Finally, the period from 2017 to 2021, where the gulls' population started increasing again, coinciding with the absence of predators. In this manuscript we will focus on the dynamics between 1981 and 2017.

136 The field data for the Audouin's gulls at Punta de la Banya during the period of study is shown in the table and the plot below (see also Figure **S1.2** A, green dots):

year	pop.	year	pop.	year	pop.	year	pop.
1981	36	**1991**	3950	**2001**	11666	**2011**	8688
1982	200	**1992**	6174	**2002**	10122	**2012**	7571
1983	546	**1993**	9373	**2003**	10355	**2013**	6983
1984	1200	**1994**	10143	**2004**	9168	**2014**	4778
1985	1200	**1995**	10327	**2005**	13988	**2015**	2067
1986	2200	**1996**	11328	**2006**	15329	**2016**	1586
1987	1850	**1997**	11725	**2007**	14177	**2017**	793
1988	2861	**1998**	11691	**2008**	13031	**2018**	1225
1989	4266	**1999**	10189	**2009**	9762	**2019**	1355
1990	4300	**2000**	10537	**2010**	11271	**2020**	1837
						2021	2319

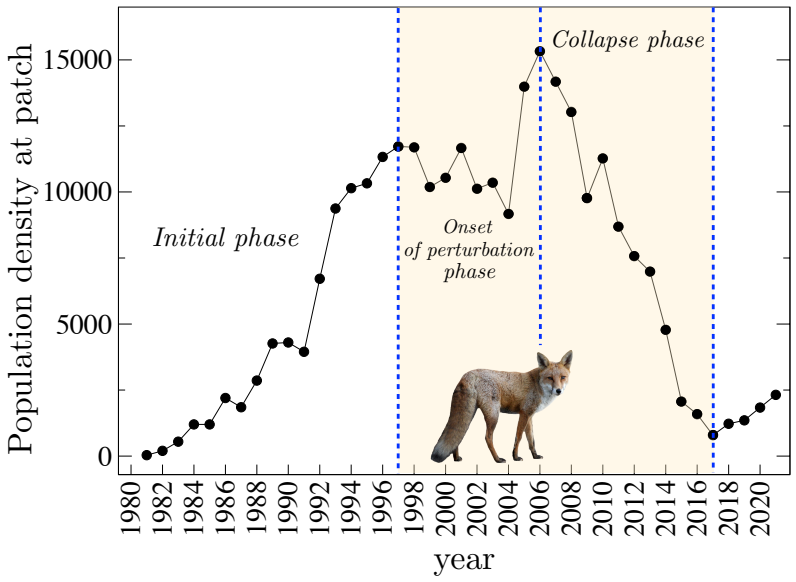


Figure S1.3: The Audouin's population data at La Banya from 1981 to 2021. The period 1981–1997, labelled as *initial phase*, was characterised by a logistic growth due to the absence of predators and the fact that the population did not approach the expected equilibrium (see Section S4.2 below). Predators (foxes) colonized the patch in 1997, causing a qualitative change in the dynamics and a decreasing tendency in the population until 2004, with a large fluctuation in 2005-2006. The period 2006–2017, labelled as *collapse phase*, was characterised by a fast population decline due to dispersal until 2017, when predators were not found at all in the patch. Notice that after the absence of predators from 2017 onwards, the population of birds started increasing again (2018-2021).

137

Supplementary Section 2

Mathematical modeling with dispersal by social copying

In this section we introduce the mathematical model used to investigate the population dynamics of Audouin's gulls in the patch of study. The model describes the population dynamics of the birds as a single-patch system considering immigration and dispersal of individuals. Other modelled ecological processes are birds' intra-specific competition for resources and density-independent death. As we thoroughly explain below, the model incorporates dispersal dynamics considered in the last two terms at the right hand side of Equation (2.1). Let us denote the birds' population size at time t by $x(t)$. Then, the model reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = (\vartheta+\omega)x(t)\left(1-\frac{x(t)}{K}\right) - \varepsilon x(t) - \Big[\rho x(t) + \lambda D\big(x(t)\big)\Big] \tag{2.1}$$

$$= \underbrace{\boxed{\varphi x(t)}}_{\substack{\text{Immigration,}\\ \text{growth and}\\ \text{death}}} - \underbrace{\boxed{\beta x(t)^2}}_{\substack{\text{Nonlinear}\\ \text{competition}\\ \text{term}}} - \underbrace{\boxed{\lambda D(x(t))}}_{\substack{\text{Dispersal by}\\ \text{social}\\ \text{copying}}}$$

where the parameters are described in Table 2.1 in the next page.

Equation (2.1) considers an initial exponential increase of the population proportional to parameters $\vartheta+\omega$, including both the reproduction of birds (ϑ) and the immigration rate (ω) of new individuals from other patches of the metapopulation (not explicitly considered) to the patch of study. Notice that, for simplicity, we have lumped these two parameters by setting $\gamma = \vartheta + \omega$. The population growth is constrained by a logistic function with carrying capacity K, introducing intra-specific competition for resources. The competition term can also be expressed as $\beta x(t)^2$, with $\beta = \gamma/K$. The death rate is fixed to $\varepsilon = 0.11$ corresponding to the annual mortality rate estimated from long-term monitoring (1988-2015) using capture-recapture modeling [10]. Finally, two terms related to dispersal are included. Exponential dispersal proportional to constant ρ and the function $D(x(t))$ that will be used to introduce dispersal by social copying. This dispersal term mimicking social copying will generically consider an inverse, density-dependent departure of the birds from the patch. That is, the less number of birds at the patch the higher their departure (see Section S5).

The model in compact form is shown framed with different colours: the blue box displays all of the processes related to population growth, including reproduction, immigration, death, and exponential dispersal of individuals now with $\varphi = \gamma - \varepsilon - \rho$. The green box corresponds to the intra-specific competition, while the dispersal term including social copying is represented within the red box. The model will be used to investigate the field data, focusing on three phases. We call the period 1981–1997, which corresponds to the establishment of the local population before the arrival of predators (i.e., before the perturbation), the *initial phase*. It is characterised by a logistic growth due to the absence of predators and the fact that the population did not exhaust the food carrying capacity. Here, we will not consider dispersal by social copying triggered by the perturbation since predators were not present in the patch during this period. The second period, labeled *onset of perturbation phase*, will incorporate the hypothesized social copying dynamics triggered by the arrival of predators in the patch, and will last till the beginning of the collapse phase starting in 2006. This phase is characterised by a change in the growing tendency of the population in 1998 and a sustained decrease until 2004, with a very large increase in the years 2005-2006. Despite the available data for this period is scarce, we will estimate the tendency of the population density fitting the model under different assumptions (Section S7). Finally, we will investigate the dynamics from 2006–2017 (*collapse phase*), when predators were still present at the patch and the gulls' population experienced the collapse. Here, we will consider the full

Parameter	Units	Range or value	Ecological meaning or description
ϑ	year^{-1}	$[0, +\infty)$	Intrinsic reproduction rate
ω	year^{-1}	$[0, +\infty)$	Rate of entry of individuals from other patches
K	birds	$[1, +\infty)$	Carrying capacity
ε	year^{-1}	0.11	Death rate estimated from field data [10]
ρ	year^{-1}	$\mathbb{R}^+$	Linear (exponential) dispersal rate
$\gamma = \vartheta + \omega$	year^{-1}	$[0, +\infty)$	Population growth rate due to reproduction and immigration
$\alpha = \gamma - \varepsilon$	year^{-1}	$(-\infty, \gamma]$	Neat population growth rate without linear dispersal
$x(0)$	birds	$[0, K]$	Initial condition of Eq. (2.1)
$\varphi = \alpha - \rho$	year^{-1}	$(-\infty, \alpha]$	Population growth rate including linear dispersal
$\beta = \frac{\gamma}{K}$	$(\text{birds} \times \text{year})^{-1}$	$[0, +\infty)$	Intrinsic growth rate over the carrying capacity
λ	year^{-1}	$\mathbb{R}^+$	Dispersal rate by social copying

Table 2.1: Model parameters for the general model used to investigate the local dynamics of Audouin gulls at la Punta de la Banya from 1981 to 2017.

model given by Equation (2.1). That is, considering dispersal terms. The analytic and qualitative study of the model for these three phases and the corresponding fitting of parameters will be done in the next sections.

Supplementary Section 3

Dynamics before the perturbation

3.1 Model dynamics

In this section we study the model given by Equation (2.1) in the initial phase, ranging from the establishment of the population at the patch of study in 1981 until the arrival of predators in 1997. The model given by Equation (2.1) to study the initial phase will not include dispersal by social copying since we hypothesize that this behavioural dispersal is triggered by the presence of predators ($\lambda = 0$). Moreover, we will also assume no linear dispersal from the study patch to other patches of the metapopulaiton ($\rho = 0$) since, as we discussed in Section 1.2 above, the initial phase was dominated by high immigration rates from outside [12]. Under these considerations, we get

$$\frac{dx(t)}{dt} = \alpha x(t) - \beta x(t)^2. \tag{3.1}$$

Equation (3.1) is a particular case of a Ricatti Equation with constant coefficients. Its closed analytical solution is obtained in the next lemma by integrating the model as an equation of separable variables.

Lemma 3.1 (A Ricatti Equation with constant coefficients)**.** *According to table 2.1 above we know that β must be non-negative. Then, the solution $x(t)$ of Model* (3.1) *is the following:*

If $\alpha \neq 0$,

$$x(t) = \frac{\alpha x(0)\exp(\alpha t)}{\alpha + \beta x(0)(\exp(\alpha t) - 1)} = \frac{\alpha x(0)}{\alpha \exp(-\alpha t) + \beta x(0)(1 - \exp(-\alpha t))};$$

and if $\alpha = 0$,

$$x(t) = \frac{x(0)}{x(0)\beta t + 1}.$$

The dynamics of Model (3.1) is well-known. However, for the sake of completeness we here analyse its qualitative dynamics for the case $\alpha, \beta > 0$. This is indeed not a restrictive assumption since as we will see below the observed data is only compatible with the positivity of parameters α and β.

The proof of the next lemma is a simple exercise (see Figure **S3.1**).

Lemma 3.2. *Assume that $\alpha, \beta > 0$. Then the function $f(x) = x(\alpha - \beta x)$ verifies: $f(0) = f\left(\frac{\alpha}{\beta}\right) = 0$, $f(K) = -K\varepsilon \leq 0$, and has a unique critical point at $x = \frac{\alpha}{2\beta}$. Hence, f is unimodal with $f\big|_{\left(0, \frac{\alpha}{2\beta}\right]}$ strictly increasing and positive, $f\big|_{\left[\frac{\alpha}{2\beta}, K\right]}$ strictly decreasing, $f\big|_{\left[\frac{\alpha}{2\beta}, \frac{\alpha}{\beta}\right)}$ positive and $f\big|_{\left(\frac{\alpha}{\beta}, K\right]}$ negative. Additionally, $0 < \frac{\alpha}{2\beta} < \frac{\alpha}{\beta} \leq K$, and $\frac{\alpha}{\beta} = K$ if and only if $\varepsilon = 0$.*

A consequence of Lemma 3.2 is that model (3.1) has two stationary solutions computed from $\dot{x} = 0$. They are $x(t) = 0$ and $x(t) = \frac{\alpha}{\beta}$ when $t \to +\infty$. Equilibrium 0 (labeled $x_0^\star$) involves, whenever stable, no population at the patch, while equilibrium $\frac{\alpha}{\beta}$ (labeled $x_1^\star$) will involve, provided is stable, the persistence of the population. Generically, the (local) stability of a given equilibrium solution $x(t \to +\infty) = x^\star$ of a one-variable differential equation $\frac{dx(t)}{dt} = f(x(t))$ can be computed from the sign of $\frac{df(x)}{dx}\big|_{x=x^\star}$. More precisely, the equilibrium is a local attractor when $\frac{df(x)}{dx}\big|_{x=x^\star} < 0$ or unstable when $\frac{df(x)}{dx}\big|_{x=x^\star} > 0$. From the previous expressions we obtain

$$\left.\frac{df(x)}{dx}\right|_{x=x^\star} = \alpha - 2\beta x^\star = \begin{cases} \alpha & \text{if } x^\star = x_0^\star = 0, \text{ and} \\ -\alpha & \text{if } x^\star = x_1^\star = \frac{\alpha}{\beta}. \end{cases}$$

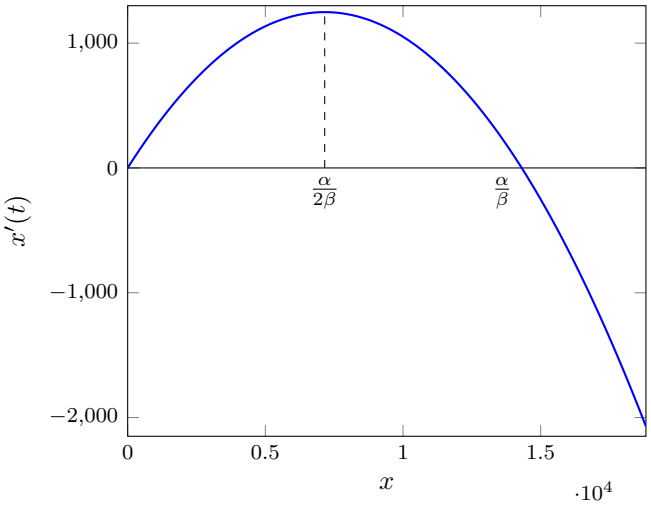


Figure S3.1: The vector field (3.1) for $\alpha = 0.348949408396403$, $\beta = 0.0000243826353653$ and $K = 18822.8$.

Hence, for $\alpha > 0$ (intrinsic growth rate of the population larger than decline rate $x_0^\star$ is unstable and $x_1^\star$ stable. We notice that for $\alpha = 0$ there exists a transcritical bifurcation involving a collision and an exchange of stability between the two equilibria. As expected, for $\alpha < 0$ (decline rate larger than population growth rate) the stable equilibrium corresponds to $x_0^\star = 0$. This behaviour is illustrated in Figure **S3.2**A as a function of α. As we will see in Section 4, the case of interest is given by $\alpha > 0$, value obtained for the field data.

Finally, an illustrative way of visualizing the stability of a dynamical system with one variable is to compute the so-called potential function, given by:

$$U(x) = -\int f(x)\,dx, \qquad \text{with} \quad U(x) = x^2\left(\frac{\beta x}{3} - \frac{\alpha}{2}\right) \text{ for Eq. (3.1).} \tag{3.2}$$

Figure **S3.2**B displays three potential functions computed from Equation (3.1) for different values of $\alpha > 0$. Specifically, the field data for the initial phase reveals that the equilibrium of the population was not when predators colonised the patch, thus being in a transient state (see Figure 3B in the main manuscript and Section S 4).

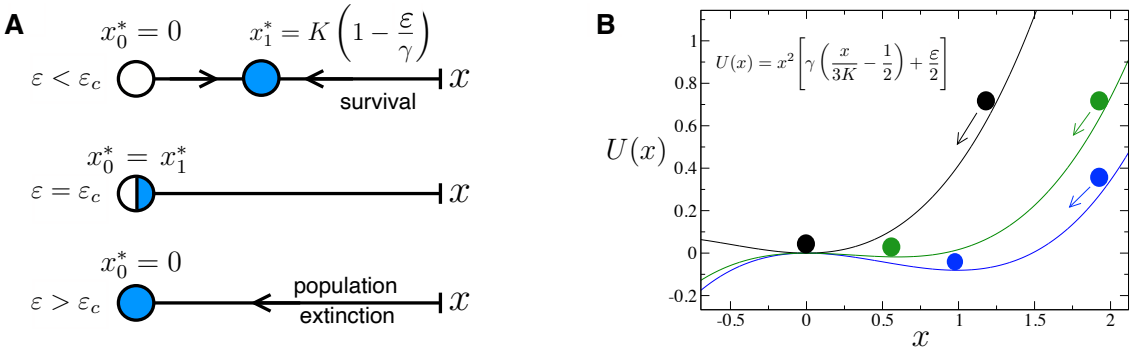


Figure S3.2: (A) Schematic diagram of the dynamics of Equation (3.1) in the phase space as a function of model parameters. Here three possible scenarios are found: $\alpha < 0$ with population vanishment (x_0^* stable and x_1^* unstable); $\alpha = 0$, bifurcation value at which the transcritical bifurcation occurs; and $\alpha > 0$, with persistence of the population (x_1^* stable and x_0^* unstable). Stable and unstable points are indicated with blue and white marbles. (B) Potential function given by Equation (3.2) computed with $\alpha = 0.4995$ (blue); $\alpha = 0.3$ (green); $\alpha = -0.495$ (black). Here we use $K = 1$.

Supplementary Section 4

Model fitting and parameters estimation: Initial phase 1981–1997

4.1 On the positivity of structural parameters: analytical proof

For the study and fitting of Model (3.1) in the initial phase (Figure **S1.3**) we need to introduce some appropriate notation. The observed population of Audouin's gulls at the years 1981 to 1997 will be denoted by

$$\eta(t, t=0:16) = \mathsf{Audouin's\_Gulls\_Observed\_Population\_at\_year}(1981+t, t=0:16) =$$
$$\underset{[0]}{36}, \underset{[1]}{200}, \underset{[2]}{546}, \underset{[3]}{1200}, \underset{[4]}{1200}, \underset{[5]}{2200}, \underset{[6]}{1850}, \underset{[7]}{2861}, \underset{[8]}{4266}, \underset{[9]}{4300}, \underset{[10]}{3950}, \underset{[11]}{6714}, \underset{[12]}{9373}, \underset{[13]}{10143}, \underset{[14]}{10327}, \underset{[15]}{11328}, \underset{[16]}{11725}\big].$$

The solution of the above Model (3.1) with $\beta \geq 0$ and initial condition $\kappa \in \mathbb{R}^+$ will be denoted by $x(t) = x_{\kappa,\alpha,\beta}(t)$. Observe that $\kappa = x_{\kappa,\alpha,\beta}(0)$ must be considered a free parameter as well.

Now we define the parameter space

$$\mathscr{F} := \mathbb{R}^+ \times \big\{(\beta K - \varepsilon, \beta) : \beta \in \mathbb{R}^+\big\}$$

(recall that $\alpha = \gamma - \varepsilon = \beta K - \varepsilon \in (-\varepsilon, \infty)$), and a map

$$\begin{array}{rccc} \mathsf{L}\colon & \mathscr{F} & \longrightarrow & \mathbb{R}^+ \\ & \big(\kappa, (\alpha,\beta)\big) & \longmapsto & \sqrt{\sum_{t=0}^{16}\big(x_{\kappa,\alpha,\beta}(t) - \eta(t)\big)^2}. \end{array}$$

The map L measures the agreement between the solution of Model (3.1) with initial condition κ and parameters α and β, and the observed data $\eta(t, t=0:16)$, through the Euclidean norm

$$\sqrt{\sum_{t=0}^{16}\big(x_{\kappa,\alpha,\beta}(t) - \eta(t)\big)^2}.$$

Observe that the map $\big(\kappa, (\alpha,\beta)\big) \longmapsto \sqrt{\sum_{t=0}^{16}\big(x_{\kappa,\alpha,\beta}(t) - \eta(t)\big)^2}$ can be decomposed in two steps:

$$\big(\kappa, (\alpha,\beta)\big) \longmapsto x_{\kappa,\alpha,\beta}(t, t=0:16) \longmapsto \sqrt{\sum_{t=0}^{16}\big(x_{\kappa,\alpha,\beta}(t) - \eta(t)\big)^2}.$$

The first step is computed with the help of Lemma 3.1, taking into account whether $\alpha < 0$; $\alpha = 0$ or $\alpha > 0$.

Of course, if the dynamics of the Audouin's gulls population size during the years 1981 to 1997 is governed by some instance of Model (3.1) with parameters $x(0) = \kappa^*$, $\alpha = \alpha^*$ and $\beta = \beta^*$, then the value of $\mathsf{L}\big(\kappa^*, (\alpha^*, \beta^*)\big)$ must be small and likely it must correspond to

$$\begin{array}{c} \min \mathsf{L}\big(\kappa, (\alpha,\beta)\big) \\ \text{subject to } \big(\kappa, (\alpha,\beta)\big) \in \mathscr{F}, \\ x(0) = \kappa \geq 0, \\ \text{and } x(t) \geq 0 \text{ for } t = 1, 2, \ldots, 16. \end{array} \tag{4.1}$$

The solution of this problem is called the *fitting of the model* and identifies a valid analytical model for the dynamics of the Audouin's gulls at the years 1981 to 1997 (of course provided that the value $\min \mathsf{L}\big(\kappa,(\alpha,\beta)\big)$ is small). Observe that the set

$$\Big\{\mathsf{L}\big(\kappa,(\alpha,\beta)\big) : \big(\kappa,(\alpha,\beta)\big) \in \mathscr{F}; x_{\kappa,\alpha,\beta}(0)=\kappa \text{ and } x_{\kappa,\alpha,\beta}(t)\geq 0 \text{ for } t=1,2,\dots,16\Big\} \subset \mathbb{R}^+$$

has 0 as a lower bound. Hence, it has a minimum element, and Problem (4.1) has at least one solution.

Next we consider a reduced (and better) parameter space

$$\mathscr{O} := \mathbb{R}^+ \times \Big\{\big(\alpha, \tfrac{\alpha+\varepsilon}{K}\big) : \alpha \in (0,\infty)\Big\} \subset \mathscr{F}$$

(here we use again that $\beta = \frac{\gamma}{K} = \frac{\alpha+\varepsilon}{K}$), and the associated reduced optimization problem becomes:

$$\begin{aligned} &\min \mathsf{L}\big(\kappa,(\alpha,\beta)\big)\\ &\text{subject to } \big(\kappa,(\alpha,\beta)\big)\in\mathscr{O},\\ &\qquad x(0)=\kappa\geq 0,\\ &\qquad \text{and } x(t)\geq 0 \text{ for } t=1,2,\dots,16. \end{aligned} \tag{4.2}$$

The next lemma reduces the search space to find the optimum fit of the model and, since $\alpha>0$ whenever $\big(\kappa,(\alpha,\beta)\big)\in\mathscr{O}$, Lemma 3.1 tells us that in the first step of the computation of the map L we have

$$x_{\kappa,\alpha,\beta}(t) = \frac{\alpha x(0)}{\alpha\exp(-\alpha t)+\beta x(0)(1-\exp(-\alpha t))}.$$

Lemma 4.1. *The solutions of Problem* (4.1) *and Problem* (4.2) *coincide.*

Proof. We have to see that for every $\big(\kappa,(\alpha,\beta)\big)\in\mathscr{F}\setminus\mathscr{O}$ there exist $\big(\widetilde{\kappa},(\widetilde{\alpha},\widetilde{\beta})\big)\in\mathscr{O}$ such that

$$\mathsf{L}\big(\widetilde{\kappa},(\widetilde{\alpha},\widetilde{\beta})\big) < \mathsf{L}\big(\kappa,(\alpha,\beta)\big).$$

The following is obtained by direct computation: $\big(200,\big(0.3,10^{-5}\big)\big)\in\mathscr{O}$, and

$$\mathsf{L}\big(200,\big(0.3,10^{-5}\big)\big) = 6317.69\cdots.$$

On the other hand, for every $\big(\kappa,(\alpha,\beta)\big)\in\mathscr{F}\setminus\mathscr{O}$ we have $\alpha\leq 0$ (in fact $\alpha\in[-\varepsilon,0]$). Since β is non-negative, this implies $x'(t)=x(t)\big(\alpha-\beta x(t)\big)\leq 0$ because $x(t)\geq 0$ for every t. Then, for every $t\geq 0$, by the Mean Value Theorem, there exists a $\xi\in(t,t+1)$ such that $x(t+1)=x(t)+x'(\xi)\leq x(t)$. Consequently,

$$\begin{aligned} x_{\kappa,\alpha,\beta}(0)\geq x_{\kappa,\alpha,\beta}(1)\geq\cdots\geq x_{\kappa,\alpha,\beta}(6)\geq x_{\kappa,\alpha,\beta}(7)\geq x_{\kappa,\alpha,\beta}(8)\geq\\ x_{\kappa,\alpha,\beta}(9)\geq x_{\kappa,\alpha,\beta}(10)\geq x_{\kappa,\alpha,\beta}(11)\geq\cdots\geq x_{\kappa,\alpha,\beta}(16). \end{aligned}$$

Assume first that $3950=\eta(10)\leq x_{\kappa,\alpha,\beta}(7)$. Then

$$\eta(t)<\eta(10)\leq x_{\kappa,\alpha,\beta}(7)\leq x_{\kappa,\alpha,\beta}(t) \quad \text{for } t=0,1,\dots,7.$$

Then,

$$\mathsf{L}\big(\kappa,(\alpha,\beta)\big)\geq\sqrt{\sum_{t=0}^{7}\big(x_{\kappa,\alpha,\beta}(t)-\eta(t)\big)^2}\geq\sqrt{\sum_{t=0}^{7}\big(\eta(10)-\eta(t)\big)^2}=8046.89\cdots>\mathsf{L}\big(200,\big(0.3,10^{-5}\big)\big).$$

This ends the proof of the lemma in this case.

Now assume that $3950=\eta(10)\geq x_{\kappa,\alpha,\beta}(7)$. In this case, $\eta(t)\geq\eta(10)\geq x_{\kappa,\alpha,\beta}(7)\geq x_{\kappa,\alpha,\beta}(t)$ for $t=8,9,\dots,16$. Hence,

$$\mathsf{L}\big(\kappa,\alpha,\beta\big)\geq\sqrt{\sum_{t=8}^{16}\big(\eta(t)-x_{\kappa,\alpha,\beta}(t)\big)^2}\geq\sqrt{\sum_{t=8}^{16}\big(\eta(t)-\eta(10)\big)^2}=15204.468\cdots>\mathsf{L}\big(200,\big(0.3,10^{-5}\big)\big).$$

□

4.2 Estimation of the structural population parameters

Here, we estimate the structural population parameters better explaining the dynamics of the initial phase. These include the initial condition, κ, and parameters α and β (see Table 2.1), taking the value of $\varepsilon = 0.11$ estimated from the field data [10]. In order to obtain the structural parameters of the population we need to focus on the local dynamics before the perturbation and not considering external perturbations. That is, focusing on the period 1981-1997. From the discussion in Section 3.1 (see Lemma 4.1) we have to solve Problem (4.2). To do so, we have used a standard *trust region method* and also the Levenberg-Marquardt algorithm to solve the trust region sub-problem (see the GNU Scientific Library (GSL) *Nonlinear Least-Squares Fitting* documentation). As it can be guessed from the last reference we have used a GSL standard library function for this computation, with numerical approximation of derivatives of the objective function. The obtained results are the following:

Parameters:

$$\kappa = x_{\kappa,\alpha,\beta}(0) = 288.04096 \pm 117.9663$$
$$\alpha = 0.348949408396403 \pm 0.04958259$$
$$\beta = 0.0000243826353653 \pm 0.00000598145$$
$$\varepsilon = 0.11 \text{ (estimated directly from original data [10])}$$
$$\gamma = \alpha + \varepsilon = 0.458949408396403\cdots$$
$$K = \gamma/\beta = 18822.7975\cdots$$

Error measures:

$$\mathsf{L}\big(\kappa, (\alpha, \beta)\big) = 2593.0536\cdots$$

Year	Population Data	
	Observed	Predicted
1981	36	288.040965
1982	200	404.917274
1983	546	567.299154
1984	1200	791.095773
1985	1200	1096.137858
1986	2200	1505.703290
1987	1850	2044.623822
1988	2861	2735.234085
1989	4266	3590.827289
1990	4300	4607.530642
1991	3950	5757.502035
1992	6714	6987.807435
1993	9373	8228.125049
1994	10143	9405.847993
1995	10327	10462.226531
1996	11328	11362.442042
1997	11725	12096.688873

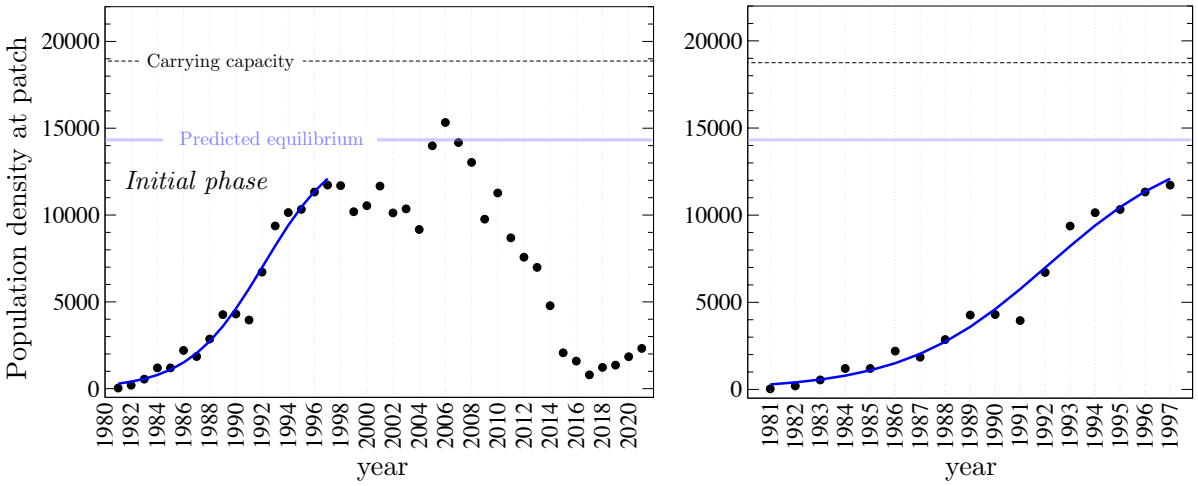


Figure S4.1: (Left) Fitting of the initial phase from the establishment of the local population at La Banya since the arrival of the predators (1981-1997) using Eq. (3.1). The blue line shows the best fit obtained with the optimization method (least-squares $\mathsf{L}\big(\kappa, (\alpha, \beta)\big) = 2593.0536$), which has given the parameter values shown above. The horizontal dashed line shows the predicted carrying capacity, while the horizontal blue line shows the predicted equilibrium point of the population. (Right) Zoom in the period 1981-1997.

Figure **S4.1** (left) displays the dynamics of the local population for the estimated parameter values (blue line). The predicted equilibrium, computed from

$$x_1^\star = \tfrac{\alpha}{\beta} = 14312.936833$$

is shown with a horizontal blue line, while the carrying capacity is shown with a horizontal dashed line. The model predictions suggest that the population have not reached the steady state on the onset of the perturbation, and that the large population increase suffered in 2005-2006 did not surpass the carrying capacity. Figure **S4.1** (right) shows an enlarged view of the dynamics obtained with the parameters best fitting the initial phase data.

Supplementary Section 5

Dynamics after the perturbation: dispersal by social copying

Here, we introduce and investigate the dynamics of the population with the model including linear dispersal and dispersal by social copying. The model reads:

$$\frac{\mathrm{d}}{\mathrm{d}t}x(t) = \varphi x(t) - \beta x(t)^2 - \lambda D(x(t), \mu, \sigma, \delta) \tag{5.1}$$

with the following parameters.

Parameter	Range or value	Ecological meaning or description
ε	0.11	Death rate estimated from field data [10]
α	0.348949408396403	Population growth rate including death of individuals (without linear dispersal)
$\gamma = \alpha + \varepsilon$	0.458949408396403	See Table 2.1 for details
$K = \frac{\gamma}{\beta}$	18822.7975	Carrying capacity
$x(0)$	$[0, K]$	Initial condition
ρ	$\mathbb{R}^+$	Linear (exponential) dispersal rate
$\varphi = \alpha - \rho$	$(-\infty, \alpha]$	Neat population growth rate including linear dispersal
β	$2.43826353653 \times 10^{-5}$	Intrinsic growth rate over the carrying capacity
Parameters concerning dispersal rate by social copying		
λ	$\mathbb{R}^+$	Dispersal rate
μ	$\mathbb{R}^+$	Tendency of dispersal function for small population sizes
σ	$\mathbb{R}^+$	Sharpness and smoothness of the dispersal function
δ	$\mathbb{R}^+$	Transition between small and large population sizes

Next, we introduce the chosen dispersal function for Model (2.1) able to mimic dispersal by social copying. We have chosen the so-called Elliot function which includes different parameters which allow to obtain multitude of different shapes. The key point for choosing this function is that it typically increases at decreasing population values. Generically, the less the population at the patch, the largest value for this function and thus the higher dispersal rates. However, due to its plasticity, other behaviours can be found showing this increasing tendency at low population numbers in certain parts. For instance, some values of the function can decrease at low population values; for some other parameters one can obtain density-independent dispersal i.e., constant dispersal independent of the population density (e.g., Fig. **S5.1** for $\sigma = 0.001$); as well as different sigmoidal shapes.

5.1 Modelling dispersal by social copying

The nonlinear dispersal function that we propose for Model (5.1) (and model (2.1)) is given by:

$$D(x,\mu,\sigma,\delta) := \begin{cases} \frac{1-\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)}{1-\mathcal{E}_{\mathsf{dir}}(0,\mu,\sigma,\delta)} & \text{when } 0 \le x \le \delta, \\ \frac{1-\mathcal{E}(x,\sigma,\delta)}{1-\mathcal{E}_{\mathsf{dir}}(0,\mu,\sigma,\delta)} & \text{when } x \ge \delta, \end{cases} \tag{5.2}$$

where

$$\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta) := \left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right)+\frac{x}{\delta}\right)\mathcal{E}(x,\sigma,\delta), \tag{5.3}$$

and

$$\mathcal{E}(x,\sigma,\delta) := \frac{\sigma(x-\delta)}{\Theta+\sigma|x-\delta|}, \tag{5.4}$$

is an *Elliot sigmoid* Θ–scaled, σ–strengthened, and δ–displaced. All the parameters of the dispersal function are non-negative and we have fixed $\Theta := 1000$ (this parameter controls the scale in the independent variable x which is related with the order of magnitude of the carrying capacity K).

Below, we describe the meaning of the other parameters of the dispersal function by providing a brief mathematical description and displaying some examples of graphs for several illustrative sets of values of parameters. A more detailed and technical description of this function can be found in Section **??**.

Proposition 5.1 (On the function $\boldsymbol{D(x,\mu,\sigma,\delta)}$)**.** *For every $\mu,\delta \ge 0$ and $x \ge 0$ we have $D(x,\mu,0,\delta) \equiv 1$. Moreover, for $\sigma > 0$ we have*

(a) The function $D(x,\mu,\sigma,\delta)$, as a function of x, is continuous, differentiable, and strictly positive.

(b) $D(0,\mu,\sigma,\delta) = 1$ and $\lim_{x\to+\infty} D(x,\mu,\sigma,\delta) = 0$.

(c) If $\mu \ge 1$, then $D(x,\mu,\sigma,\delta)$ is strictly decreasing as a function of x. Moreover, $\frac{\mathrm{d}}{\mathrm{d}x} D(x,\mu,\sigma,\delta)\big|_{x=0}$ is 0 when $\mu = 1$ and negative when $\mu > 1$.

(d) For $0 \le \mu < 1$ and $\delta > 0$, $D(x,\mu,\sigma,\delta)$ is a unimodal function with a maximum at $x^* \in (0,\delta)$ (that is, D is strictly increasing in $[0,x^*]$ and strictly decreasing in $[x^*,+\infty)$). In particular, $\frac{\mathrm{d}}{\mathrm{d}x} D(x,\mu,\sigma,\delta) > 0$ for every $x \in [0,x^*)$. On the other hand, for every $x \in \mathbb{R}^+$, $D(x,\mu,\sigma,\delta) \le D(x^*,\mu,\sigma,\delta) < 2$.

As described by Proposition 5.1, the function D is normalized to one at zero population size, and when the population size tends to infinity, the tendency to disperse converges to zero. Furthermore, D is designed so that the dispersal response of the population of birds generically increases when the population numbers at the patch diminish.

Parameter σ controls the sharpness of the jump and the smoothness of the dispersal function. In Figure **S5.1** we display several graphs for several values of σ. On the left panel, for $\sigma \ge 1$, it can be observed that as σ goes to infinity the graph of D becomes less smooth and the transition from high to small values of the dispersal function is more quick and abrupt. On the right panel, for $\sigma < 1$, it is shown that as σ decreases, the dispersal function becomes flatter and as σ converges to zero the graph of D converges to the constant function one. This latter case corresponds to a constant i.e., independent of the size of the population, dispersal.

Parameter μ controls the tendency (derivative) of the dispersal function at population size $x = 0$. Indeed, for $\mu < 1$ the curve starts at $x = 0$ with increasing tendency; for $\mu = 0$ the curve starts with derivative zero; while for $\mu > 1$ the curve starts with decreasing tendency [see Proposition 5.1(c,d) and Figure **S5.2**]. Observe that the dispersal curve whenever $\mu \ge 1$ is globally strictly decreasing while for $\mu < 1$ it is unimodal. From a more ecological point of view, parameter μ controls how fast dispersal is initiated by the individuals after the ecological perturbation below a given population threshold (modelled by parameter δ). The reason for designing the dispersal function so that it is increasing for low population sizes and low values of μ is to allow the model to deal with a wider range of nonlinear behaviours.

Finally, the dispersal function has been built in such a way that for low population values the tendency to disperse is large while it becomes smaller for large population sizes. Indeed, this is how this function models dispersal by social copying. Parameter δ controls the transition between these two scenarios (see Figure **S5.3**).

5.1.1 Properties of the dispersal function $D(x,\mu,\sigma,\delta)$

The goal of this sub-subsection is to prove Proposition 5.1, which summarizes the most relevant properties of the dispersal function $D(x,\mu,\sigma,\delta)$. To do this, we will first study the functions $\mathcal{E}(x,\sigma,\delta)$ and $\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)$.

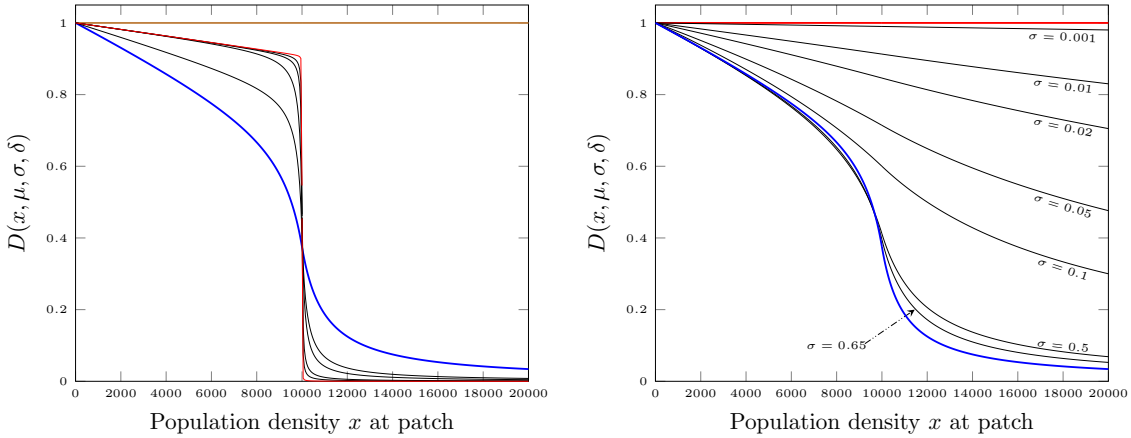


Figure S5.1: Shapes of the function $D(x, \mu, \sigma, \delta)$ used to model social copying behaviour during dispersal. We explore the ranges of $\sigma \geq 1$ (left panel) and $\sigma < 1$ (right panel) fixing $\delta = 10^4$. (Left) The brown graph corresponds to $\sigma = 0$. The sigmoidal-like blue graph has been obtained with $(\mu, \sigma) = (2, 1)$. The red curve, which is in some sense the limiting graph, is obtained with $(\mu = 1.2, \sigma = 10^3)$. The four black curves correspond, from bottom to top, to the following parameter values: $(\mu, \sigma) = (1.5, 5)$, $(\mu, \sigma) = (1.2, 10)$, $(\mu, \sigma) = (1.2, 40)$, and $(\mu, \sigma) = (1.2, 10^2)$. (Right) The red curve in this case is the limiting graph which corresponds to $\sigma = 0$. The blue graph is the same than the one in the left panel. The black curves have been obtained fixing $\mu = 2$, and $\sigma = 0.65, 0.5, 0.1, 0.05, 0.02, 10^{-2}, 10{-}3$.

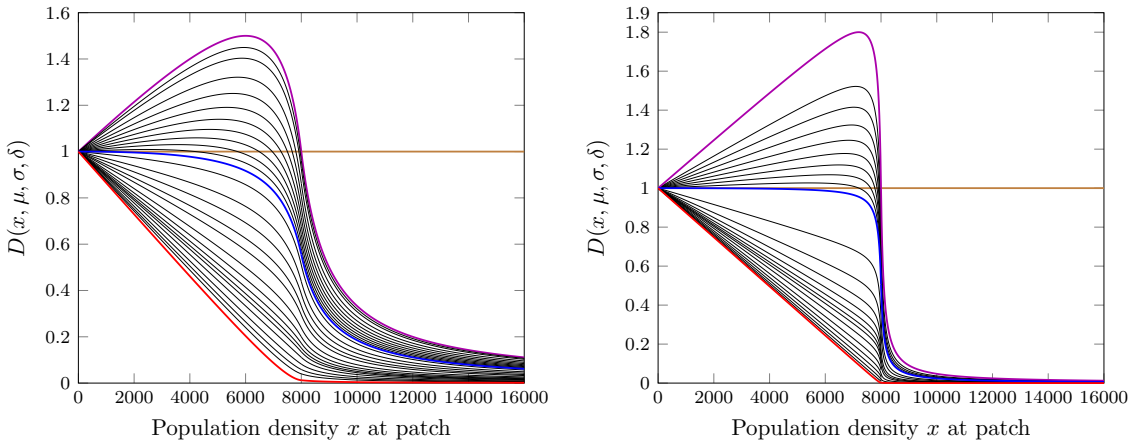


Figure S5.2: Shapes of the dispersal function $D(x, \mu, \sigma, \delta)$ for $\delta = 8 \times 10^3$ and (left panel) $\sigma = 1$ and (right panel) $\sigma = 10$. The violet and the blue curves correspond to $\mu = 0$ and $\mu = 1$, respectively. The red graph is, in some sense, the limiting case: $\mu = 100$ for the left panel and $\mu = 500$ for the right panel. All black curves are organised, from top to bottom, by increasing value of μ. Thus, all black curves between the violet and the blue curves correspond to $\mu < 1$ while the blue curves between the blue and red curves are obtained for $\mu > 1$.

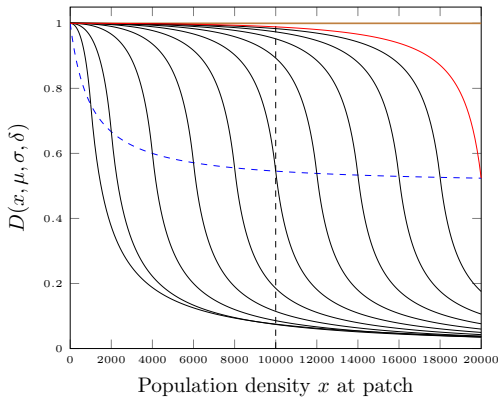


Figure S5.3: Shapes of the dispersal function $D(x, \mu, \sigma, \delta)$ for $\sigma = \mu = 1$ and different values of δ. Each curve is obtained by using the value of δ given by the x coordinate of the intersection of the blue dashed line with the curve. The vertical dashed line identifies the curve obtained with $\delta = 10^4$.

Lemma 5.2 (On the functions $\mathcal{E}(x,\sigma,\delta)$ and $\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)$)**.** *For all $\mu,\sigma,\delta \geq 0$ and $x \geq 0$ we have*

(1) $\mathcal{E}(0,\sigma,\delta) = -\frac{\sigma\delta}{\Theta+\sigma\delta}$, *and* $\mathcal{E}_{\mathsf{dir}}(0,\mu,\sigma,\delta) = -\mu\frac{\sigma\delta}{2\Theta+\sigma\delta}$,

(2) $\mathcal{E}_{\mathsf{dir}}(\delta,\mu,\sigma,\delta) = \mathcal{E}(\delta,\sigma,\delta) = 0$,

(3) $\mathcal{E}_{\mathsf{dir}}(x,\mu,0,\delta) = \mathcal{E}(x,0,\delta) \equiv 0$ *for every* $x \geq 0$,

(4) $-1 < \mathcal{E}(x,\sigma,\delta) < 1$,

(5) $\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{E}(x,\sigma,\delta) = \frac{\Theta\sigma}{(\Theta+\sigma|x-\delta|)^2} > 0$, *and*

(6) $\lim_{x\to+\infty}\mathcal{E}(x,\sigma,\delta) = 1$ *provided that* $\sigma > 0$.

When $\sigma > 0$, $\mathcal{E}$ and $\mathcal{E}_{\mathsf{dir}}$ are continuous as functions of x. Moreover, for $\mu \geq 0$ and $0 \leq x \leq \delta$,

$$\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta) = \frac{\sigma}{\delta(2\Theta+\sigma\delta)(\Theta+\sigma z)}\left(-\Gamma z + \Big(\mu\delta(\Theta+\sigma\delta) + \Gamma(\delta-z)\Big)\frac{\Theta}{\Theta+\sigma z}\right),$$

where $\Gamma := (2-\mu)\Theta + (1-\mu)\sigma\delta$ and $z = \delta - x$.

Proof. Statements (1–6) are obtained by direct computation. The fact that when $\sigma > 0$, $\mathcal{E}$ and $\mathcal{E}_{\mathsf{dir}}$ are continuous follows easily from the definitions of $\mathcal{E}$ and $\mathcal{E}_{\mathsf{dir}}$. Now we prove the last statement of the lemma. From the definition of $\mathcal{E}_{\mathsf{dir}}$ and (5) we have:

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta) &= \frac{1}{\delta}\left(1 - \frac{\mu(\Theta+\sigma\delta)}{2\Theta+\sigma\delta}\right)\mathcal{E}(x,\sigma,\delta) + \\
&\qquad \left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta} + \frac{x}{\delta}\left(1 - \frac{\mu(\Theta+\sigma\delta)}{2\Theta+\sigma\delta}\right)\right)\frac{\mathrm{d}}{\mathrm{d}x}\mathcal{E}(x,\sigma,\delta) \\
&= \frac{1}{\delta}\frac{\Gamma}{2\Theta+\sigma\delta}\frac{\sigma(x-\delta)}{\Theta+\sigma(\delta-x)} + \\
&\qquad \left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta} + \frac{\Gamma}{\delta(2\Theta+\sigma\delta)}x\right)\frac{\Theta\sigma}{(\Theta+\sigma(\delta-x))^2} \\
&= \frac{\sigma}{\delta(2\Theta+\sigma\delta)(\Theta+\sigma z)}\left(-\Gamma z + \Big(\mu\delta(\Theta+\sigma\delta) + \Gamma(\delta-z)\Big)\frac{\Theta}{\Theta+\sigma z}\right).
\end{aligned}$$

□

We are ready to prove Proposition 5.1.

Proof of Proposition 5.1. The fact that $D(x,\mu,0,\delta) \equiv 1$ and Statement (b) follow from the definition of D and Lemma 5.2(3,6). Now we assume $\sigma > 0$. The continuity of D as a function of x follows from its definition and Lemma 5.2. Moreover, by Lemma 5.2(1),

$$\begin{aligned}
D(x,\mu,\sigma,\delta) = \frac{1}{1+\frac{\mu\sigma\delta}{2\Theta+\sigma\delta}}\begin{cases}\big(1-\mathcal{E}(x,\sigma,\delta)\big) & \text{when } x \geq \delta,\\ \big(1-\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)\big) & \text{when } 0 \leq x \leq \delta,\end{cases}\Bigg\} = \\
\frac{2\Theta+\sigma\delta}{2\Theta+(1+\mu)\sigma\delta}\begin{cases}\big(1-\mathcal{E}(x,\sigma,\delta)\big) & \text{when } x \geq \delta,\\ \big(1-\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)\big) & \text{when } 0 \leq x \leq \delta.\end{cases}
\end{aligned} \tag{5.5}$$

Observe that $\frac{2\Theta+\sigma\delta}{2\Theta+(1+\mu)\sigma\delta} > 0$ and, by Lemma 5.2(4), $1-\mathcal{E}(x,\sigma,\delta) > 0$ for $x \geq \delta$. Moreover, for $0 \leq x \leq \delta$,

$$\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right) + \frac{x}{\delta} \geq 0 \tag{5.6}$$

and, by Lemma 5.2(2,5), $\mathcal{E}(x,\sigma,\delta) \leq 0$. Consequently, for $0 \leq x \leq \delta$,

$$1 - \mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)\big) = 1 - \left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right) + \frac{x}{\delta}\right)\mathcal{E}(x,\sigma,\delta) \geq 1 \tag{5.7}$$

Thus, $D(x,\mu,\sigma,\delta)$ is strictly positive.

For $\sigma, \delta > 0$, from the definition of $\mathcal{E}_{\mathsf{dir}}$ and Lemma 5.2(2) we get,

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)\big|_{x=\delta} = \frac{\mathrm{d}}{\mathrm{d}x}\left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right)+\frac{x}{\delta}\right)\Bigg|_{x=\delta}\mathcal{E}(\delta,\sigma,\delta) + \left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right)+\frac{x}{\delta}\right)\Bigg|_{x=\delta}\frac{\mathrm{d}}{\mathrm{d}x}\,\mathcal{E}(x,\sigma,\delta)\big|_{x=\delta} = \frac{\mathrm{d}}{\mathrm{d}x}\,\mathcal{E}(x,\sigma,\delta)\big|_{x=\delta}.$$

Hence, by Equation (5.5), $D(x,\mu,\sigma,\delta)$ is differentiable as a function of $x \geq 0$, and (a) holds.

To prove (c) and (d) we need to study the monotonicity of the map D. For $\sigma > 0$ and $x \geq \delta$, $D(x,\mu,\sigma,\delta)$ is strictly decreasing by Equation (5.5) and Lemma 5.2(5).

On the other hand, for $\sigma, \delta > 0$ and $0 \leq x \leq \delta$ we have by Lemma 5.2 and Equation (5.5),

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\,D(x,\mu,\sigma,\delta) &= -\frac{2\Theta+\sigma\delta}{2\Theta+(1+\mu)\sigma\delta}\frac{\sigma}{\delta(2\Theta+\sigma\delta)(\Theta+\sigma z)}\left(-\Gamma z+\Big(\mu\delta(\Theta+\sigma\delta)+\Gamma(\delta-z)\Big)\frac{\Theta}{\Theta+\sigma z}\right)\\ &= \frac{\sigma}{\delta(\Theta+\sigma z)^2(2\Theta+(1+\mu)\sigma\delta)}\Big(\Gamma z(\Theta+\sigma z)-\Theta\Big(\mu\delta(\Theta+\sigma\delta)+\Gamma(\delta-z)\Big)\Big)\\ &= \frac{\sigma}{\delta(\Theta+\sigma z)^2(2\Theta+(1+\mu)\sigma\delta)}\Big(\Gamma\big(\Theta(2z-\delta)+\sigma z^2\big)-\mu\Theta\delta(\Theta+\sigma\delta)\Big).\end{aligned}$$

Moreover, since $\Theta, \sigma, \delta > 0$ and $\mu \geq 0$, we have

$$\frac{\sigma}{\delta(\Theta+\sigma z)^2(2\Theta+(1+\mu)\sigma\delta)} > 0.$$

Since $z = \delta$ whenever $x = 0$, from the above expression for $\frac{\mathrm{d}}{\mathrm{d}x}\,D(x,\mu,\sigma,\delta)$ we obtain,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\,D(x,\mu,\sigma,\delta)\big|_{x=0} &= \frac{\sigma}{\delta(\Theta+\sigma\delta)^2(2\Theta+(1+\mu)\sigma\delta)}\Big(\Gamma\big(\Theta\delta+\sigma\delta^2\big)-\mu\Theta\delta(\Theta+\sigma\delta)\Big)\\ &= \frac{\sigma}{(\Theta+\sigma\delta)(2\Theta+(1+\mu)\sigma\delta)}\Big(\Gamma-\mu\Theta\Big)\\ &= \frac{\sigma}{(\Theta+\sigma\delta)(2\Theta+(1+\mu)\sigma\delta)}\Big((2-\mu)\Theta+(1-\mu)\sigma\delta-\mu\Theta\Big)\\ &= \frac{\sigma}{(\Theta+\sigma\delta)(2\Theta+(1+\mu)\sigma\delta)}(1-\mu)(2\Theta+\sigma\delta).\end{aligned}$$

Since $2\Theta + \sigma\delta > 0$ it follows that $\frac{\mathrm{d}}{\mathrm{d}x}\,D(x,\mu,\sigma,\delta)\big|_{x=0}$ is positive when $0 \leq \mu < 1$, zero when $\mu = 1$, and negative when $\mu > 1$.

Now we study the monotonicity of $D(x,\mu,\sigma,\delta)$ for $0 < x \leq \delta$ (which is equivalent to $0 \leq z < \delta$). When $\mu \geq 1$ we have $\Gamma \leq \Theta$, and hence

$$\begin{aligned}\Gamma(\Theta(2z-\delta)+\sigma z^2)-\mu\Theta\delta(\Theta+\sigma\delta) &\leq \Theta(\Theta(2z-\delta)+\sigma z^2)-\Theta\delta(\Theta+\sigma\delta)\\ &= \Theta\big(2\Theta(z-\delta)+\sigma(z^2-\delta^2)\big) < 0.\end{aligned}$$

Thus, in summary, $\frac{\mathrm{d}}{\mathrm{d}x}\,D(x,\mu,\sigma,\delta) < 0$ for $0 < x \leq \delta$. Since we already know that $D(x,\mu,\sigma,\delta)\big|_{[\delta,+\infty)}$ is strictly decreasing, it follows that $D(x,\mu,\sigma,\delta)$ is globally strictly decreasing as a function of $x \in \mathbb{R}^+$, and (c) is proved.

Next, to prove (d), we study the shape of $D(x,\mu,\sigma,\delta)\big|_{[0,\delta]}$ when $0 \leq \mu < 1$. To do this notice that $\Gamma(\Theta(2z-\delta)+\sigma z^2)-\mu\Theta\delta(\Theta+\sigma\delta) = 0$ is equivalent to

$$\Gamma(\Theta(2z-\delta)+\sigma z^2) = \mu\Theta\delta(\Theta+\sigma\delta)$$

which, in turn, is equivalent to

$$z(2\Theta+\sigma z)-\Theta\delta = \Theta(2z-\delta)+\sigma z^2 = \frac{\mu\Theta\delta(\Theta+\sigma\delta)}{\Gamma},$$

and to

$$z(2\Theta+\sigma z) = \Theta\delta+\frac{\mu\Theta\delta(\Theta+\sigma\delta)}{\Gamma}.$$

Now observe that, for $0 \le \mu < 1$ we have $\Theta < \Gamma$, and hence

$$z(2\Theta+\sigma z)\Big|_{z=0} = 0 < \Theta\delta + \mu\frac{\Theta}{\Gamma}\delta(\Theta+\sigma\delta) < \delta(2\Theta+\sigma\delta) = z(2\Theta+\sigma z)\Big|_{z=\delta}.$$

Consequently, since $z \mapsto z(2\Theta+\sigma z)$ is a continuous strictly increasing function of $z \ge 0$, there exists a unique $x^*(\mu) = \delta - z^*(\mu) \in (0,\delta)$ such that

$$(\delta - x^*(\mu))(2\Theta + \sigma(\delta - x^*(\mu))) = \Theta\delta + \frac{\mu\Theta\delta(\Theta+\sigma\delta)}{\Gamma},$$

which is equivalent to

$$\Gamma(\Theta(2(\delta - x^*(\mu)) - \delta) + \sigma(\delta - x^*(\mu))^2) - \mu\Theta\delta(\Theta+\sigma\delta) = 0,$$

and to $\frac{\mathrm{d}}{\mathrm{d}x} D(x^*(\mu),\mu,\sigma,\delta) = 0$. Then, the unicity of the critical point x^* and the fact that $\frac{\mathrm{d}}{\mathrm{d}x} D(x,\mu,\sigma,\delta)\Big|_{x=0}$ is positive when $0 \le \mu < 1$ tells us that $D(x,\mu,\sigma,\delta)\Big|_{[0,\delta]}$ is a unimodal map with $D(x,\mu,\sigma,\delta)\Big|_{[0,x^*]}$ strictly increasing and $D(x,\mu,\sigma,\delta)\Big|_{[x^*,\delta]}$ strictly decreasing. By using again the fact that $D(x,\mu,\sigma,\delta)\big|_{[\delta,+\infty)}$ is strictly decreasing, we get (d) except for the fact that $D(x^*,\mu,\sigma,\delta) < 2$. To prove it observe that, for $0 \le x \le \delta$, in view of Lemma 5.2(5,1) and (5.6) and (5.7) we have

$$\begin{aligned}
D(x,\mu,\sigma,\delta) &= \frac{2\Theta+\sigma\delta}{2\Theta+(1+\mu)\sigma\delta}\big(1-\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)\big) \le 1-\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta)\\
&= 1-\left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right)+\frac{x}{\delta}\right)\mathcal{E}(x,\sigma,\delta)\\
&\le 1+\left(\mu\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right)+\frac{x}{\delta}\right)\big(-\mathcal{E}(0,\sigma,\delta)\big) \le 1+\left(\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\left(1-\frac{x}{\delta}\right)+\frac{x}{\delta}\right)\frac{\sigma\delta}{\Theta+\sigma\delta}\\
&= 1+\frac{\sigma\delta}{2\Theta+\sigma\delta}+\frac{x}{\delta}\frac{\sigma\delta}{\Theta+\sigma\delta}\left(1-\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta}\right) = 1+\frac{\sigma\delta}{2\Theta+\sigma\delta}+\frac{x}{\delta}\frac{\sigma\delta}{\Theta+\sigma\delta}\frac{\Theta}{2\Theta+\sigma\delta}\\
&< 1+\frac{\Theta+\sigma\delta}{2\Theta+\sigma\delta} < 2
\end{aligned}$$

because $1-\frac{x}{\delta} \ge 0$. □

Summing up the results shown in Section S 5.1 previous results about the map D we know that, for $\sigma > 0$ and $\mu \ge 1$, $D(x,\mu,\sigma,\delta)$ is a continuous, positive, strictly decreasing function such that $D(0,\mu,\sigma,\delta) = 1$. Hence the function $-\lambda D(x,\mu,\sigma,\delta)$ is strictly increasing as a function of x and since, by Lemma 3.2, we know that $f\big|_{\left(0,\frac{\alpha}{2\beta}\right]}$ is strictly increasing we easily get (see Figs. **S5.4** and **S5.5**):

Lemma 5.3. *The function $F(x) = x(\alpha-\beta x) - \lambda D(x,\mu,\sigma,\delta) = f(x) - \lambda D(x,\mu,\sigma,\delta)$ verifies*

$$\begin{aligned}
F(0) &= f(0) - \lambda = -\lambda < 0,\\
F(K) &= f(K) - \lambda D(K,\mu,\sigma,\delta) = -K\varepsilon - \lambda D(K,\mu,\sigma,\delta) < 0,
\end{aligned}$$

and

$$x(\alpha-\beta x) - \lambda \le F(x) \le x(\alpha-\beta x) - \lambda D(K,\mu,\sigma,\delta) < x(\alpha-\beta x)$$

for every $x \in [0,K]$. Moreover, $F\big|_{\left(0,\frac{\alpha}{2\beta}\right]}$ is strictly increasing.

Next we study the full shape of the function $F\big|_{[0,K]}$ (see Figures **S5.4** and **S5.5**).

Lemma 5.4 (On the shape of $F\big|_{[0,K]}$)**.** *The function $F\big|_{[0,K]}$ verifies one of the following statements:*

(A) F has at most one critical point in the interval $[0,K]$, and this critical point is an inflexion point. Hence, $F\big|_{[0,K]}$ is strictly increasing.

(B) F has a unique critical point c in the interval $[0,K]$. The critical point c is a maximum and belongs to $\left(\frac{\alpha}{2\beta},K\right]$. Hence, $F\big|_{[0,c]}$ is strictly increasing and, when $c < K$, $F\big|_{[c,K]}$ is strictly decreasing.

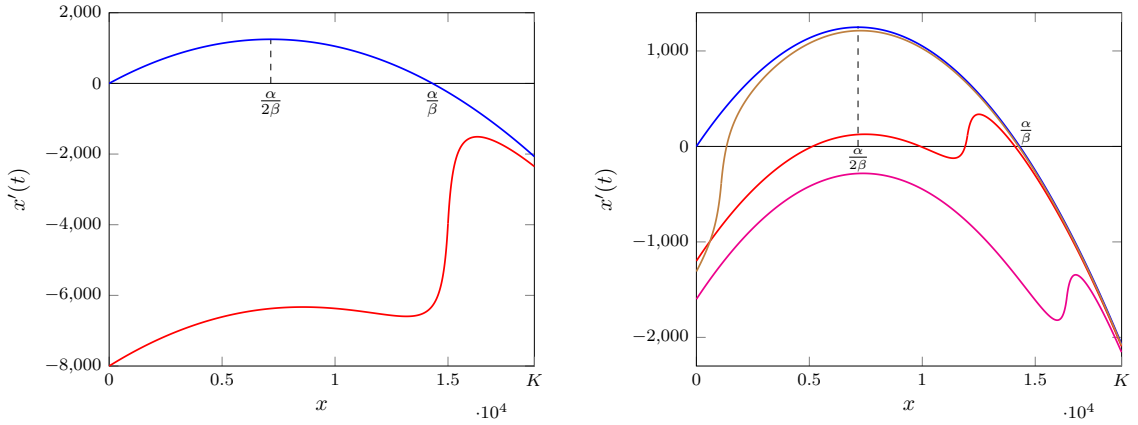


Figure S5.4: The vector field (5.1) in the x-interval $[0, K]$, in parameter's realistic cases. In blue it is shown the graph of $f(x)$ and in red (brown and magenta) several possible graphs of $F(x)$. In all cases, the values of K, α and β are the ones obtained as population's characteristics by fitting the data to the initial phase (see page 18). Likewise, for the red, brown and magenta curves, the values $\varphi = \alpha$, $\mu = 1.2$ and $\sigma = 3.2$ are fixed. The additional parameters take the following values:

- For the red curve in the left picture: $\delta = 15500$, and $\lambda = 8000$.
- For the brown curve in the right picture: $\delta = 1100$, and $\lambda = 1310$.
- For the red curve in the right picture: $\delta = 12000$, and $\lambda = 1200$.
- For the magenta curve in the right picture: $\delta = 16400$, and $\lambda = 1600$.

(C) F has at most two critical points in the interval $[0, K]$, and both critical points belong to the interval $\left(\frac{\alpha}{2\beta}, K\right)$. One of them, donoted by c, is a maximum and the other one is an inflexion point. Hence, $F\big|_{[0,c]}$ is strictly increasing and $F\big|_{[c,K]}$ is strictly decreasing.

(D) F has exactly two critical points $\frac{\alpha}{2\beta} < c^+ < c^- \leq K$ in the interval $[0, K]$. c^+ is a maximum while c^- is a minimum. Hence, $F\big|_{[0,c^+]}$ and $F\big|_{[c^-,K]}$ (when $c^- < K$) are strictly increasing while $F\big|_{[c^+,c^-]}$ is strictly decreasing.

(E) F has exactly three critical points $\frac{\alpha}{2\beta} < c_1^+ < c^- < c_2^+ < K$ in the interval $[0, K]$. c_1^+ and c_2^+ are maxima while c^- is a minimum. Hence, $F\big|_{[0,c_1^+]}$ $F\big|_{[c^-,c_2^+]}$ are strictly increasing while $F\big|_{[c_1^+,c^-]}$ and $F\big|_{[c_2^+,K]}$ are strictly decreasing.

Proof. In this proof we will use the expressions for $\frac{d}{dx}D(x,\mu,\sigma,\delta)$ from above. The whole proof amounts to control the zeros of

$$\begin{aligned} F'(x) &= f'(x) \\ &= \alpha - 2\beta x + \frac{d}{dx}\left(-M + M\begin{cases}\mathcal{E}_{\mathsf{dir}}(x,\mu,\sigma,\delta) & \text{when } 0 \leq x \leq \delta,\\ \mathcal{E}(x,\sigma,\delta) & \text{when } x \geq \delta,\end{cases}\right) \\ &= \alpha - 2\beta x + M\begin{cases}\frac{1}{\delta(2\Theta+\sigma\delta)}\frac{A - z\sigma\Gamma\left(2\Theta+\sigma z\right)}{(\Theta+\sigma z)^2} & \text{when } 0 \leq x \leq \delta,\\ \frac{\Theta\sigma}{(\Theta+\sigma(x-\delta))^2} & \text{when } x \geq \delta,\end{cases} \end{aligned}$$

where $z = \delta - x$ and

$$\begin{aligned} M &:= \frac{\lambda}{1 - \mathcal{E}_{\mathsf{dir}}(0,\mu,\sigma,\delta)} > 0 \text{ (recall that } \lambda > 0 \text{ and } 1 - \mathcal{E}_{\mathsf{dir}}(0,\mu,\sigma,\delta) > 0), \\ \Gamma &:= (2-\mu)\Theta + (1-\mu)\sigma\delta = (2\Theta + \sigma\delta) - \mu(\Theta + \sigma\delta), \text{ and} \\ A &:= \Theta\sigma\delta\big(\mu(\Theta+\sigma\delta) + \Gamma\big) = \Theta\sigma\delta\big(2\Theta + \sigma\delta\big) > 0. \end{aligned}$$

The expression,

$$-\lambda\frac{d}{dx}D(x,\mu,\sigma,\delta) = M\begin{cases}\frac{1}{\delta(2\Theta+\sigma\delta)}\frac{A - z\sigma\Gamma\left(2\Theta+\sigma z\right)}{(\Theta+\sigma z)^2} & \text{when } 0 \leq x \leq \delta,\\ \frac{\Theta\sigma}{(\Theta+\sigma(x-\delta))^2} & \text{when } x \geq \delta,\end{cases} \tag{5.8}$$

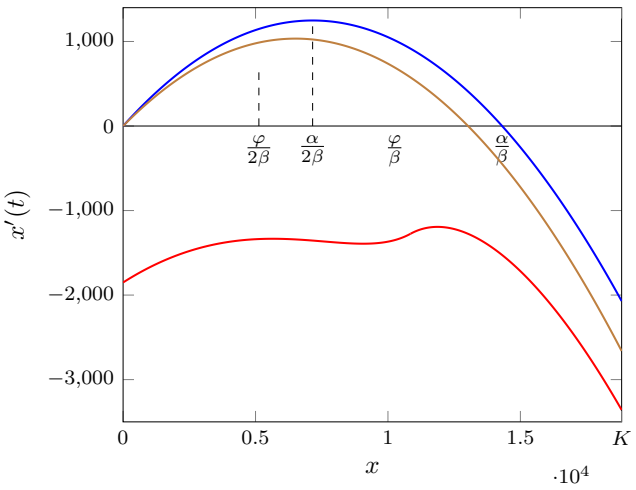


Figure S5.5: The vector field (5.1) in the interval $[0, K]$. In blue it is shown the graph of $f(x)$, in brown it is shown the graph of the map $x(\varphi - \beta x)$, and in red the graph of $F(x)$. The parameter values used in this picture are the following: K, α and β take the values from the table in Subsection **??** as before. For the red and brown curves we take the parameter's values corresponding to the best fit with the phase 2006–2017 data. These are: $\varphi = 0.317559956641$, $\lambda = 1850.9189$, $\mu = 0.00131428$, $\sigma = 0.207329$ and $\delta = 10851.606$.

is strictly positive because $-\lambda D(x, \mu, \sigma, \delta)$ is strictly increasing as a function of x. Therefore, $F'(x) > 0$ for every $x \in \left[0, \frac{\alpha}{2\beta}\right]$ (see Lemma 5.3).

So, if F has a critical point at $x \in [0, K]$, then $x > \frac{\alpha}{2\beta}$ and $F'(x) = 0$, which is equivalent to

$$2\beta x - \alpha = M \begin{cases} \frac{1}{\delta(2\Theta+\sigma\delta)} \frac{A - z\sigma\Gamma\left(2\Theta+\sigma z\right)}{(\Theta+\sigma z)^2} & \text{when } 0 \le x \le \delta, \\ \frac{\Theta\sigma}{(\Theta+\sigma(x-\delta))^2} & \text{when } x \ge \delta. \end{cases} \tag{5.9}$$

Concerning the monotonicity properties of (5.8) we have:

$$\begin{aligned} \frac{d}{dz} \frac{A - z\sigma\Gamma\left(2\Theta + \sigma z\right)}{(\Theta + \sigma z)^2} = -\frac{2\sigma(\Gamma\Theta^2 + A)}{(\Theta + \sigma z)^3} = \\ - \frac{2\sigma\Theta\left(\Theta(2\Theta + \sigma\delta) - \mu\Theta(\Theta + \sigma\delta) + \sigma\delta(2\Theta + \sigma\delta)\right)}{(\Theta + \sigma z)^3} = \\ - \frac{2\sigma\Theta\left((2\Theta + \sigma\delta)(\Theta + \sigma\delta) - \mu\Theta(\Theta + \sigma\delta)\right)}{(\Theta + \sigma z)^3} = \\ - \frac{2\sigma\Theta(\Theta + \sigma\delta)\left(2\Theta + \sigma\delta - \mu\Theta\right)}{(\Theta + \sigma z)^3}. \end{aligned}$$

So, since $M, \frac{1}{\delta(2\Theta+\sigma\delta)} > 0$ we see that

$$-\lambda \left(\frac{d}{dx} D(x, \mu, \sigma, \delta)\right)\Big|_{[0,\delta]} = M \frac{1}{\delta(2\Theta + \sigma\delta)} \frac{A - (\delta - x)\sigma\Gamma\left(2\Theta + \sigma(\delta - x)\right)}{(\Theta + \sigma(\delta - x))^2}\Big|_{[0,\delta]},$$

as a function of x, is strictly increasing when $\mu < 2 + \frac{\sigma\delta}{\Theta}$, constant when $\mu = 2 + \frac{\sigma\delta}{\Theta}$, and strictly decreasing otherwise (when $\mu > 2 + \frac{\sigma\delta}{\Theta}$).

On the other hand,

$$-\lambda \left(\frac{d}{dx} D(x, \mu, \sigma, \delta)\right)\Big|_{[\delta,+\infty)} = M \frac{\Theta\sigma}{(\Theta + \sigma(x - \delta))^2}$$

is strictly decreasing as a function of $x \ge \delta$ because $M > 0$.

Summarising, the right hand side of Equation (5.9) restricted to the interval $\left[\frac{\alpha}{2\beta}, K\right]$ is strictly positive and (see Figure **S5.6**):

(A) decreasing when either $\delta \le \frac{\alpha}{2\beta}$ or $\mu \ge 2 + \frac{\sigma\delta}{\Theta}$,

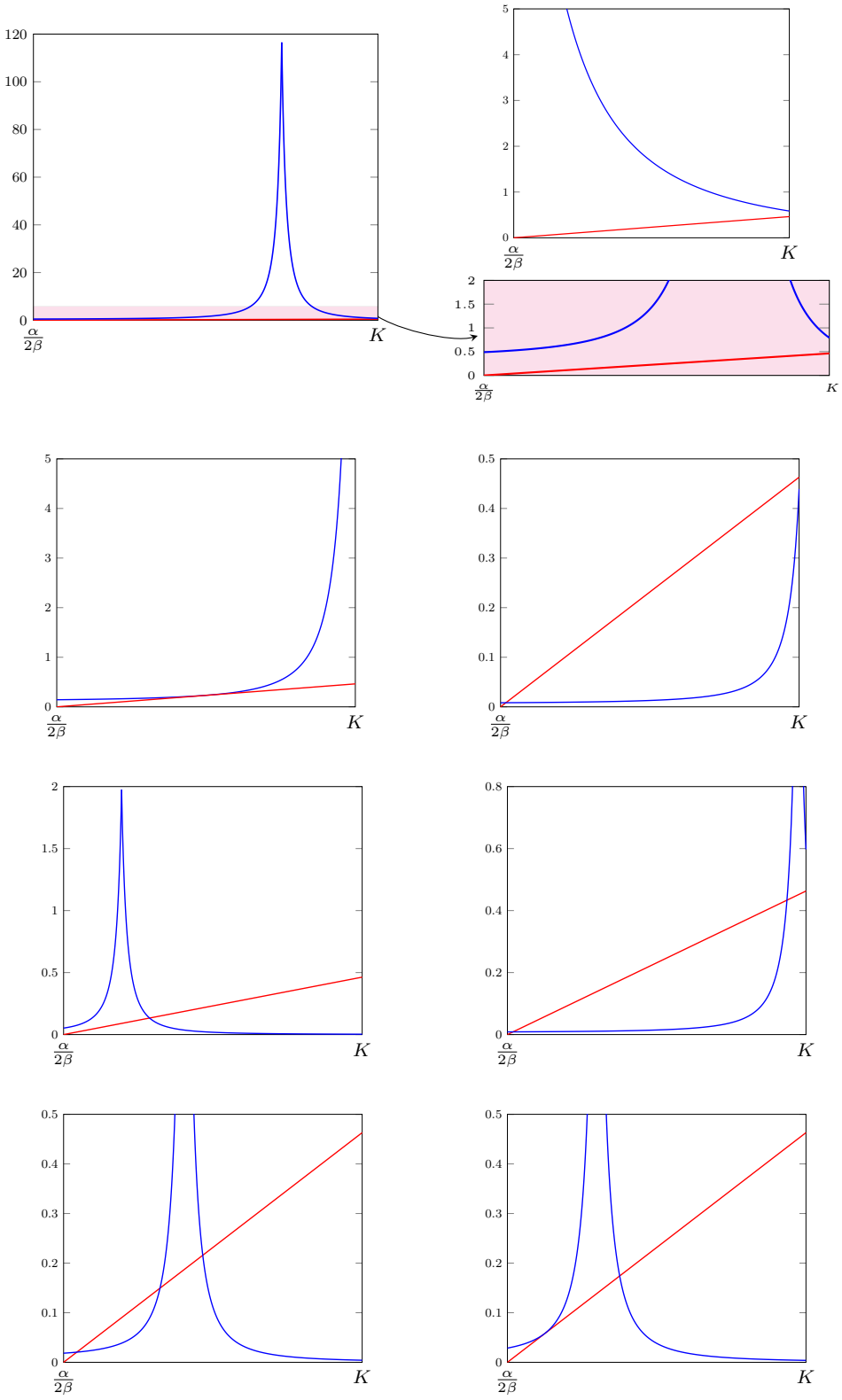


Figure S5.6: The graphs of $-f'(x) = 2\beta x - \alpha$ (in red) for the estimated values of α and β, and the right hand side of Equation (5.9) (in blue) for several realistic values of the parameters $\lambda > 0$, $\mu \geq 1$; $\sigma > 0$ and $\delta > 0$.

(B) strictly increasing when $\delta \geq K$ and $\mu < 2 + \frac{\sigma\delta}{\Theta}$, and

(C) strictly increasing on $\left[\frac{\alpha}{2\beta}, \delta\right]$ and strictly decreasing on $[\delta, K]$, when $\delta \in \left(\frac{\alpha}{2\beta}, K\right)$ and $\mu < 2 + \frac{\sigma\delta}{\Theta}$.

Observe also that $-f'(x) = 2\beta x - \alpha$ is affine with positive slope and vanishes at $x = \frac{\alpha}{2\beta}$. So,

$$(2\beta x - \alpha)\big|_{x=\frac{\alpha}{2\beta}} = 0 < -\lambda\big(\frac{d}{dx}D(x,\mu,\sigma,\delta)\big)\big|_{x=\frac{\alpha}{2\beta}}. \tag{5.10}$$

Consequently, the number κ of intersection points of the curves $-f'(x) = 2\beta x - \alpha$ and $-\lambda\frac{d}{dx}D(x,\mu,\sigma,\delta)$ in the interval $\left(\frac{\alpha}{2\beta}, K\right]$ verifies (see again Figure **S5.6**):

(A) $\kappa \leq 1$ when either $\delta \leq \frac{\alpha}{2\beta}$ or $\mu \geq 2 + \frac{\sigma\delta}{\Theta}$. Moreover, in this case $\kappa = 1$ if and only if

$$-\lambda\big(\frac{d}{dx}D(x,\mu,\sigma,\delta)\big)\big|_{x=K} \leq 2\beta K - \alpha.$$

(B) When $\delta \geq K$ and $\mu < 2 + \frac{\sigma\delta}{\Theta}$ we have $\kappa \in \{0, 1, 2\}$.
Indeed, if $-\lambda\big(\frac{d}{dx}D(x,\mu,\sigma,\delta)\big) \geq 2\beta x - \alpha$ for every $x \in \left(\frac{\alpha}{2\beta}, K\right]$, then $\kappa \in \{0, 1\}$, and if $-\lambda\big(\frac{d}{dx}D(x,\mu,\sigma,\delta)\big)\big|_{x=K} < 2\beta K - \alpha$ then $\kappa = 1$ (see the two pictures in the second row of Figure **S5.6**). The remaining case is when

$$\begin{aligned} &-\lambda\big(\frac{d}{dx}D(x,\mu,\sigma,\delta)\big)\big|_{x=K} \geq 2\beta K - \alpha \text{ and, simultaneously,} \\ &-\lambda\big(\frac{d}{dx}D(x,\mu,\sigma,\delta)\big)\big|_{x=y} < 2\beta y - \alpha \text{ for some } y \in \left(\tfrac{\alpha}{2\beta}, K\right). \end{aligned}$$

Clearly we have $F'(y) < 0$ and, in view of (5.10), $F'\left(\frac{\alpha}{2\beta}\right) > 0$. Thus, there exists $\widetilde{y} \in \left(\frac{\alpha}{2\beta}, y\right)$ such that $F'(\widetilde{y}) = 0$ and $F'(x) > 0$ for every $x \in \left(\frac{\alpha}{2\beta}, \widetilde{y}\right)$. On the other hand, observe that

$$F''(x)\big|_{\left[\frac{\alpha}{2\beta}, K\right]} = M\frac{1}{\delta(2\Theta + \sigma\delta)}\frac{2\sigma\Theta(\Theta + \sigma\delta)\big(2\Theta + \sigma\delta - \mu\Theta\big)}{(\Theta + \sigma z)^3} - 2\beta, \tag{5.11}$$

is strictly increasing as a function of x. Hence, it must be that $F'' < 0$ in the interval $\left(\frac{\alpha}{2\beta}, \widetilde{y}\right]$ since, otherwise, $F'' \geq 0$ on $\left[\widetilde{y}, K\right]$. By the Mean Value Theorem we know that there exists $\xi \in (\widetilde{y}, y)$ such that

$$F'(y) = F'(\widetilde{y}) + F''(\xi)(y - \widetilde{y}) = F''(\xi)(y - \widetilde{y}) \geq 0;$$

a contradiction. By putting all together we see that there exists a point $t \in (\widetilde{y}, K)$ such that $F''(t) = 0$, $F'' < 0$ on the interval $\left(\frac{\alpha}{2\beta}, t\right)$ and $F'' > 0$ on the interval (t, K). Again by the Mean Value Theorem we have that $F' < 0$ on $(\widetilde{y}, t]$. That is, there exists a unique intersection point of the curves $-f'(x) = 2\beta x - \alpha$ and $-\lambda\frac{d}{dx}D(x,\mu,\sigma,\delta)$ in the interval $\left(\frac{\alpha}{2\beta}, t\right]$. Analogously, $F'(K) \geq 0$, and there exists an intersection point of the curves in the interval $(t, K]$. Since F'' is positive in this interval this point must be unique.

(C) When $\delta \in \left(\frac{\alpha}{2\beta}, K\right)$ and $\mu < 2 + \frac{\sigma\delta}{\Theta}$ we have $\kappa \in \{0, 1, 2, 3\}$.
This follows by using (A) on the interval $[\delta, K]$ where the right hand side of Equation (5.9) is strictly decreasing, and by using (B) on $\left[\frac{\alpha}{2\beta}, \delta\right]$ where it is strictly increasing.

Then the proof follows by noticing that (A) fits into Statements (a,b), (B) fits into Statements (a,b,c,d), and (C) fits into Statements (a,b,c,d,e) because when $\kappa = 3$ none of the critical points can be an inflexion point. □

From the above two lemmas we get:

Corollary 5.5. *The vector field $F\big|_{[0,K]}$ has at most 4 stationary states (zeroes) and every possible cardinality of stationary states can be realized with non-degenerate zeroes (that is, zeroes where the map F is locally strictly monotone) with appropriate parameter values. Consequently the potential function of the vector field F has at most 4 critical points (of course, at most two maxima and at most two minima).*

Supplementary Section 6

Model fitting and parameters estimation: Collapse phase 2006–2017

Here, we consider the period from 2006 to 2017, thus focusing on the collapse phase involving the dispersal of almost all the individuals present at the patch of study. In what follows we will consider Model (5.1) with the parameters computed in Section 4.2, since these estimations from the initial phase are considered as the intrinsic population's characteristics. The solution of Model (5.1) with β fixed to $2.4382635446 \times 10^{-5}$, and its parameters belonging to the ranges shown in the table in Section 5 (page 15) will be denoted by $x(t) = x_{\varphi,\lambda,\mu,\sigma,\delta}(t)$, $t = 0, 1, \ldots, 11$. Observe that the solution $x(t)$ depends on the initial condition $x(0) = x_{\varphi,\lambda,\mu,\sigma,\delta}(0) \in [0, K]$, that must be considered as a free parameter as well.

On the other hand, we denote the observed population of Audouin's gulls at the years 2006 to 2017 by

$$\pi(t, t = 0:11) = \mathsf{Audouin's\_Gulls\_Observed\_Population\_at\_year}(2006 + t, t = 0:11) =$$
$$\underset{[0]}{15329}, \underset{[1]}{14177}, \underset{[2]}{13031}, \underset{[3]}{9762}, \underset{[4]}{11271}, \underset{[5]}{8688}, \underset{[6]}{7571}, \underset{[7]}{6983}, \underset{[8]}{4778}, \underset{[9]}{2067}, \underset{[10]}{1586}, \underset{[11]}{793}\big].$$

Now we define the parameter space $\mathscr{P} := [0, K] \times [-\infty, \alpha] \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$, and a map

$$\begin{aligned} \mathsf{F}\colon \quad \mathscr{P} &\longrightarrow \mathbb{R}^+ \\ \big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) &\longmapsto \sqrt{\textstyle\sum_{t=0}^{11} \big(x(t) - \pi(t)\big)^2}. \end{aligned}$$

The *fitting of the model* consists in solving

$$\begin{aligned} &\min \mathsf{F}\big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \\ &\text{subject to } \big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \in \mathscr{P}, \\ &\text{and } x(t) \in [0, K] \text{ for } t = 1, 2, \ldots, 11, \end{aligned} \tag{6.1}$$

and checking that this minimum is as low as possible to guarantee the validity of the model. As it has been noted in the case of the initial phase data in Section S 4, the set

$$\Big\{\mathsf{F}\big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) : \big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \in \mathscr{P} \text{ and } x(t) \in [0, K] \text{ for } t = 1, 2, \ldots, 11\Big\} \subset \mathbb{R}^+$$

has 0 as a lower bound. Hence, it has a minimum element, and Problem (6.1) has at least one solution.

Remark 6.1. The evaluation of the function $\mathsf{F}\big(z(0), \varphi, \lambda, \mu, \sigma, \delta\big)$ has to go through the computation of a solution of the ODE (5.1) whose dispersal term is highly non-linear. When $\lambda > 0$ the approximate solution of (5.1) is computed numerically using the Runge-Kutta-Fehlberg-Simó integrator of order 7–8 with adaptive step-size for speed and efficiency. The version of the integrator that we use has been specially implemented for more optimal speed by one of the authors.

6.1 A first approach to fit the collapse phase: a Sparse Anisotropic Grid Search

As said above, as a first approach to find the solutions of Problem (6.1), it is convenient to perform a brute force exploration of a reasonable region of the parameter space $\mathscr{P}$. The motivation for this exploration is

twofold: first, to have a rough idea of the landscape (graph) of the function F, and, second, to find a point in $\mathscr{P}$, reasonably close to the optimum of Problem (6.1). This point will be used as a fulcrum to determine a compact relatively small set $\mathscr{K} \subset \mathscr{P}$ that contains the minimum (or equivalently that $\mathscr{P} \setminus \mathscr{K}$ does not contain the minimum) of function F. The existence of the compact set $\mathscr{K}$ has two important consequences. First, Bolzano-Weierstrass Theorem tells us that the fitness function F has a minimum in $\mathscr{K}$. Thus, by the choice of $\mathscr{K}$, this minimum must be the solution of Problem (6.1). Second, the reduction of the parameters' search space from $\mathscr{P}$ to $\mathscr{K}$ will greatly simplify the minimization algorithms (which is specially useful when using GAs).

The Grid Searching Method has been implemented sparse and anisotropic on a compact region of the space $\mathscr{P}$ to reduce its computational complexity, i.e. the number of evaluations of the function F. The need for the compacity of the search domain is obvious. Moreover, a smaller domain implies less computational complexity for the grid search. At a first step, the ranges of parameters that determine the compact domain have been chosen arbitrarily, with the hope that the result of the exploration and the subsequent deductions will justify the validity of the choice. The reduction of the computational complexity of the grid search is clearly achieved by choosing a sparse grid but also by the anisotropy. By *anisotropy* we mean that, for certain parameters, the step used to construct the grid is not constant. It rather depends on the zone where the parameter lyes, and on the desired precision in that zone. Of course the sparseness and anisotropy reduce the quality and exhaustiveness of the Grid Searching but, as we will see, the results of this simple search are useful and, indeed, very satisfactory.

The Sparse Anisotropic Grid (SAG) search that we have performed is completely specified in Table 6.1 below, and the main results obtained are summarized in the next lemma.

Parameter	**Theoretical Range**	**Effective Search Range**	**Parameter Search Step**	**# of grid points**
$x_{\varphi,\nu,\lambda,\mu,\sigma,\delta}(0)$	$[0, K]$	$[12600, 18800]$	200	32
φ	$(-\infty, \alpha]$	$[0.13, 0.34]$	0.01	22
λ	$\mathbb{R}^+$	$[300, 3000]$	100	28
μ	$\mathbb{R}^+$	$[0, 10]$	$\begin{cases} 0.1 & \text{when } \mu \in [0, 0.9], \text{ and} \\ 1 & \text{when } \mu \in [1, 10] \end{cases}$	20
σ	$\mathbb{R}^+$	$[0, 50]$	$\begin{cases} 1 & \text{when } \sigma \in [0, 10], \text{ and} \\ 5 & \text{when } \sigma \in [11, 50] \end{cases}$	19
δ	$\mathbb{R}^+$	$[0, 20000]$	10	2001

Table 6.1: Full specification of the Sparse Anisotropic Grid (SAG) search. For every parameter it is given the *effective search range* together with the step (anisotropic in the case of μ and σ) used in the search, and the number of grid points. The SAG search has performed 14,988,610,560 evaluations of $\mathsf{F}\big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big)$ or, equivalently, it has explored 14,988,610,560 points of the feasible space $\mathscr{P}$, all of them contained in the compact set

$$[12600, 18800] \times [0.13, 0.34] \times [300, 3000] \times [0, 10] \times [0, 50] \times [0, 20000].$$

Lemma 6.2. *We have*

$$\mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big) = 2537.417627\cdots$$

and $x_{(15800,0.22,1700,0.3,1,8720)}(t) \in [0, K]$ *for* $t = 1, 2, \ldots, 11$.

In view of the above lemma, it makes sense to graphically explore the vicinity of the optimum point found by the anisotropic grid search to measure the "landscape complexity" of this vicinity. This is done in the Fig. **S6.1**.

To improve the SAG search we have performed a Nonlinear Least-Squares Fitting by using a Levenberg-Marquardt Trust-Region Algorithm taking as seed the SAG optima shown in Lemma 6.2. Unsurprisingly (see Figure **S6.1**) this latter algorithm has not been able to improve the previously found SAG optima (even it has not been able to improve perturbed versions of the seed, used to increase the exploratory character of the whole search).

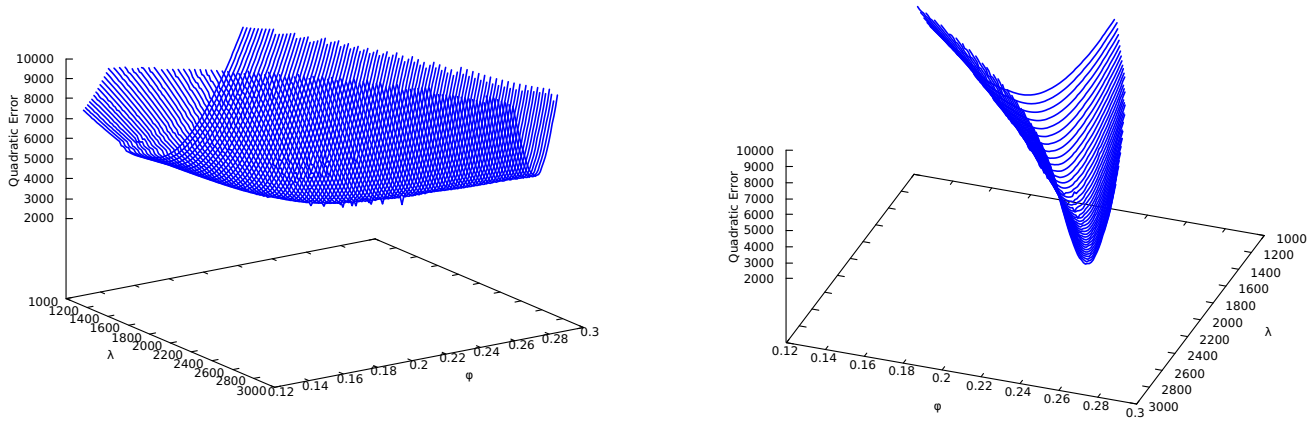


Figure S6.1: Two views of the landscape of the function F around the point $(15800, 0.22, 1700, 0.3, 1, 8720)$. The landscape shows only points whose $\mathsf{F}-$value is lower than 10000. In both pictures we have fixed the following 4 parameters: $x(0) = 15800$, $\mu = 0.3$, $\sigma = 1$ and $\delta = 8720$, λ ranges from 1000 to 3000 while φ ranges from 0.13 to α.

Conclusion: In this search/computation we greatly need to increase the exploratoryness of our algorithms. This leads us in a natural way to GAs, tuned to be highly exploratory.

Also, in this context it would be good to improve our knowledge of the parameter's space and, if possible, to try to reduce it as much as possible to decrease the complexity of the GAs as well as improving their computational efficiency.

6.2 Analytic and heuristic estimates of a compact domain that contains the optimum

The goal of this subsection is to justify that the search space for solving Problem (6.1) can be greatly reduced to the compact set

$$\mathscr{S} := [12791, 17867] \times [0.13, \alpha] \times [300, 3000] \times [0, 10] \times [0, 50] \times [0, 20000].$$

To this end, we introduce the *reduced minimization problem*

$$\begin{aligned} &\min \mathsf{F}\big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \\ &\text{subject to } \big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \in \mathscr{S}, \\ &\text{and } x(t) \in [0, K] \text{ for } t = 1, 2, \ldots, 11, \end{aligned} \tag{6.2}$$

and we will semi-analytically justify (with the help of an heuristic reasoning) that the next results holds:

Proposition 6.3. *The solutions of Problem* (6.1) *and Problem* (6.2) *coincide.*

To justify Proposition 6.3 we will use the following technical lemma. Denote,

$$\widetilde{\mathscr{S}} := [12791, 17867] \times [0.13, \alpha] \times [300, +\infty] \times \mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+ \subset \mathscr{P}.$$

Lemma 6.4. *Let* $\vec{\boldsymbol{\theta}} \in \mathscr{P} \setminus \widetilde{\mathscr{S}}$. *Then,*

$$\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) > \mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big).$$

Consequently,

$$\arg\min_{\vec{\boldsymbol{\theta}} \in \mathscr{P}} \mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) \in \widetilde{\mathscr{S}}.$$

Remark 6.5. The above lemma shows that the solution of Problem 6.1 verifies $\lambda \geq 300 \gg 0$, thus proving analytically the hypothesized highly non-linear migratory behaviour of both, the Audouin's gulls and the theoretical model.

To prove Lemma 6.4 we will use the following analytical result. Let $R \subset \mathbb{R}$ be an interval, let $f\colon R \longrightarrow \mathbb{R}$ be continuous and let $x(t)$ denote the solution of the differential equation $\frac{\mathrm{d}}{\mathrm{d}t}\, x(t) = f\big(x(t)\big)$ with initial condition $x(t_0) = \xi \in R$. Assume that this solution exists and is bounded in an interval $[t_0, T]$. Then, for every $t \in [t_0, T]$,

$$x(t) = \xi + \int_{t_0}^{t} f(x(s))\, \mathrm{d}s. \tag{6.3}$$

Lemma 6.6. *Let $R \subset \mathbb{R}$ be an interval, let $f, g\colon R \longrightarrow \mathbb{R}$ be continuous functions and let $x(t)$ and $y(t)$ denote the solutions of the differential equations $\frac{\mathrm{d}}{\mathrm{d}t}\, x(t) = f\big(x(t)\big)$ and $\frac{\mathrm{d}}{\mathrm{d}t}\, y(t) = g\big(y(t)\big)$ with initial conditions $x(t_0) \in R$ and $y(t_0) \in R$, respectively. Assume that these solutions are defined and bounded in an interval $[t_0, T]$. Suppose that $y(t_0) \le x(t_0)$ and $g(x) \le f(x)$ for every $x \in R$. Then, $y(t) \le x(t)$ for every $t \in [t_0, T]$.*

Now we are ready for the

Proof of Lemma 6.4. Let $\vec{\boldsymbol{\theta}} = (\kappa, \varphi, \lambda, \mu, \sigma, \delta) \in \mathscr{P} \setminus \widetilde{\mathscr{S}}$. We start by assuming that $\kappa \le 12791$. By Lemma 6.2 we have,

$$\mathsf{F}\big(\kappa, \varphi, \nu, \lambda, \mu, \sigma, \delta\big) = \sqrt{\sum_{t=0}^{11} \big(x(t) - \pi(t)\big)^2} \ge \sqrt{\big(\pi(0) - x(0)\big)^2} =$$

$$\pi(0) - \kappa \ge 15329 - 12791 > \mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big).$$

Analogously, if $\kappa \ge 17867$,

$$\mathsf{F}\big(\kappa, \varphi, \nu, \lambda, \mu, \sigma, \delta\big) \ge \kappa - \pi(0) \ge 17867 - 15329 > \mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big).$$

Thus, in what follows we may assume that $x(0) = \kappa \in (12791, 17867)$.

Now assume that $\varphi \le 0.13$. We denote by $u(t)$, $t = 0, 1, \dots, 11$, the solution of Model (3.1) with α replaced by $\widetilde{\varphi} = 0.13$, $\beta = 2.4382635446\cdots \times 10^{-5}$, and initial condition $u(0) = 17867$. By Lemma 3.1,

$$u(t) = \frac{\widetilde{\varphi} u(0) \exp(\widetilde{\varphi} t)}{\widetilde{\varphi} + \beta u(0)(\exp(\widetilde{\varphi} t) - 1)},$$

which is clearly defined and bounded on the interval $[0, 11]$. By direct computation, we get

$$\begin{aligned} u(1) &\approx 13886.87\cdots &&< \pi(1) = 14177,\\ u(2) &\approx 11614.90\cdots &&< \pi(2) = 13031,\\ u(4) &\approx \ 9146.96\cdots &&< \pi(4) = 11271,\\ u(5) &\approx \ 8413.05\cdots &&< \pi(5) = \ \ 8688, \end{aligned}$$

and

$$\sqrt{\sum_{t \in \{1,2,4,5\}} \big(\pi(t) - u(t)\big)^2} = 2583.929\cdots\ .$$

By Proposition 5.1(a),

$$\varphi x - \beta x^2 - \lambda D(x, \mu, \sigma, \delta) \le \widetilde{\varphi} x - \beta x^2,$$

for every $x \in \mathbb{R}^+$. Then, since $x(0) \le 17867 = u(0)$, by Lemma 6.6,

$$x(t) \le u(t) < \pi(t)$$

for $t = 1, 2, 4, 5$. Hence, by Lemma 6.2,

$$\mathsf{F}\big(\kappa, \varphi, \nu, \lambda, \mu, \sigma, \delta\big) = \sqrt{\sum_{t=0}^{11} \big(x(t) - \pi(t)\big)^2} \ge \sqrt{\sum_{t \in \{1,2,4,5\}} \big(\pi(t) - x(t)\big)^2} \ge$$

$$\sqrt{\sum_{t \in \{1,2,4,5\}} \big(\pi(t) - u(t)\big)^2} > 2583 > \mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big).$$

So, in what follows we additionally may assume that $\varphi > 0.13$.

Next we denote by $z(t)$, $t = 0, 1, \ldots, 11$, the solution of

$$\frac{dx(t)}{dt} = \widetilde{\varphi}x(t) - \beta x(t)^2 - \nu, \tag{6.4}$$

with $\widetilde{\varphi} = 0.13$, $\beta = 2.4382635446\cdots\times 10^{-5}$, $\nu = 600$ and initial condition $z(0) = 12791$. By direct computation, we get

$$\begin{array}{lll} z(9) & \approx 3783.02\cdots & > \quad \pi(9) = 2067, \\ z(10) & \approx 3336.48\cdots & > \quad \pi(10) = 1586, \\ z(11) & \approx 2904.36\cdots & > \quad \pi(11) = \ 793, \end{array}$$

and

$$\sqrt{\sum_{t\in\{9,10,11\}} \big(\pi(t) - z(t)\big)^2} = 3235.24\cdots\ .$$

By Proposition 5.1 we have

$$D(x,\mu,\sigma,\delta) \le \begin{cases} D(0,\mu,\sigma,\delta) = 1 < 2 & \text{when } \mu \ge 1, \text{ and} \\ D(x^*,\mu,\sigma,\delta) < 2 & \text{when } 0 \le \mu < 1. \end{cases}$$

Consequently, when $\lambda \le 300$ we have $\lambda D(x,\mu,\sigma,\delta) < 2\lambda \le \nu$, and

$$\varphi x - \beta x^2 - \lambda D(x,\mu,\sigma,\delta) > \widetilde{\varphi}x - \beta x^2 - \nu$$

for every $x \in \mathbb{R}^+$. Then, since $x(0) > 12791 = z(0)$, by Lemma 6.6,

$$x(t) \ge z(t) > \pi(t)$$

for $t = 9, 10, 11$. Hence, by Lemma 6.2,

$$\mathsf{F}\big(\kappa,\varphi,\nu,\lambda,\mu,\sigma,\delta\big) = \sqrt{\sum_{t=0}^{11}\big(x(t)-\pi(t)\big)^2} \ge \sqrt{\sum_{t\in\{9,10,11\}}\big(\pi(t)-x(t)\big)^2} \ge$$
$$\sqrt{\sum_{t\in\{9,10,11\}}\big(\pi(t)-z(t)\big)^2} > 3235 > \mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big).$$

414 □

Justification of Proposition 6.3. From Lemma 6.4 we see that

$$\min\Big\{\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) : \vec{\boldsymbol{\theta}} \in \mathscr{P} \text{ and } x_{\vec{\boldsymbol{\theta}}}(t) \in [0,K] \text{ for } t = 0,1,2,\ldots,11\Big\} =$$
$$\min\Big\{\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) : \vec{\boldsymbol{\theta}} \in \widetilde{\mathscr{P}} \text{ and } x_{\vec{\boldsymbol{\theta}}}(t) \in [0,K] \text{ for } t = 0,1,2,\ldots,11\Big\}.$$

On the other hand, we have performed another SAG search on the set $\mathscr{D} \setminus \mathscr{S}$, where

$$\mathscr{S} \subset \mathscr{D} := [12791, 17867] \times [0.13, \alpha] \times [300, 100000] \times [0, 600] \times [0, 5000] \times [0, 159000] \subset \widetilde{\mathscr{P}};$$

415 with the following parameter's search steps:

Parameter	**Parameter Search Step**
$x_{\varphi,\nu,\lambda,\mu,\sigma,\delta}(0)$	100
φ	0.05
λ	$\begin{cases} 100 & \text{when } \lambda \in [300, 2900], \text{ and} \\ 1000 & \text{when } \lambda \in [3000, 100000] \end{cases}$
μ	$\begin{cases} 0.1 & \text{when } \mu \in [0, 0.9], \\ 1 & \text{when } \mu \in [1, 49], \\ 10 & \text{when } \mu \in [50, 90], \text{ and} \\ 100 & \text{when } \mu \in [100, 600] \end{cases}$
σ	$\begin{cases} 0.1 & \text{when } \sigma \in [0, 1.9], \\ 2 & \text{when } \sigma \in [2, 38], \\ 10 & \text{when } \sigma \in [40, 90], \text{ and} \\ 100 & \text{when } \sigma \in [100, 5000] \end{cases}$
δ	1000

416

The results of this second Sparse Anisotropic Grid Search are the following:

Lemma 6.7. *For every point* $\vec{\boldsymbol{\theta}} \in \mathscr{D} \setminus \mathscr{S}$ *whose components belong to the grid described in the above table we have*

$$\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) > 2664 > \mathsf{F}\big(15800, 0.22, 1700, 0.3, 1, 8720\big).$$

Hence, with the help of an appropriate continuity argument we get

$$\min\Big\{\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) : \vec{\boldsymbol{\theta}} \in \mathscr{D} \text{ and } x_{\vec{\boldsymbol{\theta}}}(t) \in [0, K] \text{ for } t = 0, 1, 2, \ldots, 11\Big\} = \min\Big\{\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) : \vec{\boldsymbol{\theta}} \in \mathscr{S} \text{ and } x_{\vec{\boldsymbol{\theta}}}(t) \in [0, K] \text{ for } t = 0, 1, 2, \ldots, 11\Big\}.$$

By a Montecarlo exploration on the region $\widetilde{\mathscr{S}} \setminus \mathscr{D}$, a continuity argument as above heuristically gives

$$\min\Big\{\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) : \vec{\boldsymbol{\theta}} \in \widetilde{\mathscr{S}} \text{ and } x_{\vec{\boldsymbol{\theta}}}(t) \in [0, K] \text{ for } t = 0, 1, 2, \ldots, 11\Big\} = \min\Big\{\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) : \vec{\boldsymbol{\theta}} \in \mathscr{D} \text{ and } x_{\vec{\boldsymbol{\theta}}}(t) \in [0, K] \text{ for } t = 0, 1, 2, \ldots, 11\Big\}.$$

This ends the justification of Proposition 6.3. □

6.3 Fitting the collapse phase using artificial intelligence: Genetic Algorithms

As it has been already explained, we want to minimize the function F (or, ideally, to find a vector of parameters $\vec{\boldsymbol{\theta}} \in \mathscr{P}$ such that $\mathsf{F}\big(\vec{\boldsymbol{\theta}}\big) = 0$). This amounts solving Problem 6.1 which, in view of Proposition 6.3, is equivalent to solve the *reduced minimization problem*

$$\begin{aligned} &\min \mathsf{F}\big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \\ &\text{subject to } \big(x(0), \varphi, \lambda, \mu, \sigma, \delta\big) \in \mathscr{S}, \\ &\quad \text{and } x(t) \in [0, K] \text{ for } t = 1, 2, \ldots, 11. \end{aligned} \tag{6.2}$$

To solve Problem (6.2) we will use the following Standard Genetic Algorithm (GA) tuned to be highly exploratory, with F as its fitness function. A diagram of the GA displayed in Box 1 below with the pseudo-code, the used variables, the implemented functions, and the execution flow.

Box 1. A Standard Exploratory GA

```
popsize ← desired population size                                   ▷ Must be even
P ← {}                                                              ▷ Initializing empty first generation
for popsize times do
    P ← P ∪ {new random individual}
end for                                                             ▷ End of first generation initialization

procedure GeneticAlgorithm(P, popsize)
    bestfitness ← ∞                                                 ▷ bestfitness initialization
    repeat
        for each individual p ∈ P do
            fit ← F(p)
            if fit < bestfitness then                               ▷ True when bestfitness is ∞. Found best individual so far
                bestfitness ← fit
                Best ← p
            end if
        end for
        Q ← {}                                                      ▷ Initializing empty next generation
        for popsize/2 times do
            Parent p_a ← SelectWithReplacement(P)
            Parent p_b ← SelectWithReplacement(P)
            Children c_a, c_b ← Crossover(Copy(p_a), Copy(p_b))
            Q ← Q ∪ {Mutate(c_a),Mutate(c_b)}
        end for
        P ← Q          ▷ The population is replaced completely at each generation: we want to be exploratory
    until Best is the ideal solution or the maximum number of generations has been exhausted
    return Best
end procedure
```

To completely specify such an algorithm, a number of its elements have to be defined and revised. Namely: how individuals are coded, the operations SelectWithReplacement, Crossover, Mutate, and finally the stopping criteria. Also, a number of parameters of the algorithm have to be introduced and discussed: the population size (*popsize*), the maximum number of generations and the mutation probability. Additionally, the stopping criteria and the function select with replacement (SelectWithReplacement) depend on internal parameters that will be explained whenever these two procedures are surveyed.

The next subsections will be devoted to explain our implementation of the items listed above.

6.3.1 Population parameters and stopping criteria

Since the algorithm must be tuned to be highly exploratory it is necessary to have large *popsize* and number of generations. We have taken 20000 as tentative value for *popsize* (although this will be better dealt in our variant of the algorithm developed in Section 6.3.7), and the maximum number of generations *MaxNumGen* is taken to be 1000. Accordingly, the mutation probability has been set to 0.1 (it must be small but not too small for exploratoriness).

The basic stopping criteria depends on a new parameter which controls the maximum number of generations without improving *bestfitness*. This parameter, called, *MaxNumGenWithoutImproving* is set to 100. It works essentially as indicated by its name: after 100 generations without improvement (or modification) of *bestfitness* the algorithm stops and returns the best individual found. This can be taken as our definition of the term ideal solution in the above version of the GA.

6.3.2 Individuals

An individual in the population is specified by six "genes" corresponding to the six free parameters: the initial condition $x(0)$, φ, λ, μ, σ and δ. In the framework established by the proof of Holland's Convergence Theorem [16] it is convenient to write the genes as unsigned integers expressed in binary, with its range depending on the true range of the real parameters and its sensitivity. This leads to the distinction between the individual's *phenotype* which corresponds to the real (human readable) parameter values, and the *(discretized)*

genotype which is composed of the same parameters but written as unsigned integers in binary. Of course the *translation procedures* from phenotype to genotype and vice-versa must be specified. In Table 6.3.2 we specify, for each phenotypic parameter, its *Theoretical Range*, the associated *Effective Search Range* determined by the set $\mathscr{S}$ (which is the feasible space of Problem (6.2)), a reasonable sensitivity, or, better said, *Precision or Discretization Step* (which determines the precision of the parameter's estimate), and the associated genotype which consists on two elements: the *Unsigned Integer Upper Limit* which, as we will see, should be always taken as a power of two 2^ℓ (where ℓ depends on the effective search range and discretization step — it determines also an *unsigned integer search range* of the form $0, 1, 2, \ldots, 2^\ell - 1$), and the *translation function* from genotype to phenotype.

Table 6.3.2. **Full specification of individuals for the GA and their genes coding**

		Phenotype		Genotype		
Parameter	Theoretical Range	Effective Search Range	Precision or Discretization Step	Unsigned Integer Upper Limit	Effective Discretization Step	Translation map from genotype to phenotype
$x(0)$	$[0, K]$	$[12791, 17867]$	10^{-2}	2^{19}	$\frac{5076}{2^{19}-1}$	$u \longmapsto 12791 + u\ \frac{5076}{2^{19}-1}$
φ	$(-\infty, \alpha]$	$[0.13, \alpha]$	10^{-10}	2^{32}	$\frac{\alpha-0.13}{2^{32}-1}$	$u \longmapsto 0.13 + u\ \frac{\alpha-0.13}{2^{32}-1}$
λ	$\mathbb{R}^+$	$[300, 3000]$	10^{-2}	2^{19}	$\frac{2700}{2^{19}-1}$	$u \longmapsto 300 + u\ \frac{2700}{2^{19}-1}$
μ	$\mathbb{R}^+$	$[0, 10]$	10^{-6}	2^{24}	$\frac{10}{2^{24}-1}$	$u \longmapsto u\ \frac{10}{2^{24}-1}$
σ	$\mathbb{R}^+$	$[0, 50]$	10^{-4}	2^{19}	$\frac{50}{2^{19}-1}$	$u \longmapsto u\ \frac{50}{2^{19}-1}$
δ	$\mathbb{R}^+$	$[0, 20000]$	0.5	2^{16}	$\frac{20000}{2^{16}-1}$	$u \longmapsto u\ \frac{20000}{2^{16}-1}$

Remark 6.8 (On the determination of the Unsigned Integer Upper Limit and the Effective Discretization Step)**.** Assume that a parameter has an *Effective Search Range* of the form $[A, B]$ and a desired *Precision or Discretization Step* ξ. The integer range corresponding to A, B, and ξ is $0, 1, 2, \ldots, \left\lceil \frac{B-A}{\xi} \right\rceil$, where $\lceil \cdot \rceil$ denotes the *ceiling function*. The *Genotype Unsigned Integer Upper Limit* is defined to be the smallest power of two 2^ℓ such that $\left\lceil \frac{B-A}{\xi} \right\rceil \leq 2^\ell$.

Then, the available range of genotypic values for the parameter is $0, 1, 2, \ldots, 2^\ell - 1$, and the *Effective precision or Discretization Step* is $\frac{B-A}{2^\ell-1}$. Consequently, the translation formula from genotype to phenotype is

$$u \longmapsto A + u\ \frac{B - A}{2^\ell - 1},$$

and hence $0 \longmapsto A$ and $2^\ell - 1 \longmapsto A + (2^\ell - 1)\ \frac{B-A}{2^\ell-1} = B$.

Remark 6.9. All *Genotype Unsigned Integer Upper Limits* in Table 6.3.2 above have exponent less than or equal to 32. This means that the base data type to store the genotypic values of all genes can be `unsigned int`'s of 32 bits.

Observe that *in this framework the restrictions on the parameters are verified automatically.* Indeed, we are restricting the genotypic values of a parameter to integers of the form $0, 1, 2, \ldots, 2^\ell - 1$ with a translation formula from genotype to phenotype of the form

$$u \longmapsto A + u\ \frac{B - A}{2^\ell - 1} \in [A, B].$$

Since the *effective search ranges* in Table 6.3.2 are contained in the *Theoretical Ranges* and all values of parameters constructed by the GA are valid in the genotypic sense (i.e. belong to $\{0, 1, 2, \ldots, 2^\ell - 1\}$), the phenotypic parameter values must belong to the *Theoretical Ranges*, and hence verify all restrictions.

Remark 6.10 (On why we want the *Genotype Unsigned Integer Upper Limit* to be a power of two)**.** Observe that all genotypic values in the range $0, 1, 2, \ldots, 2^\ell - 1$, written in binary have a string of $32 - \ell$ consecutive zeroes at the most significant bits part of the number, and a string of ℓ least significant bits. Eventually, for the number $2^\ell - 1$, all ℓ least significant bits are set to 1. This eases the programming of crossovers and mutations, and will help avoiding complicate feasibility tests.

A final comment referring to the individuals' genotypes is that, for computational efficiency, is crucial to define an appropriate *data type* for them. In our case an individual is a `struct` composed by a vector of 6 *unsigned integers* (the genotype) and a variable to store the fitness value of the individual. This is accompanied (at the level of the whole population not of each individual) by the six translation formulae from genotype to phenotype shown in Table 6.3.2, and the list of exponents of the *Genotype Unsigned Integer Upper Limits* that, as explained in Remark 6.10, is crucial when setting the crossover and mutation procedures.

Finally, if we denote by $\left(u_{x(0)}, u_\varphi, u_\lambda, u_\mu, u_\sigma, u_\delta\right)$ the genes vector of an individual then, the *genotypic fitness function* is

$$\left(u_{x(0)}, u_\varphi, u_\lambda, u_\mu, u_\sigma, u_\delta\right) \longmapsto$$
$$\mathsf{F}\Big(12791 + u_{x(0)}\ \tfrac{5076}{2^{19}-1}, 0.13 + u_\varphi\ \tfrac{\alpha-0.13}{2^{32}-1}, 300 + u_\lambda\ \tfrac{2700}{2^{19}-1}, u_\mu\ \tfrac{10}{2^{24}-1}, u_\sigma\ \tfrac{50}{2^{19}-1}, u_\delta\ \tfrac{20000}{2^{16}-1}\Big).$$

6.3.3 Selection with replacement

We use tournament algorithm with tournament parameter 10 (to increase exploratoriness). The detailed explanation of the procedure (in pseudocode) is the following:

Tournament Selection Algorithm

Require: P, t ▷ The population and the tournament size, $t \geq 1$
  $Best \leftarrow$ individual picked at random from P with replacement
  for $i = 2$ **to** t **do**
    $p \leftarrow$ individual picked at random from P with replacement
    if $\mathsf{F}(p) < \mathsf{F}(Best)$ **then**
      $Best \leftarrow p$
    end if
  end for
  return $Best$

6.3.4 Random initial population

Randomness in this setting is crucial although our variant of the algorithm will slightly improve — for good reasons — the initial population thus breaking its "pure randomness".

The initial population plays the role of a sample and, if it is not distributed uniformly in the whole search space, the optimum can be far from this initial sample and therefore missed[1] or, at least, the whole search can be delayed[2].

To assure the randomness of the initial population we avoid the use of congruential random number generators. We use a completely different approach. First we have designed a high quality random bits generator. This is done with a standard (i.e. not "high tech") random numbers generator modified for binary lotteries (i.e. giving only 0's and 1's). Then as a second step we use the random binary lotteries generator to perform lotteries in pairs and use the John von Neumann trick: if both results in the pair coincide, the roll is discarded; if, on the contrary, the results in the pair are different we take the first one as the generated resultant bit. This very clever von Neumann's strategy gives an unbiased random bits generator but it is somehow inefficient[3].

Equipped with the unbiased random bits generator, to build the initial population in generation zero, we fill the *Genotype Unsigned Integer Upper Limit exponent*–least significant bits of the six genes of every one of the *popsize* individuals.

6.3.5 Mutation

We use a very simple but aggressive mutation scheme (recall that we have to be highly exploratory). For *every gene (genotypic parameter) of every generated child* we swap a single random bit (among the *Genotype Unsigned Integer Upper Limit exponent*–least significant bits) with probability $MutationProbability = 0.1$.

[1] However, if the algorithm is exploratory enough, it can discover regions "hidden" to the initial population and "repare" this problem.

[2] An initial population well randomised is therefore a good investment in computational efficiency.

[3] The more biased it is the initial random binary lotteries generator, more rolls will have to be discarded and more inefficient the random bits generator becomes.

6.3.6 Crossover

Here, we perform one-point crossover among the *Genotype Unsigned Integer Upper Limit exponent*–least significant bits of *every gene (genotypic parameter)* of the two parents. This is best explained in the following picture

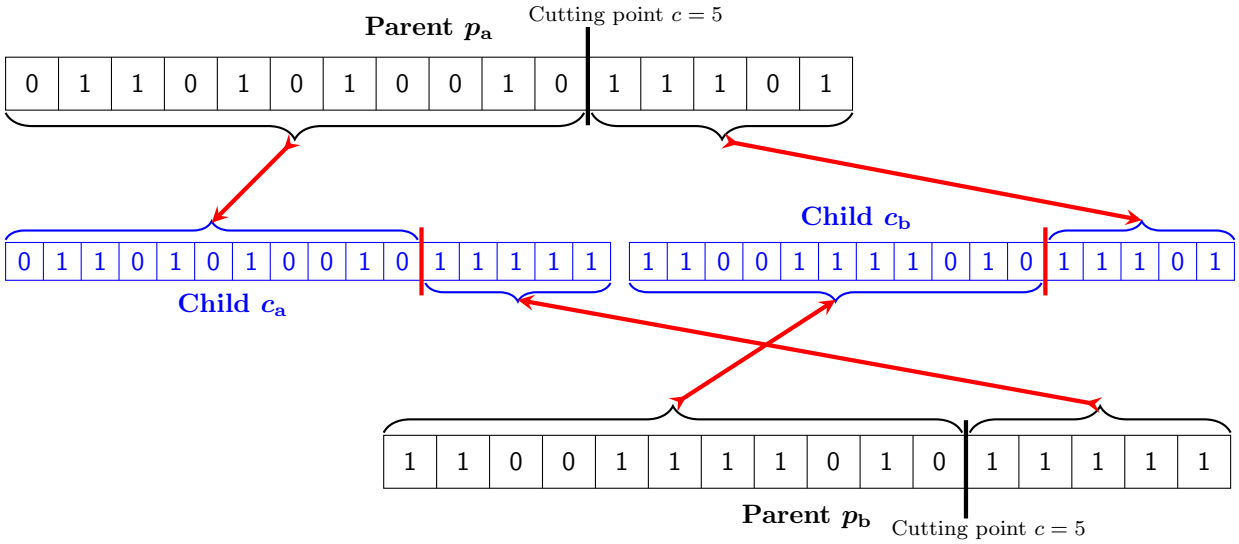


Figure S6.2: For both parents we show the *Genotype Unsigned Integer Upper Limit exponent*–least significant bits of the same gene (say gene 6 that would correspond to parameter δ) of the genotype. In accordance with Table 6.3.2, the *Genotype Unsigned Integer Upper Limit exponent* is 16. We also show the one-point crossover with cutting point at bit $c = 5$.

The *crossover cutting point* c is selected randomly for every gene among the least significant bits; but in a way that there is effective crossover (i.e. c must be different from 0 and the *Genotype Unsigned Integer Upper Limit exponent*).

6.3.7 The Set of Superior variants of the Genetic Algorithm and its execution flow

Usually, the execution flow of a Genetic Algorithm (GA) is to run a batch of instances of the algorithm (in our case the *Standard Exploratory GA*) with different sets of algorithmic parameters (*popsize*, *MaxNumGen*, the mutation probability and others) and giving as a candidate to the optimum the best individual found in the whole batch.

However, as seen in Figure **S6.1** (see also the *Full specification of individuals for the GA and their genes coding* table below) on the one hand, he search space is enormous: it has 2^{129} possible individuals and the dimesion three landscape given by φ, λ and F it has already a lot of very narrow local minima. We cannot imagine the complicacy of the landscape in dimension 7, taking into account the fact that some parameters (such as φ) are highly sensitive, while others, such as $x(0)$ and δ, have milder effects on the solution generated by the model.

These considerations tell us that finding a solution candidate to Problem (6.2) is really difficult and, as it has been said, it must be done with heuristic highly exploratory algorithms. However, as one can clearly see by looking at the results of the first batch of executions, it is good to "anchor" the now-not-so-random initial population to the fittest individuals when some is discovered. This adds an "elitist" ingredient to our algorithm for efficiency in local search. The fittest individuals just described are called the Superior ones, and the implementation of this idea gives a new meta-algorithm described in pseudo-code in the Box 2 below:

Box 2. **GA in Recurring Batches with Initial Population Reinforced by the Set of Superiors**

```
BestSuperior ← Fittest individual from Sparse Anisotropic Grid Search
BestSuperiorFitness ← F(BestSuperior)          ▷ Set of Superiors best fitness initialization
FF ← {BestSuperior}                            ▷ Set of Superiors initialization
while true do
    batchsize ← desired batch size for current iteration
    bestbatchfitness ← ∞                       ▷ Best batch fitness initialization
    for b ← 1 to batchsize do
        popsize_b ← desired population size for the b–batch iteration   ▷ Must be even
        P ← {}                                 ▷ Initializing empty first generation
        for popsize_b times do
            P ← P ∪ {new random individual}
        end for
        for each individual w ∈ FF do
            P ← w at a random place   ▷ Individuals from the Set of Superiors added at random places of P
        end for
        Best ← GeneticAlgorithm(P, popsize_b)   ▷ Standard Genetic Algorithm with P as initial population
        if F(Best) < bestbatchfitness then      ▷ Computing the best individual in the whole batch
            bestbatchfitness ← F(Best)
            BestInBatch ← Best
        end if
    end for                                    ▷ End of batch
    if bestbatchfitness < BestSuperiorFitness then   ▷ Updating the Set of Superiors, if necessary
        FF ← FF ∪ {BestInBatch}
        BestSuperior ← BestInBatch
        BestSuperiorFitness ← bestbatchfitness
    else
        return BestSuperior                    ▷ The End when there is no improvement
    end if
end while
```

Observe that in the above algorithm the Set of Superiors is formed by the fittest individuals of every batch and it is nitialized to the best phenotypic individual found in the Sparse Anisotropic Grid Search. Observe also that for every run of the *Standard Exploratory GA* in a batch, a single instance of every Superior individual is added to the now-not-so-random initial population at a random place (in particular a Superior individual can replace another Superior individual previously added to the initial population). In other words, the random initial population of every *Standard Exploratory GA* is anchored to the "optimal search space zone" by means of the Set of Superiors.

6.3.8 The results

In the table below we explain the execution flow of our *GAs in Recurring Batches with initial population reinforced by the Set of Superiors* which consists in the Sparse Anisotropic Grid Search and 6 batches.

The Set of Superiors of every batch is the result of the SAG search and a Best Batch Individual from every one of the previous batches (in the case of batch 4 we add two Superiors to the set instead of one because this batch gave a lot of better fitted individuals). Every row shows the best result from the batch (i.e. Best Batch Individual): columns 2–7 show the individual's phenotype and column 8 the individual's fitness (i.e. Least Squares norm). The last batch (number 6) is used only as a stopping condition, i.e. to check that Best Batch Fitness does not improve.

Remark 6.11. The fact that in the last batch (i.e. for 11000 runs of the algorithm) we obtain a *unique Best Batch Individual*, and these coincide *for all population sizes*, tells us that *probably* the result we have found is the true global optimum of Problem (5.1).

Next, welow we show, for the optimum given in the last row of the table below (with blue background), the data prediction from year 2006 to 2017 with a picture of the data fitting. The shape of the dispersal term $\lambda D(x(t), \mu, \sigma, \delta) = 1850.9189 \times D(x(t), 0.00131428, 0.207329, 10851.606)$, is also displayed below the data fitting.

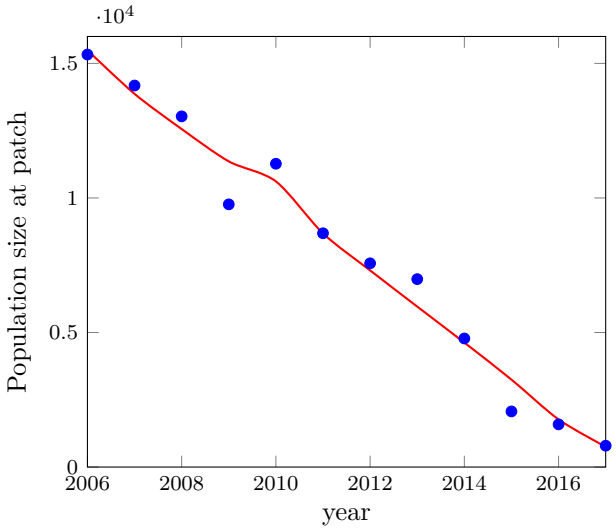
First Superior: the output of the Sparse Anisotropic Grid Search (Lemma 6.2)							
Number of Fitness evaluations	$x(0)$	φ	λ	μ	σ	δ	Fitness
14,988,610,560	15800	0.22	1700	0.3	1.0	8720	2537.417627

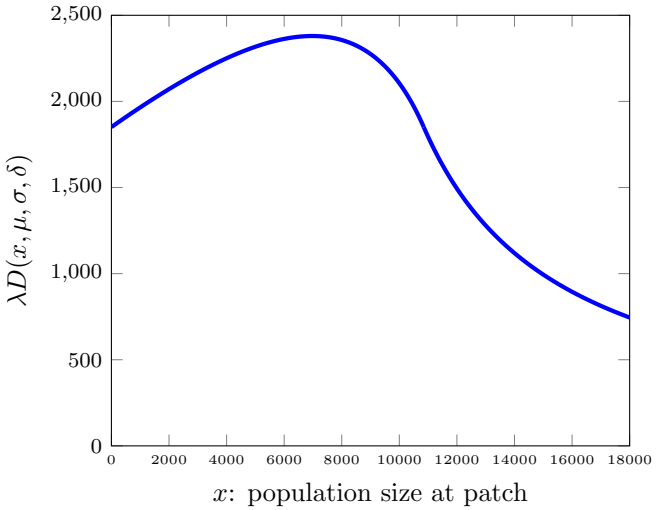
The rest of the Set of Superiors: a constant improvement of Best Batch Fitness

Batch #	Batch size	$x(0)$	φ	λ	μ	σ	δ	Best Batch Fitness
1	5500	15666.5193	0.250033585091	1628.8779	0.00006859	0.482179	8761.425	2473.513453
2	5500	15660.2552	0.254830499985	1649.3537	0.00081102	0.430394	8806.592	2470.298789
3	7700	15539.5629	0.312792844321	1872.6615	0.06318748	0.195313	10703.288	2468.449046
4	5500	15489.7795	0.314718281218	1829.1813	0.00075221	0.216484	10841.230	2430.918962
		15489.1114	0.317495907002	1858.0257	0.01334488	0.205422	10850.385	2431.153957
5	5500	15489.3244	0.317559956641	1850.9189	0.00131428	0.207329	10851.606	2427.004082
6	11000	15489.3244	0.317559956641	1850.9189	0.00131428	0.207329	10851.606	2427.004082

Table 6.2: A full account of the execution flow of the algorithm, together with the building of the final Set of Superiors. Every batch is divided into 11 sub-batches of equal size *batchsize*/11. In each sub-batch the population size is constant and ranges from 15000 to 25000 in steps of 1000. Note that from batch 4 we add two Superiors to the set instead of one. Every row shows the result (i.e. Best Batch Individual) obtained in the corresponding batch: columns 2–7 show the individual's phenotype and column 8 the individual's fitness i.e., Least Squares norm.

Year	Population Data	
	Observed	Predicted
2006	15329	15489.32
2007	14177	13867.81
2008	13031	12560.67
2009	9762	11358.90
2010	11271	10615.23
2011	8688	8682.89
2012	7571	7307.27
2013	6983	5962.40
2014	4778	4625.06
2015	2067	3249.43
2016	1586	1769.00
2017	793	754.32





Supplementary Section 7

A change in the tendency of gulls' population increase at the onset of perturbation

The aim of this section is to explore the change in the tendency of gulls' population increase coinciding with the onset of the perturbation, when predators arrived at the patch. This section has been placed after the analyses and computations carried out for the collapse phase because the fitting after the onset of the perturbation has been carried out using the parameters of the Elliot function obtained in Section 6. Despite the amount of data for this period is very limited since the decline of the population ranges from 1998 to 2004, we will analyse this period of time considering dispersal. Before doing so, we will explore the period from the establishment of the population in 1981 to 2004, in order to see how a logistic model may provide or not a good fitting of the population dynamics for this period. Figure **S7.1** displays the dynamics predicted by Eq. (3.1) until 2004 using the structural population parameters estimated in Section 4.2. Notice that by extending the time series until 2004 using the estimated values of the initial phase, the field data after 2007 clearly deviates from the dynamics obtained with the parameter values before the arrival of predators.

Year	Population Data	
	Observed	Predicted
1997	11725	12096.688873
1998	11691	12674.456135
1999	10189	13116.386478
2000	10537	13447.142449
2001	11666	13690.683293
2002	10122	13867.859101
2003	10355	13995.627921
2004	9168	14087.185201

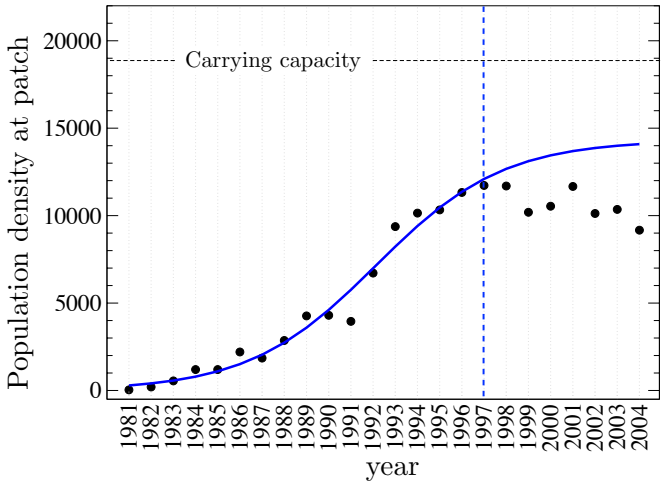


Figure S7.1: (Left) Predictions for the period 1981-2004 using Eq. (3.1) with the estimated structural parameters given in Section 4.2. (Right) Dynamics obtained with the structural parameters (blue line). Notice the change in the tendency at the onset of the perturbation.

Next, we fit the model to the period 1981-2004 using the same methodology than in Section 4.2. Here, we have used the value of β obtained in the initial phase ($\beta = 2.43826\cdots\times 10^{-5}$), leaving as free parameters the initial condition and α. We have done this way to allow the population to decrease towards lower population values since $\alpha = \gamma - \varepsilon$. The best fit has been obtained for $x_0 = 603.14051$ and $\alpha = 0.293142$. However, the error is much higher than the one obtained for the initial phase (i.e., 2593.0536), now given by 5639.340402 for the period 1981-2004. The predicted versus the observed values are displayed in Fig. **S7.2**.

Finally, we will fit the period 1998-2004 taking into account dispersal (both linear and by social copying).

Year	Population Data	
	Observed	Predicted
1981	36	603.140507
1982	200	795.004531
1983	546	1042.331142
1984	1200	1357.299582
1985	1200	1752.255749
1986	2200	2238.022399
1987	1850	2821.459603
1988	2861	3502.548751
1989	4266	4271.719872
1990	4300	5108.525121
1991	3950	5982.725025
1992	6714	6858.134632

Year	Population Data	
	Observed	Predicted
1993	9373	7698.370369
1994	10143	8472.663059
1995	10327	9159.866941
1996	11328	9749.725865
1997	11725	10241.674833
1998	11691	10642.218530
1999	10189	10962.004175
2000	10537	11213.338500
2001	11666	11408.448034
2002	10122	11558.462699
2003	10355	11672.955468
2004	9168	11759.845444

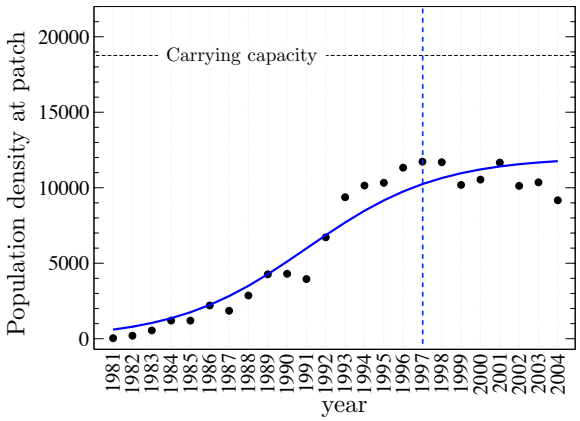


Figure S7.2: (Left) Predictions for the period 1981-2004 obtained from the best fit of Eq. (3.1) to the field data taking the value β estimated from the initial phase and leaving free $x(0)$ and α. (Right) Dynamics obtained for the best fitting (blue line). The fitting of the dynamics from 1981 to 2004 results in a least-squares error = 5639.340402. Notice that the predicted initial condition largely departs from the observed one.

Here, as well, we do not aim at providing an exhaustive fitting for this short period of time, since the field data are scarce, but evaluate the tendency evaluating the weight of exponential dispersal versus dispersal by social copying. We want to emphasise that the period of interest is the local collapse observed from 2006 to 2017 investigated in Section S6.

To fit the dispersal by social copying we will use the parameters for the Elliot function $D(x, \mu, \sigma, \delta)$ obtained from the collapse period. We are doing so for two reasons. First, we are assuming that the shape of the dispersal function by social copying is an inherent trait of this species and thus the values of μ, σ and δ will be approximately constant. Second, we are only leaving λ as free parameter to avoid overfitting, since the Elliot function has three parameters. Figure **S7.3** displays the best fit obtained for $\varphi = 0.2960168$ and $\lambda = 685.2774$. Also, the initial condition for this period must be $x(16) = x(1997) = 12096.688873$. More precisely, the model we consider for the period 1998-2004 is

$$\frac{\mathrm{d}}{\mathrm{d}t} z(t) = \varphi z(t) - \beta z(t)^2 - \lambda D(z(t), \mu, \sigma, \delta) \tag{7.1}$$

with the following parameters

Parameter	Range or value	Ecological meaning or description
α	0.348949408396403	Population growth rate including death of individuals (without linear dispersal)
β	$2.43826353653 \times 10^{-5}$	Intrinsic growth rate over the carrying capacity
K	18822.7975	Carrying capacity
$z(0)$	$x(16) = 12096.688873$	Initial condition
μ	0.00131428	Tendency of dispersal function for small population sizes
σ	0.207329	Sharpness and smoothness of the dispersal function
δ	10851.606	Transition between small and large population sizes
ρ	$\mathbb{R}^+$	Linear (exponential) dispersal rate
$\varphi = \alpha - \rho$	$(-\infty, \alpha]$	Population growth rate including linear dispersal
λ	$\mathbb{R}^+$	Dispersal rate

The solution of the above Model (7.1) with initial condition $x(16) = x(1997) = 12096.688873$ will be denoted by $z(t) = z_{\varphi,\lambda}(t)$. Clearly, $z_{\varphi,\lambda}(0) = x(16) = 12096.688873$.

Observe that the fitting in this setting will consist in estimating two parameters: the population growth rate including a possible linear dispersal φ, and the dispersal rate λ. In a similar way as before we define the map

$$\begin{aligned} \mathsf{T}: \quad \mathscr{K} &\longrightarrow \mathbb{R}^+ \\ (\varphi, \lambda) &\longmapsto \sqrt{\textstyle\sum_{t=1}^{7} \big(z_{\varphi,\lambda}(t) - \psi(t)\big)^2}, \end{aligned}$$

where $\mathscr{K} := [-\infty, \alpha] \times \mathbb{R}^+$, and

$$\psi(t, t = 1:7) = \mathsf{Audouin's\_Gulls\_Observed\_Population\_at\_year}(1997 + t, t = 1:7) = \overset{[1]}{[}11691, \overset{[2]}{10189}, \overset{[3]}{10537}, \overset{[4]}{11666}, \overset{[5]}{10122}, \overset{[6]}{10355}, \overset{[7]}{9168}].$$

Now, the *fitting of the model* consists in solving

$$\begin{aligned} &\min \mathsf{T}(\varphi, \lambda) \\ &\text{subject to } (\varphi, \lambda) \in \mathscr{K}, \\ &\text{and } z(t) \in [0, K] \text{ for } t = 1, 2, \ldots, 7. \end{aligned} \tag{7.2}$$

To solve this problem again we have used a standard *trust region method* with the Levenberg-Marquardt algorithm to solve the trust region sub-problem (see the GNU Scientific Library (GSL) *Nonlinear Least-Squares Fitting* documentation). As before, we have used numerical approximation of derivatives of the objective function. The obtained results are given by $\varphi = 0.29601680927544571$ and $\lambda = 685.2774$, with error $\mathsf{T}(\varphi, \lambda) = 1709.008415\cdots$. Despite the low amount of data, these migrations rate also indicate that the dominant dispersal is due to social copying. The model fitting for this short period of time is shown in Fig. **S7.3** below.

Year	Population Data	
	Observed	Predicted
1997	11725	12096.688855
1998	11691	11612.937619
1999	10189	11202.812016
2000	10537	10841.736696
2001	11666	10514.650524
2002	10122	10218.755097
2003	10355	9952.316758
2004	9168	9712.808050

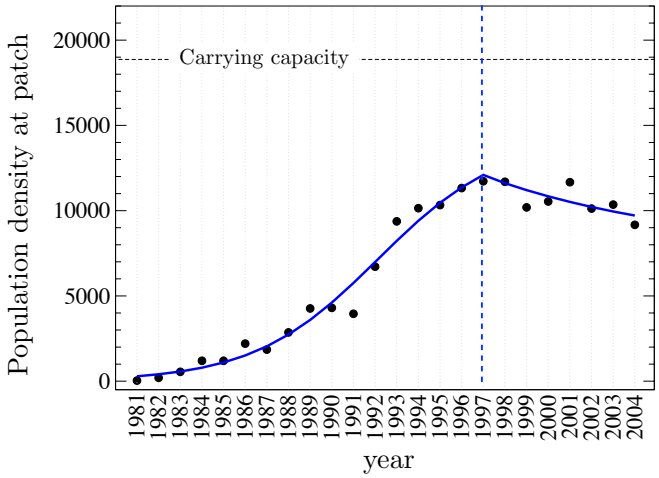


Figure S7.3: (Left) Model fitting for the period 1998-2004 obtained using Eq. (3.1) taking the value β estimated from the initial phase and taking the value from 1997 as initial condition. Here, we have left free $x(0)$, φ, and λ (using the values of μ, σ, and δ obtained with the best fit for the collapse phase). Here we have assumed that the shape of the function $D(x(t), \mu, \sigma, \delta)$ is a trait of this species. (Right) Dynamics obtained for the best fitting (blue line). The fitting of the dynamics from 1998 to 2004 results in a least-squares error = 1709.008415.

Bibliography

[1] Bartumeus, F. *et al.* Fishery Discards Impact on Seabird Movement Patterns at Regional Scales. Curr. Biol. 20, 215–222 (2010).

[2] Oro, D., Ruiz, X. Seabirds and trawler fisheries in the northwestern Mediterranean: differences between the Ebro Delta and the Balearic Is. areas. ICES J. Mar. Sci. 54, 695–707 (1997).

[3] Martínez-Abraín, A., Oro, D., Forero, M. G., Conesa, D. Modeling temporal and spatial colony-site dynamics in a long-lived seabird. Popul. Ecol. 45, 133–139 (2003).

[4] Fernández-Chacón, A. *et al.* When to stay, when to disperse and where to go: survival and dispersal patterns in a spatially structured seabird population. Ecography 36, 1117–1126 (2013).

[5] Payo-Payo, A. *et al.* Colonisation in social species: the importance of breeding experience for dispersal in overcoming information barriers. Sci. Rep. 7, 42866 (2017).

[6] Almaraz, P., Oro, D. Size-mediated non-trophic interactions and stochastic predation drive assembly and dynamics in a seabird community. Ecology 92, 1948-1958 (2011).

[7] Oro, D., Pradel, R., Lebreton, J.-D. Food Availability and Nest Predation Influence Life History Traits in Audouin's Gull, Larus audouinii. Oecologia 118, 438-445 (1999).

[8] Payo-Payo, A., Genovart, M., Bertolero, A., Pradel, R., Oro, D. Consecutive cohort effects driven by density-dependence and climate influence early-life survival in a long-lived bird. Proc R Soc B 283, 20153042 (2016).

[9] Tavecchia, G., Pradel, R., Genovart, M., Oro, D. Density-dependent parameters and demographic equilibrium in open population. Oikos 116, 1481-1492 (2007).

[10] Genovart, M., Oro, D., Tenan, S. Immature survival, fertility, and density dependence drive global population dynamics in a long-lived species. Ecology 99, 2823–2832 (2018).

[11] Payo-Payo, A. *et al.* Predator arrival elicits differential dispersal, change in age structure and reproductive performance in a prey population. Sci. Rep. 8, 1971 (2018).

[12] Oro, D., Ruxton, G. D. The formation and growth of seabird colonies: Audouin's gull as a case study. J. Anim. Ecol. 70, 527-535 (2001).

[13] Oro, D. Living in a ghetto within a local population: An empirical example of an ideal despotic distribution. Ecology 89, 838-846 (2008).

[14] Genovart, M. *et al.* Varying demographic impacts of different fisheries on three Mediterranean seabird species. Glob. Change Biol. 23, 3012–3029 (2017).

[15] Genovart, M., Doak, D., Igual, J.M., Sponza, S., Kralj, J., Oro, D. Varying demographic impacts of different fisheries on three Mediterranean seabird species. Glob. Change Biol. 23: 3012–3029 (2014).

[16] J. Holland, Hidden Order: How Adaptation Builds Complexity, Helix Books (1996)