

Who is afraid of modelling time as a continuous variable?

Hanna Kokko, University of Mainz, [hkokko@uni-mainz.de](mailto:hkokko@uni-mainz.de)

*Remember to acknowledge Jim McLean & lab members. Consider adding Gil03 and Mye18 as refs (stability of pred-prey cycles). And deal with editor's initial reaction to this which reads: "Two points for me could be added:*

*1-one summary figure compiling the approach in an appealing way. HK: yup. also, the fig will explain  $r=\text{rand}(1)$ ;  $\tau=-\log(r)/\text{sum}(\text{rates})$  as a fast alternative to  $\text{exprnd}()$ .*

*2- adding an empirical example with real data to go in the ms after the simulations or in suppl. in order to appeal to the most empiricists out there (!) but also to add more ecology in the mix. – HK: This'll be about explaining in much more detail what Deng 2018 achieved"*

*from Shikhara: Not sure if this is worth mentioning in the manuscript, but I just remembered that R provides at least two packages to run Gillespie simulations 'out of the box' (GillespieSSA and GillespieSSA2) that could be quite useful for empiricists. You actually cite the paper associated with the former package in the introduction, but without mention of the package.*

*2 comments (Zug & Shikhara) are as files in the folder, file dates 12.1.*

## Abstract

Almost all models of ecological and eco-evolutionary processes involve creating trajectories of something (be it population densities, average trait values, or environmental states) over time. This requires decision-making regarding how to represent the flow of time in models. Most ecologists are exposed to continuous-time models (typically in the form of ordinary differential equations) as part of their training, especially since the famous Lotka-Volterra predator-prey dynamics is formulated this way. However, few appear sufficiently well trained to produce their own work this way, and may lack exposure to the true versatility of methods that are available. Specifically, knowledge that discrete individuals can be modelled in continuous time using the Gillespie algorithm is not as widespread as it should be. In this review I will illustrate the flexibility of continuous-time modelling methods such that researchers can make informed choices, and not resort to discretizing time as a 'default' without a clear biological motivation to do so. I provide three example-based tutorials to help interested researchers get started: (1) a comparison of deterministic and stochastic dynamics of the Lotka-Volterra predator-prey model, (2) an evaluation of matelessness in a hypothetical insect population (and the consequences for

selection to mate more often by either searching more efficiently or by shortening the ‘time out’ after each mating), and (3) an example of within-season density dependence followed by a birth pulse leading to Beverton-Holt or Ricker dynamics depending on whether the deaths of conspecifics help reduce the mortality of others or not (compensatory mortality). I will also highlight properties of the exponential distribution that may be counterintuitive yet good to know when using the Gillespie algorithm or when deriving expected lifetime reproductive success or other similar quantities. I also show how to proceed if the so-called memorylessness assumption (Markov property) does not hold in a given situation, and show how continuous and discrete times can be freely mixed if the biological situation dictates this to be the preferred option.

## 1. Introduction

Anyone interested in modelling the ecological or evolutionary dynamics of populations has to make methodological choices. The treatment of time is a crucial one. Should one divide time into discrete steps, where various simultaneous changes are made happen, and time is then updated to a next step, where new simultaneous events can occur? Or, should one let time flow continually, so that it is possible to ask about the population state at, say,  $t = 13.67$  (instead of only at  $t = 13$  and  $t = 14$ )?

My aim in this review is to help those who may have had no explicit exposure to the choice of how to deal with time. More often than not, the training encountered by individuals within my target group directs them to discrete-time choices, perhaps without awareness that there could be alternatives (even though good starting points to broaden one’s mind are published once in a while, e.g. Pineda-Krch, 2010; Black & McKane, 2012; deLong & Gibert, 2016; deLong & Cressler, 2022). My target group are researchers who mainly work empirically but might also sometimes build an individual-based simulation or need to quantify the change of an allele frequency from one generation to the next, or young scientists who are still discovering methods and are not sure if they want to head down the route of becoming theoreticians. To help them get started, I will provide examples rather than mathematical theorems, and apologise to those readers who are already familiar with differential equations and/or Markov processes and the Gillespie algorithm: to them, the treatment will probably feel rather basic and perhaps verbose.

While there is no inherent ranking order so that one way to treat time would be superior to the other, it helps to be fully informed of just how flexibly time can be modelled, either in deterministic models or stochastic simulations. Time, after all, in reality flows continually, and many biological processes do not occur precisely synchronously. Forcing them to do so may

1 make it difficult to appreciate some processes: e.g., if a modeller forces all potential pairings  
2 between males and females to happen during one time step, it becomes cumbersome to include  
3 limited mate-searching capacity into a model. Thus, my aim is to provide an example-based  
4 tutorial to just how flexible continuous-time modelling can be in ecology.  
5 The two topics that I will provide example-based guidance on are ordinary differential equations  
6 (often abbreviated ODEs) and the Gillespie algorithm, together with variations thereof. The  
7 methods themselves have been available for centuries (Leibniz 1684, Newton 1687) and decades  
8 (Gillespie 1977), respectively. But their versatility is, to this date, underappreciated in ecology.

10 In ODEs and their stochastic simulations using the Gillespie algorithm, events can be almost  
11 anything (Masuda & Vestergaard, 2022). They may refer to a birth or a death. They can also  
12 indicate that an individual is changing its state, perhaps by migrating to a new location (Méndez  
13 et al., 2014; Abdullahi et al. 2019), by engaging in a mating (Ekrem & Kokko, 2023), or by  
14 becoming infected (Vestergaard & Génois, 2015; McCullough et al., 2016). The entities that  
15 change state do not have to be individuals. In Gillespie algorithms involving infections, even an  
16 entire species may become infected, which happens when *Wolbachia* jumps from one species to  
17 another (Zug et al., 2012). Likewise, states may refer to the state of an environmental variable  
18 that the species needs to adapt to (Berrios-Caro et al., 2021).

20 Some readers who are familiar with the assumption structure of the Gillespie algorithm might  
21 consider it a downside that computational speed is achieved by restricting the attention to types  
22 of events in a different way: they must happen at a constant rate. The constancy of rates requires  
23 the system to have a property that is called memorylessness (Box 1): the time that has already  
24 elapsed since the last event is not allowed to have any impact, positive or negative, on the rate at  
25 which the next event happens. The technical term for the lack of memory in this special,  
26 mathematical meaning, is the Markov property, after the Russian mathematician Andrey Markov  
27 (and for the mathematically inclined reader, there is much to digest in the field of Markov  
28 processes, Markov chains, and the associated properties of the Poisson distribution; Masuda &  
29 Vestergaard, 2022, provides a good starting point and plenty of references).

31 The memorylessness feature, i.e. the demands that rates remain constant no matter how long  
32 ago the last event occurred, may sound like a rather severe limitation. Consider, for example, the  
33 ageing process of individuals. Here, as long as we assume senescence to take place, one should be  
34 able to assume that an individual whose birth is long in the past should be experiencing a  
35 different mortality rate compared with earlier (Golubev, 2004; Boonekamp et al., 2015; Hughes  
36 & Hekimi, 2016). Another example that seemingly makes it impossible to apply the Gillespie

algorithm involves phenological changes. If it has been warm for long, winter is due to arrive, so the probability of cold weather should increase over time and not stay constant. But the memorylessness is a limitation that can be very easily overcome: in an extended algorithm, events can also be prescheduled to occur. For example, all adult *Daphnia* that one is simulating might die once winter commences, leaving only dormant eggs. Mixtures of discrete and continuous processes can in reality be woven into any computer program with Gillespie-like features, and I will show how.

I also hope to clarify some common misconceptions that I often encounter in the anonymous setting of being at the receiving end of reviewer comments. Reviewers of papers may be puzzled about some statements, such as a mortality rate that can take a higher value than 1, or they might be taken aback by models that specify the number of mate encounters as  $vN_MN_F$ , where  $v$  is a positive constant and  $N_M$  and  $N_F$  are the numbers (or in some case densities) of males and females, respectively. At first sight this seems to imply that the number of matings in a time step can exceed the number of available individuals, but such misconceptions arise when the reader is used to thinking about time steps during which “a lot happens”. In continuous-time modelling, discrete time steps of 1 unit length are not the focus of interest at all. Some authors admittedly have used the phrase ‘time step’ (deLong & Cressler, 2022) when describing births and deaths in a special case of the Gillespie process, perhaps as a result of not being very interested in the precise durations of time between events. In deLong & Cressler (2022), every event that happens causes a time step to happen. But because the Gillespie algorithm can also keep track of the length of times between events, and these times are typically of highly variable length, I recommend avoiding using the language of ‘steps’ as this evokes a false feeling of a regular schedule and can thus confuse. Note though, that in all that follows, it is still possible — if desired — to ask about the state of the population at those time points that are integers, e.g. once a year, measured at a suitable part of the life cycle. All this will become hopefully clear below.

## 2. From ideal gases to furry (but discrete) lemmings and foxes: How should we teach the Lotka-Volterra predator-prey model?

It is hardly possible to complete an undergraduate degree in ecology without having seen Lotka-Volterra predator-prey cycles. The two names that are linked to this model might suggest a collaboration between two scientists, but instead, the joint name recognizes independent work of two scholars (Lotka, 1925; Volterra, 1927; they did subsequently become aware of each other, to some degree seeing each other as rivals, Kingsland, 1995). Lotka referred to his work as ‘physical biology’, inspired by methods that were in place to analyze chemical reactions (called *kinetics*). His insight was to realize that a change in the numbers of predators and prey as a result

of them meeting each other is quite analogous to the kinetics of molecules bumping into another, where the availability of the original molecules is reduced and the product of the reaction is added to the pool of the newly formed ones. Volterra similarly took molecules as inspiration, but his focus was on molecules in a gas and their encounters with each other — and his inspiration to apply this to biology seems to have resulted from the observation that the First World War, which led to starkly reduced fishing effort, made predatory species markedly more abundant in the sea (Kingsland, 1995).

For chemical kinetics, the choice of continuous (as opposed to discrete) time is rather obvious: there is no reason to expect reactions to spontaneously arrange themselves into discrete pulses. The chemistry-inspired background led both Lotka and Volterra to think about rates of change in the abundance, more specifically densities, of prey  $N_1$  and predators  $N_2$ . The rates of change apply for any value of time  $t$ , with no special focus on integer-valued times. (They also did not think of  $N_1$  and  $N_2$  as integers, and this is the more debatable decision; I shall return to it below). The rates of change in the abundance are denoted  $dN_1/dt$  and  $dN_2/dt$ , respectively. Positive rates imply an increasing population, negative a decrease, and zero means no change. Notation for the coefficients for the classic Lotka-Volterra dynamics varies from text to text, but there are four coefficients to be defined, e.g.,  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ . At the time of writing, the English-language Wikipedia page lists them like that, while e.g. Wikipedia in German replaces  $\alpha$  with  $\epsilon_1$ ,  $\beta$  with  $\gamma_1$ ,  $\gamma$  with  $\epsilon_2$  and  $\delta$  with  $\gamma_2$ ; it really does not matter. Using the  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$  notation, and denoting the density of prey with  $N_1$  and that of the predator with  $N_2$ , the two equations take the form

$$\frac{dN_1}{dt} = \alpha N_1 - \beta N_1 N_2 \quad (1a)$$

$$\frac{dN_2}{dt} = -\gamma N_2 + \delta N_1 N_2 \quad (1b)$$

For the examples I will discuss, I will, however, aim for a more biologically intuitive decomposition. Consider this version:

$$\frac{dN_1}{dt} = b B_1 N_1 - a N_1 N_2 \quad (2a)$$

$$\frac{dN_2}{dt} = -\mu N_2 + \frac{a}{n} B_2 N_1 N_2 \quad (2b)$$

While I have not encountered this particular formulation in any text before, it is a useful way to make students think about the biological meaning of the values chosen for the parameters. Here,  $b$  is the rate at which the prey can breed: if this rate is, for example, 2, then prey can breed twice in one time unit. The choice of units depends on the system, but we could, for example, think of a

vertebrate where a sensible unit might be 1 year; for a microbial population, appropriate choices might be much shorter, e.g., one unit might correspond to a day. It does not matter for anything else than for the ease of interpretation of the time axis. If we settle on a year as the unit, then the organism that can fit in 2 breeding attempts per year ( $b = 2$ ) — or we can equivalently state that it breeds, on average, every  $\frac{1}{2}$  time units, more familiarly expressed as every 6 months. Each breeding event is assumed to add  $B_1$  offspring to the population, and here the subscript reminds us that we are adding prey, not predators, who have their own brood size  $B_2$ .

The coefficient  $a$  specifies the number of (successful, from the predator's perspective) encounters between predators and prey. For  $b$  and  $B_1$ , it is rather easy to think about the sensible range of values, but how about  $a$ ? The expected total number of kills per time unit that we can observe in the system is  $aN_1N_2$ , and the relevant thought process is the following. If one marks a specific prey (say, lemming) individual and a specific predator (say, arctic fox) with tags that allow us to track their fate, and they co-occur spatially in an area where predator-prey encounters do happen, then  $a = 1$  implies that it would take, on average, 1 unit of time for a specific lemming to be found and killed by a specific fox. (Any other individual fox may also get our focal lemming, of course, or the focal fox could die before finding our focal lemming, but such 'competing events' are fully accounted for by the equations as a whole, as we will see). If  $a = 0.1$ , then the time it takes for this event to happen averages  $1/a = 10$  years: we are then dealing with situation where a randomly chosen lemming and a randomly chosen fox are not that likely to meet, at least when we model fox lifespan realistically.

We must also briefly discuss the parameters that I chose for the predator (equation 2). The value of  $\mu$  gives the predator's death rate, and rather weirdly, it is not assumed to increase if there is little to eat. This is also a topic I will return to. Instead, the total rate of observable deaths is  $\mu N_2$ , and from the perspective of a focal predator, the death rate is a constant (of the  $\mu N_2$  deaths, a proportion of  $1/N_2$  concern one particular predator, thus the per capita death rate is  $\mu$ ). The deaths obviously contribute negatively to population growth, while births contribute positively. In the classic interpretation, the predator's hunting success is assumed to improve the rate at which new predators enter the system (via births), but has no effect on the death rate. The fact that the death rate  $\mu$  is not sensitive to prey abundance means that predators disappear rather sluggishly from the system if there is nothing to eat, at least if one has chosen  $\mu$  to give lifespans (of magnitude  $1/\mu$ ) that work well under normal conditions — i.e. with the assumption that prey are available — for a specific predator species. In the traditional way of presenting Lotka-Volterra equations, this weirdness goes easily unnoticed, because neither predators nor prey go

1 extinct; but once we have moved on to a stochastic version, this problem can no longer be swept  
2 under the carpet.

3  
4 We are, however, still rather far away from discussing the stochastic version. We must first finish  
5 discussing the choices of all parameters in the system and thereafter understand the  
6 deterministic expectations. To continue on the first task, let us look at the coefficient  $(a/n) B_2$ ,  
7 that I chose to represent the coefficient  $\delta$  when moving from eq. 1b to 2b to make it more  
8 ecologically interpretable. Here, I intentionally re-use the very same  $a$  that already appeared in  
9 the prey equation (2b): an attack, after all, impacts both predators and prey, so if, say, the rate at  
10 which predators find prey improves (e.g., they evolve better sensory systems), this should be  
11 reflected in both eq. 2a and 2b. However, it would be an odd choice to use the exact same  
12 coefficient value ( $a$  without any modifications) in front of the  $N_1 N_2$  terms for both predators and  
13 their prey. This would mean assuming that a new predator arises every time a prey is captured;  
14 but real arctic foxes do not produce a new cub every time their foraging succeeds. Much more  
15 food is required for that to occur. The  $a/n$  term therefore denotes the rate of reproduction, with  $n$   
16 giving the mean number of prey items that need to be captured for each new brood of predators.  
17 When the predators breed, a brood of  $B_2$  predator are added to the system.

## 19 2.1. Different perspectives on kills — all of them exponential

20 The above rates can be thought about from different perspectives. There is, as a whole, a rate  
21  $a N_1 N_2$  at which kills occur in the population. But for a particular lemming, deadly events happen  
22 at a rate  $a N_2$  (a total rate  $a N_1 N_2$  is divided among  $N_1$  unique prey individuals, which leaves  $a N_2$ ).  
23 Should we be interested in this particular kill being performed by a specific fox, we need to  
24 dividing the rate further by  $N_2$ , which — correctly — recovers the result for the rate at which a  
25 particular fox finds a particular lemming, which is simply  $a$ .

26  
27 Importantly, these rates can be translated into expected times to observing something in the  
28 system. These rates define how the luck of avoiding being predated on, i.e. the probability that a  
29 lemming is still alive (its survivorship), behaves over time. This time, let us start from the  
30 smallest-scale perspective: If we observe a particular lemming for  $t$  units of time (e.g. a month as  
31 the observation duration would imply  $t = 1/12$ , given our interpretation of one time unit as a  
32 year), the probability that the lemming has avoided a specific fox all that time is  $e^{-at}$ . Broadening  
33 the perspective somewhat, the probability that this one lemming has avoided the risks that arise  
34 through the entire population of foxes is  $e^{-a N_2 t}$ , and the probability that in the whole population  
35 no deaths have yet occurred, as a result of any lemming meeting its end through any fox, is

1  $e^{-aN_1N_2t}$  (Figure 1). The last probability plummets very quickly to zero (Figure 1a) compared  
2 with the rest, but all of them obey the same shape of an exponential decline (Figure 1b-d).

3  
4 These examples offer insights that generalize. When an event is occurring at a rate (also called  
5 hazard)  $\lambda$ , then the time to its occurrence is exponentially distributed with parameter  $\lambda$ . Note  
6 that mathematical texts often use the language of ‘hazards’ for rates, even though the events can  
7 be beneficial, at least from the perspective of a fox finding a meal. The language of hazards is an  
8 understandable choice when modelling survival, e.g., a piece of equipment may fail after a time,  
9 but it is useful to know that the language of hazards does not restrict the validity of results to  
10 events with negative consequences.

11  
12 A property of the exponential distribution is that its mean is  $1/\lambda$ . Thus, the mean time that one  
13 has to wait for the event to occur is  $1/\lambda$ . The list of above perspectives, each taking a broader  
14 look at the situation than the previous one, correspond to  $\lambda = a$ ,  $\lambda = aN_2$ , and  $\lambda = aN_1N_2$ ,  
15 respectively. But one does not have to work in a nested manner: We can deviate from it to ask  
16 questions from the perspective of an individual fox. How long does it have to wait for a meal?  
17 Now the capture events, that occur at a rate  $aN_1N_2$  in the population as a whole, are divided  
18 among all predators  $N_2$ . The resulting rate  $aN_1$  implies that our focal fox finds a meal, on average,  
19 every  $1/(aN_1)$  time units. Similarly, we need to wait on average  $1/(\mu N_2)$  time units to observe the  
20 death of a fox in the population, or  $1/\mu$  to observe the death of a specific focal fox.

## 22 2.2. Deriving the deterministic expectation: towards limit cycles

23 Lotka-Volterra equations are usually considered in their deterministic setting, meaning that we  
24 ignore stochasticity. This was also true for the original use of the equations. If one is a chemist  
25 dealing with a gas (recall Volterra’s source of inspiration), one can trust that the number of  
26 molecules in any macroscopic volume of gas is a very large number, and consequently, one does  
27 not need to worry too much about the inherent stochasticity in the timing of events. The law of  
28 large numbers guarantees that the error between prediction and observation will be small if we  
29 simply track the expected change in the number of the relevant molecules — or, in ecology, the  
30 predators and their prey. This thinking underlies ecology textbooks’ tradition of asking the  
31 students to calculate single values for  $dN_1/dt$  and  $dN_2/dt$  for each setting, without emphasis on  
32 any kind of stochasticity.

33  
34 However, should the reader now ask in his/her mind “but predators and prey do not come in as  
35 large numbers as molecules in a gas flask”, this is an extremely good point. In Figure 2 I have



therefore plotted both the deterministic dynamics, which ignores all stochastic effects, and an example of an outcome of a stochastic simulation, conducted using the Gillespie algorithm: this is the stochastic version of the kinetics, based on finding out, after each event (death or birth of predators or their prey), the identity and timing of the next individual event.

Most ecology textbooks plot the limit cycles that result from the equations (Figure 2a) but not the stochastic counterparts, even though real data (the famous hare-lynx cycles, Deng (2018) and references therein) is usually shown alongside the limit cycles. The phrase ‘limit cycle’ means that starting from any value will make the system return to the same point; neither predators nor prey go extinct. This particular result will cease being true in the presence of stochasticity, which will allow us to discuss the differences between deterministic and stochastic dynamics. But before comparing, we need to know how to get from Eq. (1) to the plot of the deterministic dynamics. It is possible to find so-called *analytical* solutions for the limit cycle, meaning that one can ask, for any  $t$ , what the value of  $N_1(t)$  and  $N_2(t)$  will be, if the values at an earlier point, for example  $t = 0$ , are known. The solutions are rather cumbersome, however (Evans & Findley 1999), and ecology textbooks tend to divert attention of students away from asking such questions. *Numerical* solutions for the deterministic dynamics are much easier to find, which means plotting the solution for very many different, closely spaced values of  $t$  and ‘joining the dots’ (interpolating). These can be derived for any accuracy, such that the outcome no longer looks jagged.

### 2.3. Euler’s method: why considering a small enough $dt$ will remove some but not all of our worries

The simplest, easiest to understand, but not the most accurate method for finding numerical solutions to differential equations is that devised by Euler (1768). The method is based on a thought process that involves assigning a small value to the  $dt$  in equations 1 and 2, where it represents an infinitesimally small duration of time (mathematically speaking, a limit, when small durations are diminished even further to approach zero). A small value is not infinitesimally small but (so the hope goes) sufficiently close, so that we can estimate how much our variables of interest change if we extrapolate from the currently valid rates of change into the very near future. With this interpretation of  $dt$  in mind, we can multiply both sides of eq. 2a-b with  $dt$ , to compute the change in  $N_1$  and  $N_2$ :

$$dN_1 = (bB_1N_1 - aN_1N_2) dt \quad (3a)$$

$$dN_2 = \left(-\mu N_2 + \frac{a}{n} B_2 N_1 N_2\right) dt \quad (3b)$$

The key idea in numerical solutions to differential equations is that over very small amounts of time, the change in numbers is proportional to the time elapsed. Let us consider  $a = 1$ ,  $N_1 = 1000$  and  $N_2 = 10$ . In such a population, we expect one kill approximately every 53 minutes. This is the inverse of approximately 0.019 lemming kills per minute, obtained by applying  $dt = 1/525,960$  in the expression  $aN_1N_2dt$ . Why  $1/525,960$ ? A year has on average 365.25 days, taking into account  $\frac{1}{4}$  of them being leap years — seen this way, one minute is  $1/525,960$  of an average year. This ignores more complicated leap year rules that happen every century or even less often, as they make no difference to the point I wish to make here. One could equally simply work on expected kills during one millionth of a year, whether this corresponds to an integer number of minutes or not.

Observing the system for two minutes would yield an expectation of 0.038 observed kills of lemmings, which is twice the expectation for one minute. But since both 1 and 2 minutes are longer than infinitesimal, all statements of this type are approximations. They are based on the quantity  $aN_1N_2$  remaining constant (at 1000 for our example) for the duration  $dt$ , even though the numbers are ever-changing. Kills, after all, are supposed to remove individuals, and there are also other changes that can happen during the minutes in question (eq. 3). The point of the Euler method is that if  $dt$  is sufficiently small, the error remains small as well. If  $dt$  is large, the approximation works very badly; for  $dt > 0.1$ , this particular example would have more lemmings being eaten than exist to begin with.

The fact that an approximation behaves badly if  $dt$  is large has an upside too: no-one with good training will expect her audience be satisfied with an approximation that has been derived with  $dt = 0.1$ , let alone  $dt = 1$ , which would amount to an attempt to go through an entire 'time step' by assuming that numbers do not change in between. Conversely, this should remove the worry of reviewers who express concern that a high rate (values higher than 1, including much higher than 1) appears to predict impossibly many encounters between predators and prey. A high rate simply means that the time between two encounters is very small. Thus, prey become depleted quickly, and the new rate, always computed as  $aN_1N_2$ , diminishes in pace with the diminishing prey population. By the time that one unit of time has flown, little remains of the original prey population, and the value of  $aN_1N_2$  no longer predicts many encounters to happen. All we need to do in principle is to update frequently. Updating populations using small enough  $dt$  values and repeating the process until the sum of all  $dt$  values is a sufficiently long timespan to be of interest is the gist of the Euler method.

But while Euler's idea is easy to follow (and thus easy to teach), it is no longer seen as more than a good initial setup to develop intuition. For serious work, Euler's method, which predates the invention of computers by centuries, is not a good enough approximation. A problem very relevant for the Lotka-Volterra example is that even if  $dt$  is small, it still means not updating the population continually enough. Graphically, this problem means that the Euler method is prone to 'miss the turn' whenever rapid changes in direction are the real property of the dynamics (exemplified in Figure 2a: the purple arrow shows where the prediction, based on extrapolation, would end up). Arrows as long as that in Figure 2a, making the prediction go massively off the true curve, would admittedly only happen if one used an exaggeratedly improper (long) value of  $dt$ . But smaller  $dt$  would still cause arrows to point somewhat too much outwards, and the error obviously then accumulates over time, if each arrow's endpoint is taken as the new starting point for the next  $dt$ .

The above is a well known limitation of the Euler method of solving differential equations. Mathematicians have therefore worked out far better numerical methods to solve the behaviour of a system of ODEs, and most programming languages offer ready-made functions to solve them to a desired accuracy (e.g. `deSolve` in R, various functions starting with 'ode' in Matlab, and `scipy.integrate.solve_ivp` in Python). Since these better methods are abundantly available, and embedding them in programming code them requires no in-depth understanding of why they work so well, I encourage the reader to become familiar with them.

## 2.4. Gillespie: From expectations to stochastic outcome when time is continuous and real-life individuals are discrete

The numerical solvers of differential equations produce the deterministic expectation of the abundances of individuals in the system. This is often desired, especially since one can think of 0.019 lemmings removed per minute as a reduction in density, rather than implying that we have somehow shaven off 2% of one particular lemming. But there are also reasons to study the consequences of the fact that real lemmings and foxes are discrete entities. Specifically, one could reasonably ask if the arrow's tendency in Figure 2 to make the system jump from a limit cycle towards a slightly larger-amplitude limit cycle should be considered a real possibility in nature; do real foxes and lemmings obey the dynamic equations such that they never 'miss a curve', or could there be stochasticity that allows that to happen. There is also the worry that the larger-amplituded limit cycles make the populations go through periodic small numbers. The outermost limit cycle in Figure 2a has so few individuals left in some part of the cycle that their ability to recover could be disputed.

Real events make discrete individuals be born or die (or, in more general models, disperse to a different state, change sex, or grow to a different size class), and this means some stochastic jumps must occur. In principle, my above discussion of Figure 1 should help here: the curves look like something one could use to choose values for random variables denoting time to an observation, such as a lemming being removed from the population. This is also true: the time until an event occurs can be sampled from an exponential distribution. But so far, one might feel an unease of an unresolved issue with Figure 1b in particular. I was happily deriving an exponential decline of the survivorship of a specific lemming as a result of being eaten by a specific fox, apparently without worrying that one lemming can only be killed once. What if another fox took it first? In general, what about all the events that can be said to have rates (and thus predicted times of occurring) but never actually happen, because the situation changed in the meanwhile? It turns out that figure 1 has to be interpreted as information on survivorship in the absence of any other events interfering with the process. If other events can happen too, there are ways to deal with this, a topic I will now, finally, turn to.

#### 2.4.1. *Memorylessness when events compete to happen*

The property of memorylessness (the Markov property) now comes into play with full force. It can be applied whenever the time since the last event has no predictive value when trying to deduce which event happens next. For stochastic simulations, memorylessness generates a very neat mathematical outcome. When two kinds of events ‘compete to happen’ to an individual, but one (event type A) occurs at a rate  $\lambda_A$  while the other (event type B) at  $\lambda_B$ , we are technically required to compare random variables sampled from two exponential distributions: the time  $\tau_A$  that we have to wait for A to occur is drawn from an exponential distribution with parameter  $\lambda_A$ , and the time to B,  $\tau_B$ , is drawn from another exponential distribution, with parameter  $\lambda_B$ . The smaller of the two outcomes reflects the event that actually happens, as it was the one that is ‘scheduled’ to happen.

This sounds simple, but it gets even better: sampling from two exponential distributions, with parameters  $\lambda_A$  and  $\lambda_B$  respectively, and choosing the event that happened first, is mathematically equivalent to sampling just once from an exponential distribution with parameter  $\lambda$  where  $\lambda = \lambda_A + \lambda_B$ . This means that the wait time sampled from an exponential distribution,  $\tau \sim \text{Exp}(\lambda_A + \lambda_B)$ , will work as the scheduling for the next event, whichever type it turns out to be. And, *irrespective of the value of  $\tau$  that this sampling ‘spat out’*, we can trust that the event that occurs at that time is of type A with probability  $\lambda_A/(\lambda_A + \lambda_B)$  and an event of type B with the complementary probability  $\lambda_B/(\lambda_A + \lambda_B)$ . This is an insight that may sound too good to be true (Box 1 lists some

counterintuitive consequences); it is key to making the Gillespie algorithm possible. It also extends to more than two event types. Thus, if a lemming can give birth (event type A), or be eaten by a fox (event type B), or be eaten by a skua (event type C), the computation simply employs  $\tau \sim \text{Exp}(\lambda_A + \lambda_B + \lambda_C)$  and the production of another litter can be added to the lemming's life with probability  $\lambda_A/(\lambda_A + \lambda_B + \lambda_C)$ .

The above logic is not only useful in stochastic simulations, but also when the aim is to compute expressions for expectations, such as how often an individual with a specific set of traits can be expected to encounter mates or food items, or moves location, before something else of biological relevance (e.g. death) happens to it. It is not necessary to painstakingly track the timing of various events if the only relevant aspect is the probability that one more event, of some specified fitness consequence, can be added before the accumulation of fitness stops. In some cases, it may be particularly useful to know the expected time spent in a particular state. As an example, Gerber & Kokko (2016) derive results for an example where asexual reproduction is possible while a female of a facultatively sexual species is still virgin, and sexual reproduction commences once mated. Death can intervene before or after her first mating. Here, expected lifetime production of asexual eggs is simply proportional to the expected wait time until either death or mating happens; the expected lifetime production of sexual eggs, in turn, is the probability that it was mating rather than death that happened, multiplied by the expected lifespan (inverse of the death rate) thereafter. Should the egg production rate (e.g. per day) be  $\lambda_E$ , and the other rates are  $\lambda_M$  for mating (which was in Gerber & Kokko, 2016, made dependent on current sex ratio), and  $\mu_0$  and  $\mu_1$  for mortalities before and after the 1<sup>st</sup> mating, the expected number of asexual eggs is proportional to  $\lambda_E/(\lambda_M + \mu_0)$  and that of sexual eggs is  $(\lambda_M/(\lambda_M + \mu_0)) \cdot 1/\mu_1$ . Similar analytical expressions can be useful in very many contexts.

Of course, the memorylessness does not allow us to pretend that the world did not change once an event did happen. Once the numbers of predators and prey have changed, the rates that apply from now on are different too (or, in the case of the facultatively sexual female, we above specified that if she has mated, her death rate is now different from before; the same could have been done for egg production rate too). Memorylessness is our friend in a different way: it means that we can safely throw away the worry that one ought to keep track of what happened *before* the event. A worried researcher might think complicated thoughts of the style “because it took some time (of length  $\tau$ ) for the first event to happen, the others that were not chosen to happen this time are now somehow more ‘impatient’ or ‘due’ to occur, but I do not know how to compute this effect”. Nothing of this sort needs to be computed – except if there truly are prescheduled

events that violate the memorylessness condition; I will deal with them in a separate section of mating dynamics (section 3 below).

#### 2.4.2. Implementing the Gillespie algorithm for Lotka-Volterra predator-prey cycles

But first, the 1<sup>st</sup> example: a model of Lotka-Volterra with the Gillespie algorithm. Originally developed to model stochasticity in chemical reactions (Gillespie, 1977), I will explain the algorithm's logic here with a stochastic version of eqn (2a-b). I will now assume that  $N_1$  and  $N_2$  are discrete numbers; in the system of Figure 2b, I have started the system with  $N_1 = 1000$  and  $N_2 = 100$ .

Recall that we now think about events competing to happen to individuals, and wish to give exact time stamps to these events. There are three types of events that can happen:

1. a prey individual breeds (rate  $bN_1$ ), and this adds  $B_1$  new prey to  $N_1$
2. a predator may find and kill a prey (rate  $aN_1N_2$ ), with two distinct consequences:  $N_1$  is reduced by one, and — for a subset of cases — the predator may have reached its  $n$ th meal, in which case  $B_2$  new predators are added to  $N_2$
3. a predator may die (rate  $\mu N_2$ ), and this removes one individual from  $N_2$

One could, in principle, track the number of meals of each predator separately; or one can approximate and state that the probability that a particular meal is the predator's  $n$ th one (since the last meal) is  $1/n$ . Choosing the latter approach is computationally efficient, though if death rates of predators are high, one should remember that deaths can interfere before  $n$  meals are reached – in which case  $1/n$  would represent an overestimate. One could, if one wanted to, track the situation more explicitly by creating subclasses of predators that have not eaten since their meal clock was reset, or eaten one, two, ...,  $n-1$  times since the reset; the Gillespie algorithm would simply keep working, but a lot more bookkeeping would be required, as there would be 1 prey class and  $n$  predator classes. For brevity's sake, however, I will proceed with just two types of individuals (predators, and prey) and accept the  $1/n$  as an approximation.

We start the simulation from a desired combination of  $N_1$  and  $N_2$  individuals, and time  $t$  set to zero. The population sizes are now integers. We know that the total rate of anything happening is  $bN_1 + aN_1N_2 + \mu N_2$ . This is interpretable as a parameter of an exponential distribution, and we use the exponential with  $\lambda = bN_1 + aN_1N_2 + \mu N_2$  to sample the time to the next event, which will happen after a wait time that is distributed as  $\tau \sim \text{Exp}(\lambda)$ . The sampling uses a function called `rexp` in R, `expnrnd` in Matlab, `exponential` in Python – details of the syntax therefore vary, but there always is such a function available. A word of caution though: there are two alternative

parametrizations of the exponential distribution, and it is advisable to check if in a particular program  $\lambda$  or  $1/\lambda$  needs to be entered. (Perform a sanity check: If the function is fed a very high rate as an input, it must return very small values for the duration of the wait,  $\tau$ ; if it does not do so, it uses the alternative parametrization where the intended mean of  $\tau$  — also called the scale of the exponential distribution,  $1/\lambda$  — is the input, instead of the rate,  $\lambda$ ).

This might yield, e.g.,  $\tau = 0.00274885$ . This is the time that has to flow before the first event happens; we thus update time from its previous value, e.g.,  $t = 0$ , to  $t + \tau$ , so that time is now 0.00274885. But we still need to decide which type of event it was.

In principle, we already know how to do this (see section 2.4.1.). For this we do not even need to know the value of  $\tau$ , which we only stored because it is typically of interest to see how much time is needed to complete a predator-prey cycle, not only how many events happen during its completion. But technically, how to choose an event, when the chances of each occurring are unlikely to be equal? We cannot pick randomly one of them with  $1/3$  probability each. What needs to be randomized can be visualized by three line segments placed next to each other: one of length  $bN_1$ , the other of length  $aN_1N_2$ , and a third of length  $\mu N_2$ . Next, we choose a uniformly distributed random variable from the range 0 to  $bN_1 + aN_1N_2 + \mu N_2$  (the implementation is easy: choose a uniformly distributed random number between 0 and 1 and multiply the value by  $bN_1 + aN_1N_2 + \mu N_2$ ) and then see in which segment the location landed. The ‘see where it landed’ is easiest to compute with asking if the random number is smaller than the values in the cumulative sum of the values  $[bN_1, aN_1N_2, \mu N_2]$ . If it landed in the first segment, it is smaller than all of the values in the cumulative sum; if it landed in the middle, it is larger than the first cumulative sum value but smaller than the other two, and so on. Now that we know which event happened, we do the update rules as dictated above, take note of the new time ( $t + \tau$  replaces the previous  $t$ ), and we can then proceed to the next event. Integer values of time play no special role in this, though one can of course interrogate the output at integer values of time too: there will be a previous update event that is the last one before each integer, and at the integer-valued time, the system is still in that state.

### 2.4.3. *What does Gillespie show for the limit cycles?*

An example of the stochastic simulation, run until extinction, shows how the stochastic predator-prey system shifts continually between different amplitudes of the Lotka-Volterra cycles (Figure 2b-d). The deterministic expectation is that the system always returns to the same point where we started it, but this is not true when predators and their prey are discrete entities living in a continuous-time world. Instead, the system completes – in this particular run — nine cycles

1 before the prey goes extinct (Figure 2b-d uses different colours for the first three, second set of  
2 three, and the final three). Thereafter the predator only experiences deaths, no births, and its  
3 population, too, dies off, with an exponentially declining number of individuals remaining. This  
4 final decline confirms the expectation that when there is only a constant death rate,  $\mu$ , multiplied  
5 by the number of individuals that are still available to die, the shape of the decline is exponential  
6 (just like it was in Fig. 1d, but now for predators).

7  
8 This allows me to return to my wish that the model would gain realism if predators died faster in  
9 the case where they have little or no prey to feed on. With  $\mu = 1$ , the decline of predators takes  
10 very long. For biological realism it would be best to make  $\mu$  high, to redefine it as the death rate  
11 as a starving predator, and interpret the positive term  $aN_1N_2$  not as 'births' but as a composite of  
12 births and the maintenance of already existing predators. This complicates the pedagogic value  
13 of the Lotka-Volterra framework considerably, however. Its value in explaining real predator-  
14 prey cycles requires quite a lot of modification for many other reasons too: data collection of the  
15 famous Hudson Bay predator-prey cycles involved documented kills by humans, not population  
16 estimates of the individuals that remained after human hunting effort took its toll, and this  
17 introduces many conceptual issues with the model fit (Deng 2018). How much into such detail a  
18 lecturer of population ecology should go to is debatable, and perhaps depends on the level  
19 (bachelor, master) of students too.

20  
21 None of this is the Gillespie algorithm's fault, however. It is important to note that the Gillespie  
22 approach to stochasticity is not an 'error' in the sense of Euler's method's tendency to compute  
23 the trajectories too simplistically. Gillespie really does capture the precisely correct stochastic  
24 version of the relevant set of differential equations (often called 'the master equation'). While  
25 Euler was prone to proceeding straight past a curve (Figure 2a), Gillespie's algorithm shows that  
26 the cycles under stochasticity can move both outwards and inwards from a predefined limit  
27 cycle. They can thus transition from higher to lower amplitudes — as when moving from the first  
28 three cycles (marked with green in Fig. 2b-d) to the next three (blue in Fig. 2b-d) — as well as  
29 the opposite, which is what happens in the last three cycles (red in Fig. 2b-d). Each run of the  
30 simulation is likely to come with a different time to extinction, depending on how long it takes  
31 for the population to end up on their fatal last round.

32  
33 The shape of the cycles is quite similar to the deterministic dynamics, but the ability of a  
34 population to bounce back from very small numbers is clearly overestimated by a focus on  
35 deterministic dynamics. That deterministic Lotka-Volterra cycles require populations to grow  
36 from unrealistically low densities has been called the 'atto-fox' problem (Mollison, 1991; atto is



the name of a minuscule coefficient  $10^{-18}$  that is more frequently found in physics journals than in ecology). Thus, should one already have wondered if it really is realistic to assume that a prey population, that for much of the time stays below 5000 individuals, really can sustain approximately 100 predators (see the y axis scales in Fig. 2c,d), the answer is 'no'. Extinction is the expected outcome, even if the precise time to extinction depends on how often stochasticity moves the system 'inwards' versus 'outwards'. Once a sufficiently high-amplitude cycle is reached, large increases are followed by even more spectacular crashes, and this is unsustainable.

#### 2.4.4. *Lessons learnt (so far)*

As a whole, the lesson is that Lotka-Volterra dynamics can produce cycles (and this is nicely replicated in stochastic simulations too), and it has great historic value, but the interpretation of the coefficients is not entirely easy and, in any case, something else is needed for population persistence (Abrams, 2000; Blasius et al., 2020). The intricacies of species coexistence are beyond the scope of this paper, but it is interesting to note that the Gillespie algorithm is occasionally employed to study its various facets. For example, Okuyama (2015) applied it to two consumers competing for a single resource, Abdullahi et al. (2019) to seed dispersal of two competing plant species, and Acevedo & Stewart (2023) to competing prion strains. The algorithm has even been used to predict outcomes of heterospecific ant battles (Martelloni et al., 2014). The other important lesson from the above exercise, differences between deterministic expectations and stochastic realizations of population growth, has also been considered more extensively elsewhere (Marrec et al., 2023).

There are many intriguing avenues that one could extend the simulations to, including considering evolutionary questions, either in predator-prey (Luhning & DeLong, 2020; Serrao et al., 2021) or other contexts (e.g. resource use evolution, DeLong & Gibert, 2016, or Red Queen dynamics, Schenk et al., 2020). Gillespie approaches can operate over very different timescales, ranging from gene regulatory networks (Ribeiro et al., 2006) to macroevolution (Chevin, 2016). To best learn to apply the method to one's own interest, I would suggest a play with scenarios that extend already published approaches. For example, in the present context, what if some lemmings were better camouflaged than others? The addition that is needed is the splitting of  $N_1$  into at least two categories of individuals (at minimum, 'better' and 'poorer' camouflaged ones, each with a different value for  $a$ ) and reformulating the equations to keep separately track of their births and deaths. Also, lemmings are not asexual, and one might want the genotypes to have some heritability; this would require adding two sexes and being explicit about the

genotypes of offspring if the sire has good camouflage but the female does not, or vice versa.  
Deterministic as well as stochastic versions of the model could then be produced.

Yet, the peculiarities of the Lotka-Volterra dynamics, which I listed above, would still apply. It is worth re-emphasizing that it is not the Gillespie algorithm's fault that the dynamics of the classic predator-prey model is not good at avoiding unrealistically low populations, with the consequence that the Gillespie algorithm with its in-built demographic stochasticity has no alternative but to report the consequences. I will therefore turn to a very different example, that does not share the particular challenges of Lotka-Volterra dynamics, but comes with a set of its own ones.

### 3. Towards more flexibility: predicting matelessness and selection differentials

In my 2<sup>nd</sup> model, I would now like to turn to answering the question I posed in the Introduction: what if some events break our assumption of a memoryless process? Or, what if there are aspects of the environment, or transitions between individual states, that happen at a much more regular pace than the constant rate assumption with its exponentially distributed timings would predict?

My 2<sup>nd</sup> example of the Gillespie approach shows how it can be tweaked to yield insight into many situations that are tedious to solve with differential equations only. The biological inspiration of this 2<sup>nd</sup> example is the fact that mate-searching is not guaranteed to succeed, even for females, in many insect systems (Rhainds, 2019), as there is relatively high daily mortality, and matings can take time. On the other hand, sometimes males can be phenomenally good at locating females, with moth antennae having the ability to detect concentrations of molecules more sensitively than previously reported for any organism (Angioy et al., 2003). But how do these factors play together. Does it still mean that even if males can mate repeatedly, there is a chance that not all females are found by males before their deaths? How precisely does one create predictions for how many females mate?

The modelled situation now is starkly different from the Lotka-Volterra case: we are interested in modelling, over one season (as opposed to an open-ended number of predator-prey cycles), the way individuals drop out of the pool of available ones, either through mating or through dying. To simplify matters, I am going to ignore the fact that many insect species are protandrous, with males emerging before females. I will simply assume that all adults emerge simultaneously. Thereafter, there are two types of events 'competing to happen' to a female: she may be found by a male, or she may die (e.g. by being found by a hungry bird). If the ecological situation that she found herself in was constant over time, we would have an easy time working with the rates:

recall that two events, with rates  $\lambda_A$  and  $\lambda_B$ , yield  $\lambda_B/(\lambda_A+\lambda_B)$  as the probability that B occurs before A; if B is death, then this is also the probability of ending her life without ever having been found by a male, and if males search for females just like predators searched for prey (parameter  $a$ ), an expression would take the form  $\mu_F/(\mu_F+aM)$  where  $\mu$  is a female's death rate and  $M$  is the number of searching males.

But this way to work with rates is only valid under the special case where matings take infinitesimally little time away from a male, and a male immediately afterwards also has sufficient amounts of sperm (and other components of the ejaculate) to find and fertilize new females. In that setting, males 'bounce back' from being unavailable immediately, and it is then irrelevant that they may also find other females, as this takes negligible time away from them finding the focal female. Under this special case, it is fine to not have  $N_F$ , the number of females, appear in the  $\mu_F/(\mu_F+aM)$  term.

But in the more realistic scenario where other females do represent competitors for sperm, since each mating takes time and/or depletes sperm, what should our equations look like? Now there is no shortcut: the number of males available does vary over time, and this should be tracked explicitly. For simplicity, I will here state what to do if matings take time, but there is no upper limit to how many times a male can mate in principle; if such a limit is desired, it can be done, by tracking the dynamics of males who have mated 0, 1, ...,  $L$  times, where  $L$  is the desired limit (Brunton-Martin et al., 2021 model a system with sexually deceptive orchids and their pollinators that are assumed to have such a limit; their aim was to investigate deterministic dynamic differences between haploidiploid and diploid pollinators, without a Gillespie version).

The time cost of mating is, in principle, easy to implement: we simply split the males into two categories, those who are searching (the number  $M$  can now be exclusively used to refer to such males) and those who are alive but unavailable at the moment, as a result of having recently found a female to mate with (denoted  $m$ ). Females who are still unmated are denoted  $F$ , and if we assume that they only mate once, we do not need any other variable to track mated females: if the only feature of female life we are interested in is whether they ever mated before death interfered, we only need to know the dynamics of  $F(t)$  over time. At each time, after all, the hazard  $\mu$  applies, and a proportion  $\mu F(t)$  will die without having mated. Each (infinitesimally small) piece of time,  $dt$ , thus contributes  $\mu F(t)dt$  virgin deaths to the overall pool of dead virgins, which we sum up over the entire season. Summing up infinitesimally small 'slices' is the same as integrating. Therefore, as a whole, integrating  $\int_{t=0}^{\infty} \mu F(t) dt = \mu \int_{t=0}^{\infty} F(t) dt$  then gives us the answer of the total number of unmated females. But to reach this grand goal, we first need to

understand how male availability impacts with the way  $F(t)$ , the number of females who are (a) still alive and (b) still unmated, declines over time.

### 3.1. Deterministic expectations: this time we can trust them better

Again, we could do this in two ways: derive the deterministic expectation (working directly with the differential equations), or build a Gillespie algorithm to yield stochastic outcomes. The deterministic version is easy in principle, but it comes with one worry. Let us have a look at what needs to be done:

$$\frac{dF}{dt} = -\mu_F F - vMF \quad (4a)$$

$$\frac{dM}{dt} = -\mu_M M - vMF + \frac{m}{T} \quad (4b)$$

$$\frac{dm}{dt} = -\mu_M m + vMF - \frac{m}{T} \quad (4c)$$

One can see that the number of unmated females can only decline from any initial number (e.g.  $F(0) = 1000$ ): this part of the population depletes because of female deaths (which I now denote  $\mu_F$  to emphasize that this death rate may be different from male death rate  $\mu_M$ ), and because females leave the pool of unmated females if they mate. Mate encounters happen just like predator-prey encounters did, and now I use the coefficient  $v$  with which these encounters happen (Hutchinson & Waser, 2007). The letter  $v$  can be thought to stand for velocity, but it should be thought to reflect any type of searching efficiency. For example, the high olfactory sensitivity in male moths increases  $v$ , as it undoubtedly speeds up the search.

The number of males that are actively searching for females declines if there are male deaths, or if males enter a ‘time out’ state due to having mated (also called a refractory period, which I here define to include the duration of the actual copulation too). Males in ‘time out’, then, return to the mating pool after their refractory period is over. The number of time-out males being  $m$ , we can compute their rate of returning to the pool of searching males: if each of them spends the time  $T$  (on average) in the time-out state before resuming their search, then  $m \cdot \frac{1}{T}$  is the rate at which males of category  $M$  appear after having been away from this category for a while.

This, however, is a case where we might begin to worry about the way the time to an event is modelled. Recall that all constant rates (here, individuals males resume searching at rate  $1/T$  if they are currently not searching) are only valid if the time already spent waiting for an event has no predictive value for trying to predict when the next event happens. Empirically, this holds for

1 some aspects of life but not others. It has been shown that if a cow has been lying down for a  
2 long time, we can use this information to predict it will stand up soon, but the reverse is not true:  
3 knowing the time that a cow has already been standing up does not help to predict when it will  
4 lie down (Tolkamp et al., 2010), a finding that earned the researchers an IgNobel prize. In our  
5 situation, the question is: If more time has passed since a male begun his copulation with a  
6 particular female, does this indicate an increased probability that he is ready to resume his mate  
7 search soon?

8  
9 A biologically plausible assumption is yes: copulation times are likely to follow some other  
10 distribution than exponential. Exponential distributions imply that there are many short  
11 durations and a long tail of long ones, but it is easy to find data contradicting this assumption:  
12 the range of reported values does not start at 0 but at a larger number of minutes (e.g.  
13 damselflies, Andrés & Rivera, 2000, and bumble bees, Brown & Baer, 2005). Unfortunately many  
14 studies do not report shapes of distributions at all, only summary statistics. But it is safe to  
15 assumed that matings and the subsequent recovery times (refractory period) cause the time-out  
16 of insect males to have much more regular durations than an exponential distribution would  
17 suggest. As an aside, should a reader worry about male bees only being able to mate once:  
18 bumble bee males are not in this category, as they can mate multiply (Belsky et al., 2020).

19  
20 But is the shape of the distribution a detail that does not matter, or does it change the mating  
21 dynamics substantially? The only way to know is to study this question by modelling it both  
22 ways, and then perform the comparison. We can first derive the expectations that use the  
23 exponential distribution assumption, yielding solutions for  $F(t)$ . Technically, one uses the  
24 numerical solver of one's favourite programming language for the equations (4a-c), and then  
25 computes the matelessness value for each combination of  $v$ ,  $T$ ,  $\mu_F$  and  $\mu_M$  (Figure 3 shows an  
26 example). Thereafter, one needs to know how to change the assumption sets.

### 27 28 3.2. Gillespie with a touch of memory

29 The task is one that is often encountered: one is dissatisfied with some assumptions that were  
30 necessary to make the ODE approach work, and would like to see if the key results are robust to  
31 changing this or that assumption. But does the Gillespie algorithm help here? One might think  
32 'no', since the Gillespie algorithm in its classical form relies on the memoryless assumption too,  
33 which in the present case translates to 'copulation durations should be exponentially  
34 distributed'. Yet, it is very easy to modify the algorithm to include events that are scheduled  
35 differently than the memoryless assumption dictates. And while one is at it, why not also think

about other aspects that can be relaxed (optionally, one at a time, for otherwise we would not know what change is responsible for what outcome). For example, what if males differ from each other in the values of  $v$  (searching efficiency) and/or  $T$  (perhaps some males can resume searching faster than others)? Both traits should be under sexual selection, and one might be interested in deriving just how much the values matter as determinants of mating success.

In Figure 4, the algorithm has been implemented to achieve each of these goals. The example starts with  $F(0) = 1000$  females and  $M(0) = 1000$  males, and  $m(0) = 0$  because no male is initially dealing with the consequences of a previous mating. At any time point from  $t = 0$  onwards, there are rates that relate to females dying (rate  $\mu_F F$ ), searching males dying (rate  $\mu_M M$ ), or males in time-out dying (rate  $\mu_m m$ ). These death-related events are dealt with like we already learned in the Lotka-Volterra case. The mate encounter events and the events of males resuming search have to be handled slightly differently, however.

### 3.2.1. *Mate encounters and what happens thereafter*

Let us think about mate encounters first. They are now male-specific (though all chosen to be identical in Figure 4a and c). The relevant rate is  $v_i F$  for male  $i$ . For the specific examples of Figure 4, I chose the lognormal distribution to describe within-male variation:  $v_i = e^{\ln v + \varepsilon_i}$ , where  $v$  now indicates the median search efficiency of the population and  $\varepsilon_i \sim N(0, \sigma_v^2)$ . In the lognormal, the logged values have values for a mean  $\ln v$  and variance  $\sigma^2$  that describe an underlying normal distribution; the  $v$  of the lognormal distribution then describes the median, not the mean. In case the reader is wondering why I did not simply choose a normal distribution: this would make it difficult to scale the variance across scenarios with vastly different  $v$  values. The choice of the lognormal allows us to interpret the same standard deviation value (in Figure 4,  $\sigma_v = 0.25$  whenever it is nonzero) sensibly across a wide range of values of  $v$ : on an absolute scale, a small improvement of  $v$  matters more when all males have small  $v$  values. Implementing the variation on a log scale takes this into account, and also avoids the worry that adding normally distributed values to a small number would make us have to deal with cases where some search efficiencies become negative. On a log scale they can be negative without problems, the lognormal then ‘lifts’ them to small positive search efficiencies.

Now that males have their  $v_i$  values randomized and assigned to them, how to know when an event ‘a particular male mated’ happened? In principle, this can be modelled in the very same way as before, there simply are far more types of competing events (three kinds of deaths, and a total of  $M$  types of mate encounters, one for each searching male) that all contribute to the total sum of rates in the Gillespie algorithm. An equivalent alternative way, perhaps conceptually

clearer when writing programming code — though possibly computationally a little slower — is to deviate from the approach of forming a sum of total rates. It is mathematically equivalent to first form the death time of the first individual to die as usual, and compare this candidate for an event to the mating time of the first male to mate, which is another candidate event. The latter can be done by drawing from an exponential distribution with a vector of values  $\mathbf{v}F$ , which gives one event time for each male ‘in one go’ (in programming language, this is a vectorized operation). Most programming languages allow drawing multiple samples from an exponential distribution simultaneously (without ‘for loops’), each with their own parameter (here,  $v_i$ ). Then we choose the smallest of these values, which is the candidate event that is scheduled to happen soonest. If the soonest mating is scheduled to happen earlier than the soonest of the deaths, the program proceeds to a mating, with a particular male then entering the time-out state.

Unless, of course, the soonest event to occur is some male’s return to the searching state, our final category of events. Like with  $v$ , one needs to model inter-male variation so that no negative values are permitted. I again choose the lognormal so that the male-specific time-out duration  $T_i = e^{\ln \bar{T} + \theta_i}$ , with  $\theta_i \sim N(0, \sigma_T^2)$ .

Here, finding out the time to the next event is different from before. With  $T_i$  the goal is to avoid using the exponential distribution, thus one cannot rely on its convenient memorylessness. To provide an example that deviates from the memorylessness assumption as much as possible, I will consider a case where  $T_i$  indicates a precise male-specific duration of the time out, with zero within-male variance when the same male mates repeatedly, but with variance between the males permitted ( $\sigma_T^2$  can be either zero or positive).

It is not a computationally difficult problem to make some events deviate from memorylessness. Recall that the main goal is to decide which event type occurs soonest. The task is to assign each male a time stamp which denotes the pre-scheduled value of  $t$  that makes it resume searching (transitioning from  $m$  to  $M$ ). If the male mated at time  $t$ , this time stamp takes the value  $t+T_i$ , and this value is kept in memory (without changing it when  $t$  is updated due to other events occurring). Above, we had already found which event, death or mating, happens first. This soonest-occurring of the other competing events has a time stamp of  $t+\tau$  (with the  $t$  here now referring to the last event that happened in the population as a whole). If that time stamp is later than the earliest male-specific time stamp for resuming searching, then the outcome is that the earliest male scheduled to resume searching indeed does so. Other male’s time stamps for resuming are kept unchanged no matter which type of event occurs. This keeps the memory intact for how long they still have to spend in the time-out state.

This approach can be used for any event type that has a distribution that deviates from the exponential. For example, the programming code can include a time stamp for the end of the entire season due to seasonality. Each time an event is chosen, a small check is needed: it has to have a time stamp smaller than the end-of-season time stamp. If no competing event offers such a value, then the season's end is indeed the soonest event to occur. The program can then track whatever is assumed to happen at the end of season: e.g. if the 'season' referred to an overwintering season, a birth pulse might now happen, or if the 'season' was a summer during which *Daphnia* can breed, the season end means that only dormant eggs can now survive, and all other individuals are removed.

In our case of mating dynamics, a small detail still needs to be dealt with. For males that are already searching, it is inappropriate to have a time stamp that makes them begin searching (one should not attempt to transition to a state where one is already), but for computational efficiency reasons, it may be beneficial to compare all time stamps in one go, instead of any cumbersome attempt to exclude some males from the approach. An efficient way to code the comparison is to assign all currently searching males an infinitely long resume-searching time stamp. This requires that the programming language of one's choice allows to set  $\infty$  as a value, with the desired property that  $\infty$  will never be chosen as the smallest value in a set. Thus, the results of Figure 4 were obtained by setting the initial ( $t = 0$ ) resume-search time stamps to  $\infty$  for all males, and setting one of them to  $t + T_i$  whenever a specific male was mating.

### 3.2.2. Scenario comparison: large selection differentials do not align with high matelessness

The Gillespie algorithm allows us to consider many scenarios for matelessness and the consequences for other variables in the system. In Figure (4), each panel's heatmap shows the mean outcome of 10 simulation runs for a fine grid of different values of search efficiencies  $v$  (logarithmically spaced from  $v = 0.1$  to 100, each  $v$  interpreted as the median of a distribution when there is variation) and time out ( $T$ ) (logarithmically spaced from  $T = 0.001$  to 1, again interpreted as the median when there is variation). Note that the upper end of choices for  $T$  means that typical time outs are as long as the average lifespan, since Figure 4 uses  $\mu_F = \mu_M = 1$ . The leftmost column (a) assumes no between-male variance, i.e. that the  $\{v, T\}$  combination at a given location in the heatmap is used by all males, while in (b) the searching efficiency  $v_i$  differs between males, in (c) the time out  $T_i$  varies between males, and in (d) the assumptions of (b) and (c) are combined such that males vary in both traits. The non-varying case (a) also permits us to do the comparison to the expected matelessness as predicted by numerically solving the set of



ODEs; this is done for three different contour lines of matelessness (1 out of 1000 females, 5 out of 1000 females, and 10 out of 1000 females). If these match the heatmap's colour, then the results are robust despite the ODE (biologically awkwardly) relying on memorylessness for the duration of the time out.

Unsurprisingly, the matelessness values are clearly highest when searching is inefficient (low  $v$ ) and when male time-outs are large (large  $T$ ), but the per-female risk of remaining unmated only reaches about 2%, even in the worst-performing case (darkest colours in Figure 4a, top row, refers to around 20/1000 unmated females). How well does the differential equation approach predict this result – did its biologically unjustified distribution for the time out duration matter? The prediction (contour lines in Figure 4a) is not too far off. The ODE, too, is able to outline areas of high and low matelessness and make statements about where the problem is negligible, i.e. to the right of the 1 out of 1000 females contour line, which also corresponds to the area where the Gillespie algorithm predicts the appropriately pale colour for the heatmap, indicating 1 or fewer unmated females. In other words, the assumption in the ODE that mating durations are exponentially distributed, which was made for mathematical convenience only, did not cause wildly wrong estimates for matelessness.

It is also possible to collect data, during the one-generation simulation, on the identity of sires and how their values of search efficiency and time-out differ from the mean of all males. This is known as the selection differential. In the current context it should be computed using the logarithms of the trait values,  $\ln v$  and  $\ln T$ . The reason for this choice is identical to the reason why the lognormal was chosen to be the distribution. Very small searching rates or time-outs do not permit the subset of highly performing males to have trait values that differ very much from the rest of the population, but they may still be relatively speaking much ahead of the rest. The relevant scales are different when average trait values are small or large; to use the analogue from body size evolution, one should not erroneously conclude that selection can only be strong in large-bodied animals. To make the selection differentials properly comparable across scenarios, the appropriate comparison is between the mean of logged trait values for sires with the mean of logged trait values of all males. This also restores the interpretation of  $\ln v$  and  $\ln T$  being population means, not medians; the definition of the selection differential does not use medians.

In this comparison, selection differentials for search efficiency are high whenever either  $v$  or  $T$  (or both) are small or at least not very high (Figure 4, 2<sup>nd</sup> row). This is not the same parameter region as that predicting strong matelessness (Figure 4, 1<sup>st</sup> row). Instead, selection on  $v$  is

stronger when the situation permits the best-performing male to mate with more females (data on this can also be collected during the Gillespie algorithm runs, Figure 4, 4<sup>th</sup> row). Many  $\{v, T\}$  combinations simply do not create very strong polygyny, and this limits how strongly sexual selection can operate in the modelled situation. This is logical: if search rates are high and each time-out is long, then each male finds his first female rather quickly, and spends a long time processing this one mating. During this time, the remaining males, being quick searchers too, have found their own one female each (pale areas in the 4<sup>th</sup> row of Figure 4).

The pattern is different for the conditions that predict selection for shorter time-outs to be high (Figure 3, 3<sup>rd</sup> row; the colours indicate the absolute value of the selection differential, the underlying differentials are negative since males with a shorter, not longer, time-out are favoured as sires). Here, when males as a whole have a very high search efficiency  $v$  and this combines with a long time-out  $T$ , selection on  $T$  is weak: the situation rarely offers more than one female per male, and a male with a somewhat shorter  $T$  cannot change this as females have run out by the time this male re-emerges from his first time out. The weakest selection on  $T$  occurs under the same conditions that predict weakest selection on  $v$ . But for time-out, the strongest selection on  $T$  (Figure 4, 3<sup>rd</sup> row) does not correspond to strongest polygyny (Figure 4, 4<sup>th</sup> row), but to intermediate polygyny values. Very short time outs, combined with very low search efficiencies, mean that males cannot gain much by attempting to make the time-out shorter still (even if one correctly tracks their success against log-transformed trait values, which should permit seeing the effect of subtle changes at short timescales. Here the lacking search efficiency is now the limiting factor of male mating success, and modifications of time out simply do not help that much.

### 3.3. Extensions, efficiency aspects and stopping rules

One could extend the matelessness example in numerous directions. For example, one could investigate what happens if the mate searching effort of high- $v$  males comes at a cost of higher mortality. If one desired to move the analysis in a more explicit evolutionary direction still, one could proceed beyond the analysis of a single generation. The events above tracked the numbers of individuals over one season. In an extension, each mating could be associated with the bookkeeping of eggs and their genotypes, which then hatch in a subsequent season, and the evolutionary outcomes can be tracked over multiple generations. Ekrem & Kokko (2023) provide such an extension. They also considered the question of sex-specific emergence times (protandry and protogyny) and their evolution, limited the number of times a male could mate, and showed what to do when a modelling approach cannot, for the most part, rely on memorylessness: they

1 included a seasonal peak in the availability of resources. This last assumption implies that  
2 matings are more valuable when they occur at a phenologically suitable time. Their solution was  
3 to compute all time stamps for every possible mating first, and then remove those that could not  
4 logically happen, e.g. due to death having occurred first for one or both participants, or if the  
5 male had, in previous events, depleted all his sperm supplies. The take-home message is that  
6 Gillespie-like processes can coexist with very many types of memory in the system too. Although  
7 this may cause some reduction of computational efficiency, the computational cost need not be  
8 huge.

10 Specifically, the efficiency of the algorithm can nevertheless be maintained quite high, if one  
11 realizes that the search for the soonest event to occur can usually be vectorized (simultaneous  
12 computing of many scenarios). Even though the algorithm itself only considers one event at a  
13 time, these can be computed in parallel for multiple populations, either for the same set of  
14 parameter values, in which case we are computing replicates, or for a large number of different  
15 ones. The heatmaps in Figure 4 were created that way, and the 10 replicates for 3720 different  
16 parameter values across the 4 scenarios took no longer than one lunch break to compute on a  
17 small, moderately powerful tablet, as this figure needed to be computed while travelling. In case  
18 of larger problems where computing time becomes a critical resource, one can read up on more  
19 advanced topics such as tau-leaping (Nobile et al. 2014), where several events are computed in  
20 one go, trusting (or hoping) that each of them has only a minor effect on modifying the rates – a  
21 trade-off between accuracy and speed, in other words.

23 The only unusual aspect compared to discrete-time steps is that each replicate, in a Gillespie  
24 simulation, yields a uniquely different sequence of time points during which events happen.  
25 Therefore the total time that each population has been simulated for takes a different value by  
26 the time a pre-specified number of events (e.g. 10 million) has been computed. One can thus  
27 decide on an appropriate stopping rule: perhaps it makes sense to run a pre-specified number of  
28 events for every population, or stop the simulation when all populations have run for at least a  
29 certain pre-specified number of time units (and some will then have run for longer – but one can  
30 still ask about their state at the pre-specified minimum), or then a question-specific stopping  
31 rule may be implemented, such as no observation of a significant covariation between the time at  
32 the last 1000 events and an evolving trait value, which can be used to give an indication of  
33 convergence with no further evolutionary change. In Figure 4, the logical criterion was to run the  
34 simulation until the fate of all of the initial  $F(0) = 1000$  females was known (either mated or  
35 dead).

#### 4. Appetite for more than just the numerical solution? Closed form solutions offer insight when they can be derived

A nice feature of the exponential distribution is its simplicity: it is easy to predict the expected number of events that occur over a specific time (the relevant mathematical literature is that of the Poisson process), and the probability that an event has occurred by time  $t$ , such as the death of an animal, is very easy to derive: as long as nothing, such as the number of predators interested in attacking a focal prey, changes, the rates predict that the focal prey is alive with a probability  $e^{-\lambda t}$ , where  $\lambda$  is a rate that appropriately describes the situation. This is an example of a closed form solution for something (in this case, the survivorship for an animal): it is a neat result to have, because it means that to get an answer to the question “by when do we have only a fraction  $x$  (e.g.  $x = 0.001$ ) of the initial population left?” one does not have to proceed painfully slowly from one time ‘step’ to another, probing if there still are more than 0.001 left; one can simply solve  $e^{-\lambda t} = x$ , which is done by taking logs on both sides, thus  $-\lambda t = \ln x$  and the answer is  $t = -\ln x / \lambda$ , an answer that is valid for any  $x$  and  $\lambda$  (and the  $t$  takes positive values despite the presence of a minus sign; this happens because when  $x < 1$ ,  $\ln x$  is negative).

But solutions are not always that simple: after all, in most interesting models, more event types than just 1 are competing to happen, e.g. the number of predators keeps changing as does that of their prey. Consequently, in Figure 2d, the simple expectation of an exponential decline in individual numbers was visible only at the very end of the simulation — in this case for the predator population, subject to a death rate  $\mu$  — as the predators had run out of prey and this left only one event type possible (death). As a whole, the Lotka-Volterra equations already have so many simultaneous processes happening in them that the closed form solutions are unwieldy and tend to be discussed in journals with a mathematical rather than an ecological focus (e.g. Evans & Findley, 1999).

This does not, however, mean that it is generally not worth looking for closed form solutions to ecologically interesting problems. My final example below shows how this can be done (for other examples see e.g. Allen & Dytham, 2009), and it additionally provides an example of interesting links between continuous and discrete time modelling. For two very famous discrete-time models of density dependence, there is a known but perhaps underappreciated link between the dynamics that is visible from one birth pulse to another, and the underlying continuous-time processes that can, when viewed at intervals of  $t = 1$ , produce exactly the type of density dependence that the discrete-time version describes.

To highlight correspondences between discrete and continuous time formulation requires some mental gymnastics in notation, as two different interpretations of time need to be hybridized. I will aim for maximum clarity by using capital letters  $N$  and  $T$ , with subscript notation  $N_T$ , to refer to the discrete-time tradition, and lowercase  $n(t)$ , with time indicated in brackets, to refer to what happens if a non-integer time  $0 < t < 1$  has passed since the last birth pulse. (In calculus, sometimes  $\tau$  is used for the latter, which would be useful here too, had I not already used it to refer to the values of random variables denoting time to an event in the Gillespie algorithm. Overlaps between different models are sometimes hard to avoid: here  $n(t)$  is hopefully sufficiently different from the  $n$  that I used in the Lotka-Volterra model to denote the number of prey captures needed to form a new predator, so that confusion can be avoided).

In both cases that I will discuss, the idea is that there is a birth pulse where the number of breeding individuals,  $N_T$ , is simply multiplied with a variable  $> 1$ , which I will call  $R_0$ , to yield the post-breeding population,  $R_0 N_{T+1}$ . Thereafter, there is a non-breeding season, during which any of the  $R_0 N_{T+1}$  individuals may die. There are various possibilities how deaths could be density dependent, but let us consider two possibilities.

#### 4.1. Events during the non-breeding season can lead to Beverton-Holt or Ricker dynamics from one birth pulse to the next

First, it could be the case that the instantaneous death rate,  $\mu(t)$ , depends on the number of conspecific competitors that are still alive at time  $t$ ; logically, when all interactions are competitive, the death rate should be an increasing function of the number of competitors. It also then follows that populations diminish rapidly at the beginning of the non-breeding season, and less rapidly later on, not only because there are fewer individuals still available to die (total deaths observable during a small amount of time,  $\mu(t)n(t)dt$ , obviously diminish when  $n(t)$  diminishes), but also because  $\mu(t)$  itself is now assumed to shrink as  $n(t)$  shrinks. An interesting choice is  $\mu(t) = \mu_0 n(t)$ , such that the death rate takes the value  $\mu_0$  if there is only one individual left, and increases linearly with the number of competitors present. Now, we can ask, how many individuals are left at  $t = 1$  if we started at  $t=0$  with  $n(0) = R_0 N_T$  individuals? This requires solving the ODE

$$\frac{dn}{dt} = -\mu(t)n(t)dt = -\mu_0(n(t))^2 dt \quad (5)$$

over the time interval from 0 to 1. This is an ODE with a closed form solution; if one is not fluent enough to solve differential equations using pen and paper, then one can ask mathematical

software capable of analytical solutions to provide the answer, e.g. in Mathematica the relevant command is `DSolve[n'[t] == -mu0 n[t]^2, n[t], t]`. The solution is  $n(t) = 1/(\mu_0 t - c)$ , where  $c$  is a constant. Its role is to choose, among the infinitely many solutions (one for each value of  $c$ ), the relevant one: we know the state of the population at one time point, namely  $t = 0$  (where it is  $n(t) = R_0 N_T$ ), and  $c$  can be adjusted so that the trajectory is 'nailed' to go through the desired  $\{t, n(t)\}$  combination of values. (Solving for the relevant value of  $c$  can also be added to the request that programs such as Mathematica perform for us when we wish to find the ODE solutions, but I consider it useful to write it out explicitly here so that one can see that there is nothing particularly difficult about it.) Since  $1/(\mu_0 t - c)$  simplifies to  $-1/c$  when  $t = 0$ , we have  $-1/c = R_0 N_T$ , thus  $c = -1/(R_0 N_T)$  and we can use this  $c$  to predict the value of  $n(t)$  at the end of the non-breeding season where  $t = 1$ . Evaluating  $n(t) = 1/(\mu_0 t - c)$  at  $t = 1$  and  $c = -1/(R_0 N_T)$  yields  $n(1) = \frac{1}{\mu_0 + \frac{1}{R_0 N_T}}$ , which can also be written as  $n(1) = N_T \frac{R_0}{1 + \mu_0 R_0 N_T}$ . Realizing that  $n(1)$ , the population size at the end of the non-breeding season, is simultaneously the next time step's pre-breeding population,  $n(1) = N_{T+1}$ , we have arrived at

$$N_{T+1} = N_T \frac{R_0}{1 + \mu_0 R_0 N_T}. \quad (6)$$

This happens to be the functional form of a standard discrete-time model of population growth, the Beverton-Holt model (Beverton & Holt, 1957), though it was not originally written with the aid of a mortality coefficient  $\mu_0$ . This result (originally described in Kokko & Lindström, 1998) generalizes to various situations: one can perform an explicit tracking of the abundance of the resource that the population depends on (Geritz & Kisdi, 2004); others have extended the dynamics to include e.g. cannibalism and two-sex dynamics (Eskola & Parvinen, 2007). Interestingly, the dynamics can arise through many processes, and not all of them rely on a continually depleting resource at all (Royama, 1992, Brännström & Sumpter, 2005) – from which we can conclude that for any actual situation, one needs to know more than just the population sizes at specific points in time to infer the properties of the process that produced the sequence observations.

To complete my aim of showing how another discrete-time classic model can arise from continuous dynamics within a season, consider an alternative assumption to the above logic: now assume that the number of individuals after the first birth pulse, which is  $R_0 N_T$ , determines the state of the environment that, in turn, determines the death rate such that the rate is proportional to the number of post-breeding individuals; however, now we assume that this death rate applies in an unchanged manner throughout the non-breeding season, i.e., the death

of competitors some time into the non-breeding season does not help to vacate resources to be distributed among the survivors (thus we make a radically different assumption regarding compensatory mortality). Now the relevant ODE is

$$\frac{dn}{dt} = -\mu(t)n(0)n(t)dt \quad (7)$$

It is solved with a similar procedure as above, but the solution is simpler: the closed form is  $n(t) = ce^{-\mu_0 n(0)t}$ , and the constraint  $n(0) = R_0 N_T$  makes us deduce  $c = R_0 N_T$ , thus the population declines throughout the season as  $R_0 N_T e^{-\mu_0 n(0)t}$ , and when  $t = 1$ , this simplifies to  $R_0 N_T e^{-\mu_0 n(0)}$ . Remembering, again, that  $n(0) = R_0 N_T$  and  $n(1) = N_{T+1}$ , we have established the link between two discrete-time points:

$$N_{T+1} = R_0 N_T e^{-\mu_0 N_T}. \quad (8)$$

This is the equation defining the Ricker dynamics (Ricker, 1954), even though once again the notation is usually different. It may require some practice to see the fundamental equivalence between a relatively standard Ricker formulation  $N e^{r(1-N/K)}$  and what we just saw,  $R_0 N_T e^{-\mu_0 N_T}$ . But the former can be written as  $N e^{r(1-N/K)} = N e^{r-rN/K} = N e^r e^{-r/KN}$ , thus the equivalence arises once noting that  $R_0$  simply needs to take the value  $e^r$ , while the same  $r$  will be modified to yield  $\mu_0$  by dividing by  $K$ , i.e.  $\mu_0 = r/K$ .

All these results could, of course, also be put into the context of the Gillespie algorithm. To keep this paper of readable length, I will do not describe the procedure here, and instead simply hope that the reader appreciates that without the knowledge of what kind of within-season dynamics can yield the Ricker or the Beverton-Holt equations, one could not even proceed to study the way that stochastic survival of individuals could yield temporal fluctuations around the equilibria that are measured at every breeding season.

## 5. Conclusions, ending with a wish

*For example, I was not able to fully decipher  $v$ , as this was defined as a mate search rate [...] but without any explanation. I also expected this value to be  $< 1$ , as assumed in [reference] for “mate acquisition rate”, so that for all potential pairings between males and females, there is a probability of pairing. In your model, is the interpretation the same, so that  $v=100$  mean that for each possible pairwise encounter there are 100 matings? If so, what happens if there are fewer than 100 males or*

1 *100 females? What constrained your choice of  $v$  from 0.1 to 100, and the values for the other*  
2 *parameters?*

3 *anonymous reviewer*  
4

5 The above is a reviewer comment to a MS that a PhD student and I sent to a journal a while ago,  
6 and after much clarifying effort (the reviewer thought there must be time steps in the model  
7 when we did not say there were any), it was published (Ekrem & Kokko, 2023). This comment  
8 prompted me to write this review: there are efficient methods out there, but they are more often  
9 used by researchers who feel comfortable publishing in journals with a mathematical or  
10 computational angle, e.g., a focus on statistical physics. For ecologists, their use requires a little  
11 bit of refreshing of the memory of calculus classes at high school or undergraduate teaching,  
12 some additional familiarity with how modern software can help overcome the tedious bits, and  
13 perhaps a willingness to step outside familiar paths where simulations obviously ‘must’ have a  
14 time step structure. I hope this text can help a little for those willing to understand just how  
15 flexible the methods can be.  
16

## 17 Box 1. Counterintuitive but useful consequences of memorylessness

18 There are some counterintuitive properties of the exponential distribution, that all relate to the  
19 fact that the time elapsed since the last event yields no insight into the time that one still needs  
20 to wait until the next event occurs. If, for example, lifespans are exponentially distributed (which  
21 implies that there is some age-independent extrinsic mortality rate, but no senescence), then a  
22 younger and an older individual have exactly the same life expectancies from now on, and the  
23 probability that either of them outlives the other is 0.5. Or, differently put, adding one year that  
24 one has survived successfully does not deplete the expected number of years ahead. As such,  
25 these statements should not sound too paradoxical: to think that there should be a difference, i.e.  
26 an individual to whom nothing has happened for a longer time is now “due to” experience an  
27 event sooner, is the same fallacy as a gambler succumbs to, if thinking that since past dice throws  
28 have for a long time not produced the number six, this number is now “due to arrive” sooner.  
29 Humans, of course, do senesce, and perhaps the fact that our lifespans are far from being  
30 exponentially distributed makes it so very difficult to have correct intuition about variables that  
31 really are exponentially distributed.  
32

33 Another way to express the counterintuitive property of memoryless processes is that knowing  
34 which of two events is observed first – one, A, that occurs with a high rate, and another, B, that  
35 occurs at a low rate — yields no information regarding how quickly the event happened. Again,



intuition might suggest that those events that can be expected to occur soon (high rate) are overrepresented among events that occur in very the near future. Thus if the observation was that A happened, we are tempted to conclude that it probably did happen after only a short wait, or conversely, if we know that the wait was short, the temptation is to conclude that it was more likely that A rather than B happened. But neither interpretation is true: the time to the next event is exponentially distributed with the mean  $1/(\lambda_A + \lambda_B)$  and, no matter what precise value the random variable took (i.e. how long did we have to wait to observe the next event), the probability it was A is  $\lambda_A/(\lambda_A + \lambda_B)$ . But although this constancy can make one's head spin a little, these properties are also extremely nice: they allow computational efficiency in the Gillespie algorithm by making the probabilistic choice of each event a simple one, and they can simplify calculations of expected lifetime reproductive success considerably.

Perhaps the most intriguing property of a memoryless process, in the context of lifespan specifically, is what happens when sampling a population that keeps its size constant (whenever a vacancy is created by death, it is filled by a new recruit; this is called the Moran process (Moran, 1958)). When individual lifespans are exponentially distributed, with mortality rate  $\mu$ , what is the mean age of observed deaths? Clearly the expected mean for such a data set is  $1/\mu$ . Next question: what is the mean age of individuals observed to be alive at any point in time? The expected value is also  $1/\mu$ . The equivalence of these two results sounds deeply paradoxical: shouldn't the average age of individuals be half of their age at death, since every individual necessarily passes through all ages from 0 to its own lifespan, and the midpoint of that range is  $1/(2\mu)$ ? Another way to phrase the paradox is to make the additional assumption that reproduction occurs continually from age 0 to death, and phrase the measurement of ages of living individuals as a measurement of generation time. How can the mean age of parents (i.e., generation time, Ellner, 2018) be the same as the mean age of freshly dead individuals, when each parent necessarily has produced its offspring at a range of ages that are all younger ages than the age of its own death?

The paradox finds its solution once realizing that individuals are weighted differently in the two different samples. Each individual counts only once towards the age-at-death distribution. While alive, however, the lucky subset of individuals who avoid death for longer are equivalently much longer available to be part of a sampling effort than their short-lived conspecifics. This allows the ages of those who offer ages far beyond  $1/\mu$  to be sampled to pull the mean upwards. This finding applies to various distributions of lifespan, but in the exponential case, the pull upwards is strong enough to make the two means identical. Weird, but true.

## References

- Abdullahi, A., Shohaimi, S., Kilicman, A. & Ibrahim, M.H. (2019). Stochastic models in seed dispersals: random walks and birth-death processes. *Journal of Biological Dynamics*, 13, 345–361.
- Abrams, P.A. (2000). The evolution of predator-prey interactions: theory and evidence. *Annual Reviews of Ecology and Systematics*, 31:79-105.
- Acevedo, S. & Stewart, A.J. (2023). Eco-evolutionary trade-offs in the dynamics of prion strain competition. *Proceedings of the Royal Society of London B*, 290, 20230905.
- Allen, G.E. & Dytham, C. (2009). An efficient method for stochastic simulation of biological populations in continuous time. *Biosystems*, 98, 37-42.
- Andres, J.A. & Rivera, A.C. (2000). Copulation duration and fertilization success in a damselfly: an example of cryptic female choice? *Animal Behaviour*, 4, 695-703.
- Angioy, A.M., Desogus, A., Barbarossa, I.T., Anderson, P. & Hansson, B.S. (2003). Extreme sensitivity in an olfactory system. *Chemical Senses*, 28, 279-284.
- Belsky, J.E., Camp, A.A. & Lehmann, D.M. (2020). The importance of males to bumble bee (*Bombus* species) nest development and colony viability. *Insects*, 11, 506.
- Berrios-Caro, E., Galla, T. & Constable, G.W.A. (2021). Switching environments, synchronous sex, and the evolution of mating types. *Theoretical Population Biology*, 138, 28–42.
- Beverton, R.J.H. & Holt, S.J. (1957). On the dynamics of exploited fish populations. *Fisheries Investigations London, Series II*, 19, 1–533.
- Black, A.J. & McKane, A.J. (2012). Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution*, 27, 337-345.
- Blasius, B., Rudolf, L., Weithoff, G., Gaedke, U. & Fussmann, G.F. (2020). Long-term cyclic persistence in an experimental predator–prey system. *Nature*, 577, 226–230.
- Boonekamp, J.J., Briga, M. & Verhulst, S. (2015). The heuristic value of redundancy models of aging. *Experimental Gerontology*, 71, 95-102.
- Brännström, Å. & Sumpter, D.J.T. (2005). The role of competition and clustering in population dynamics. *Proceedings of the Royal Society of London B*, 272, 2065-2072.
- Brown, M.J.F. & Baer, B. (2005). The evolutionary significance of long copulation duration in bumble bees. *Apidologie*, 36, 157–167.
- Brunton-Martin, A.L., Gaskett, A.C. & Kokko, H. (2021). Resilience of haplodiploids to being exploited by sexually deceptive plants. *Oikos*, 130, 2053–2063.
- Chevin, L.-M. (2016). Species selection and random drift in macroevolution. *Evolution*, 70, 513-525.

- 1 DeLong, J.P. & Cressler, C.E. (2022). Stochasticity directs adaptive evolution toward  
2 nonequilibrium evolutionary attractors. *Ecology*, 104, e3873.
- 3 DeLong, J.P. & Gibert, J.P. (2016). Gillespie eco-evolutionary models (GEMs) reveal the role of  
4 heritable trait variation in eco-evolutionary dynamics. *Ecology and Evolution*, 6, 935-945.
- 5 Deng, B. (2018). An inverse problem: trappers drove hares to eat lynx. *Acta Biotheoretica*, 66,  
6 213-242.
- 7 Ekrem, R.K. & Kokko, H. (2023). Sexual conflict over phenological traits: selection for protandry  
8 can lock populations into temporally mismatched reproduction. *Evolution*, 77, 789-800.
- 9 Ellner, S.P. (2018). Generation time in structured populations. *American Naturalist*, 192, 105-110.
- 10 Eskola, H.T.M. & Parvinen, K. (2007). On the mechanistic underpinning of discrete-time  
11 population models with Allee effects. *Theoretical Population Biology*, 72, 41-51.
- 12 Euler, L. (1768). *Institutiones calculi integralis* (in Latin).
- 13 Evans, C.M. & Findley, G.L. (1999). A new transformation for the Lotka-Volterra problem. *Journal*  
14 *of Mathematical Chemistry*, 25, 105-110.
- 15 Gerber, N. & Kokko, H. (2016). Sexual conflict and the evolution of asexuality at low population  
16 densities. *Proceedings of the Royal Society of London B*, 283, 20161280.
- 17 Geritz, S.A.H. & Kisdi, É. (2004). On the mechanistic underpinning of discrete-time population  
18 models with complex dynamics. *Journal of theoretical Biology*, 228, 261-269.
- 19 Gillespie, D.T. (1977). Exact stochastic simulation of coupled chemical reactions. *The Journal of*  
20 *Physical Chemistry*, 81, 2340-2361.
- 21 Golubev, A. (2004). Does Makeham make sense? *Biogerontology*, 5, 159-167.
- 22 Hughes, B.G. & Hekimi, S. (2016). Different mechanisms of longevity in long-lived mouse and  
23 *Caenorhabditis elegans* mutants revealed by statistical analysis of mortality rates. *Genetics*,  
24 204, 905-920.
- 25 Hutchinson, J.M.C. & Waser, P.M. (2007). Use, misuse and extensions of "ideal gas" models of  
26 animal encounter. *Biological Reviews*, 82:335-359.
- 27 Kingsland, S.E. (1995). *Modeling Nature*. 2nd ed. University of Chicago Press, Chicago.
- 28 Kokko, H. & Lindström, J. (1998). Seasonal density dependence, timing of mortality, and  
29 sustainable harvesting. *Ecological Modelling*, 110, 293-304.
- 30 Leibniz, Gottfried (1684). Nova Methodus pro Maximis et Minimis (in Latin). *Acta Eruditorum*, 3,  
31 467-473.
- 32 Lotka, A.J. (1925). *Elements of Physical Biology*. Williams & Wilkins, Baltimore, MD.
- 33 Luhning, T.M. & DeLong, J.P. (2020). Trophic cascades alter eco-evolutionary dynamics and body  
34 size evolution. *Proceedings of the Royal Society of London B*, 287, 20200526.
- 35 Martelloni, G., Alisa, S., Bagnoli, F. & Santini, G. (2014). Modeling ant battles by means of a  
36 diffusion-limited Gillespie algorithm. *Theoretical Biology Forum*, 107, 57-76.

- 1 Masuda, N. & Vestergaard, C.L. 2022. *Gillespie algorithms for stochastic multiagent dynamics in*  
2 *populations and networks*. Cambridge University Press, Cambridge.
- 3 McCullough, K., Roberts, M.G. & Laing, C.R. (2016). Exact analytical expressions for the final  
4 epidemic size of an SIR model on small networks. *ANZIAM Journal*, 57, 429-444.
- 5 Méndez, V., Campos, D. & Bartumeus, F. (2014). *Stochastic foundations in movement ecology*.  
6 Springer Series in Synergetics. Springer, Heidelberg.
- 7 Mollison, D. (1991). Dependence of epidemic and population velocities on basic parameters.  
8 *Mathematical Biosciences*, 107, 255-287.
- 9 Moran, P.A.P. (1958). Random processes in genetics. *Mathematical Proceedings of the Cambridge*  
10 *Philosophical Society*, 54, 60-71.
- 11 Newton, I. (1687). *Philosophiæ Naturalis Principia Mathematica* (in Latin).
- 12 Nobile, M.S., Cazzaniga, P., Besozzi, D., Pescini, D. & Mauri, G. (2014). cuTauLeaping: A GPU-  
13 powered tau-leaping stochastic simulator for massive parallel analyses of biological  
14 systems. *PLoS one*, 9, e91963.
- 15 Okuyama, T. (2015). Demographic stochasticity alters the outcome of exploitation competition.  
16 *Journal of Theoretical Biology*, 365, 347-351.
- 17 Pineda-Krch, M. (2010). GillespieSSA: Implementing the stochastic simulation algorithm in R.  
18 *Journal of Statistical Software*, 25, 1-18.
- 19 Rhainds, M. (2019). Ecology of female mating failure/lifelong virginity: a review of causal  
20 mechanisms in insects and arachnids. *Entomologia Experimentalis et Applicata*, 167, 73-84.
- 21 Ricker, W.E. (1954). Stock and recruitment. *Journal of the Fisheries Research Board Canada*, 11,  
22 559-623.
- 23 Royama, T. (1992) *Analytical population dynamics*. London: Chapman & Hall.
- 24 Schenk, H., Schulenburg, H. & Traulsen, A. (2020). How long do Red Queen dynamics survive  
25 under genetic drift? A comparative analysis of evolutionary and eco-evolutionary models.  
26 *BMC Evolutionary Biology* 20, 8.
- 27 Serrao, S.R., Ritmeester, T. & Meyer-Ortmanns, H. (2021). Rare extinction events in cyclic  
28 predator–prey games. *Journal of Physics A*, 54, 235001.
- 29 Tolkamp, B.J., Haskell, M.J., Langford, F.M., Roberts, D.J. & Morgan, C.A. (2010). Are cows more  
30 likely to lie down the longer they stand? *Applied Animal Behaviour Science*, 124, 1–10.
- 31 Vestergaard, C.L. & Génois, M. (2015). Temporal Gillespie algorithm: fast simulation of contagion  
32 processes on time-varying networks. *PLoS Computational Biology*, 11, e1004579.
- 33 Volterra, V. (1927). Variazioni e fluttuazioni del numero d'individui in specie animali conviventi  
34 (in Italian). *Regio Comitato Talassografico Italiano, Memoria*, 131, 1–142.

1 Zug, R., Koehncke, A. & Hammerstein, P. (2012). Epidemiology in evolutionary time: the case of  
2 *Wolbachia* horizontal transmission between arthropod host species. *Journal of*  
3 *Evolutionary Biology*, 25, 2149-2160.

4

## 1 Figure legends

2 Figure 1. From rates to probabilities: (a) The probability that nothing has been observed to  
3 happen yet declines rapidly if the rate is high (orange curve: probability that no predator-prey  
4 kill has happened in the population since  $t = 0$ ), while the decline is less rapid for moderate rates  
5 (red curve: probability that a specific prey is still alive, if it was alive at  $t = 0$ ) and much shallower  
6 still for events that occur at a low rate (black curve: probability that a specific prey has avoided a  
7 specific predator individual, if both were alive at  $t = 0$ ). Because the curves have very different  
8 scales, (b) to (c) plot them separately over relevant time scales, revealing their exact same  
9 exponential shape. Parameters:  $a = 0.1$ ,  $N_1 = 1000$ ,  $N_2 = 100$ .

10

11 Figure 2. Comparison of deterministic and stochastic dynamics in the Lotka-Volterra predator-  
12 prey system. (a), four different limit cycles depicting deterministic predictions when starting the  
13 cycle at different densities for predators and prey, and (b-d) one run of a stochastic  
14 implementation of the same process using the Gillespie algorithm (starting at  $N_1 = 1000$ ,  $N_2 =$   
15  $100$ ). The purple arrow in (a) exemplifies the problem with the Euler equation: extrapolating  
16 from the current direction of change is prone to ‘miss the curve’. In (b-d), the first three cycles  
17 are in green, the next three in blue and the final three in red, to make it more visually clear that  
18 the system is not bound to always spiral outwards (from blue to red), the opposite can happen  
19 too (from green to blue), but the outwards movement is often observed last because it makes the  
20 population very prone to extinction. Other parameters:  $b=2$ ,  $B_1=5$ ,  $B_2=1$ ,  $a=0.1$ ;  $\mu=1$ ;  $n=150$ .

21

22 Figure 3. Deterministic dynamics predicted by equations (4a-c), exemplified for search efficiency  
23  $v=1$ , time out duration  $T = 0.1$ , and death rates  $\mu_F = \mu_M = 1$ . At  $t = 0$ , the dynamics are initialized  
24 with 1000 females and 1000 males, with no-one in ‘time out’ ( $m(0)=0$ ). The availability of  
25 searching males plummets initially with the same slope as the availability of females, but as the  
26 assumption is that males can resume searching while females do not become available again, the  
27 availability of searching males ( $M(t)$ ) begins to rise again, soon there are however no females left  
28 to mate with. The total number of males ( $m(t)+M(t)$ ) declines at a rate predicted by male  
29 mortality, thus at  $t = 0.5$  the expected number of males is  $1000e^{-0.5} = 606.53$ . Female  
30 matelessness can be computed based on the curves: the integral of female abundance  $F(t)$  with  
31 respect to time (the total area under the  $F(t)$  curve) is 2.7581, and since  $\mu_F = 1$  in this example,  
32 there are a total of  $2.7581\mu_F = 2.7581$  deaths while a female was still alive and waiting for a  
33 mating to happen (and death intervened first). Thus, the predicted matelessness here is a little  
34 less than 3 females out of 1000.

35

1 Figure 4. Heatmaps for the Gillespie algorithm's predictions for the matelessness problem, and in  
 2 (a), a comparison to the deterministic dynamics (contour lines for 1, 5 or 10 unmated females  
 3 out of 1000). Columns left to right indicate assumptions of no between-male variance ( $\sigma_v = \sigma_T =$   
 4 0), variance in searching efficiency  $v$  only ( $\sigma_v = 0.25, \sigma_T = 0$ ), variance in time-out  $T$  only ( $\sigma_v = 0,$   
 5  $\sigma_T = 0.25$ ), and variance in both ( $\sigma_v = \sigma_T = 0.25$ ). In each small panel, the x axis varies the  
 6 population-wide level of  $v$  and the y axis varies the same for  $T$  (strictly speaking these refer to  
 7 medians not means, see text for explanation). Top row: darker shades correspond to more  
 8 females remaining unmated. Second row: darker shades correspond to a higher value of the  
 9 selection differential on searching efficiency when measured as  $\ln v$ ; the figure remains entirely  
 10 white if males do not vary in this trait. Third row: darker shades indicate a larger selection  
 11 differential on time out when measured as  $\ln T$ ; the differentials are all negative, as shorter  $T$  is  
 12 beneficial, and their absolute values are shown. Again, no inter-male variation results in an  
 13 entirely white figure. Bottom row: The mating success of the most successful male in the entire  
 14 population, as the mean of 10 replicates, for each value of  $v$  and  $T$ . All other heatmap values are  
 15 likewise based on 10 replicates. Parameters that are not varied in the figure:  $\mu_F = \mu_M = 1$ .  
 16