

Stochastic modelling of prey depletion processes

Thomas Clerc^{a,*}, Anthony C. Davison^b, Louis-Félix Bersier^a

^a Unité d'écologie & évolution, Université de Fribourg, Ch. du Musée 10, CH-1700 Fribourg, Switzerland

^b Institut de Mathématiques, Ecole Polytechnique Fédérale de Lausanne, IMA-FSB-EPFL, Station 8, CH-1015 Lausanne, Switzerland

The modelling of prey–predator interactions is of major importance for the understanding of population dynamics. Classically, these interactions are modelled using ordinary differential equations, but this approach has the drawbacks of assuming continuous population variables and of being deterministic. We propose a general approach to stochastic modelling based on the concept of functional response for a prey depletion process with a constant number of predators. Our model could involve any kind of functional response, and permits a likelihood-based approach to statistical modelling and stable computation using matrix exponentials. To illustrate the method we use the Holling–Juliano functional response and compare the outcomes of our model with a deterministic counterpart considered by Schenk and Bacher [2002. Functional response of a generalist insect predator to one of its prey species in the field. *Journal of Animal Ecology* 71 (3), 524–531], who observed the depletion of *Cassida rubiginosa* due to its exclusive predator, *Polistes dominulus*. The predation was found to be Holling type III, reflecting the ability of the predator to regulate its prey. Our approach corroborates this result, but suggests that the prey depletion census should have been performed more often, and that predation features were significantly different between the two years for which data are available.

1. Introduction

The number of prey a predator kills per unit of time, termed the functional response (Solomon, 1949), is a central parameter for population dynamics, as its form governs the global stability of systems. In the biocontrol context, its shape can determine whether a predator is able to regulate its prey (Murdoch and Oaten, 1975).

A classical terminology for functional response types was proposed by Holling (1959). Apart from a few ecologists (see, e.g. Turchin, 2003, p. 82), most authors use Holling's classification of functional responses into three types, which can in principle be distinguished by plotting the functional response against the number of available prey. A type I functional response increases linearly up to a maximum and then remains constant. This form of predation seems to occur very rarely in nature and is thus almost always ignored for modelling. A type II functional response has a concave shape: it increases smoothly and tends asymptotically to a constant with a negative second derivative. In type III functional responses, the second derivative is first positive and then negative, but the curve is also asymptotically constant. Such a functional response is often called sigmoid or S-shaped (Fig. 1).

Generally, a type III functional response is said to reflect the ability of a predator to “learn” or to “switch” prey. Such predators focus on

alternative prey species when the density of the focal species is low, but switch to hunting it when this density becomes high (Murdoch, 1969). These generalist predators, unlike specialist ones, are desirable for biocontrol, so it is of high interest to determine whether a given predator has a type II or type III functional response.

Not only did Holling (1959) classify functional response types, but he also proposed models for the functional response, under which the predator is either searching for or is “handling” a prey. The two parameters involved in Holling's model are the searching rate, sometimes also called the attack rate, and interpreted as the area covered by a predator per unit of time, and the handling time, understood as the time needed for chasing, killing and eating a single prey. The general model is

$$g(N, P) = \frac{a(N)N}{1 + a(N)\tau N}, \quad (1)$$

where N and P represent the prey and the predator density, respectively, and $a(N)$ and τ denote the searching rate and the handling time.

When $a(N)$ is constant, the Holling type II functional response arises, whereas for type III the situation is less clear. Although Hassell (1978, p. 40) and more recently Juliano (2001) claim that type III functional responses can arise whenever $a(N)$ increases with the prey density N , this assertion is misleading. Even for the very general searching function of Juliano (2001),

$$a(N) = \frac{d + bN}{1 + cN}, \quad (2)$$

* Corresponding author. Tel.: +41 26 300 88 33.
E-mail address: thomas.clerc@unifr.ch (T. Clerc).

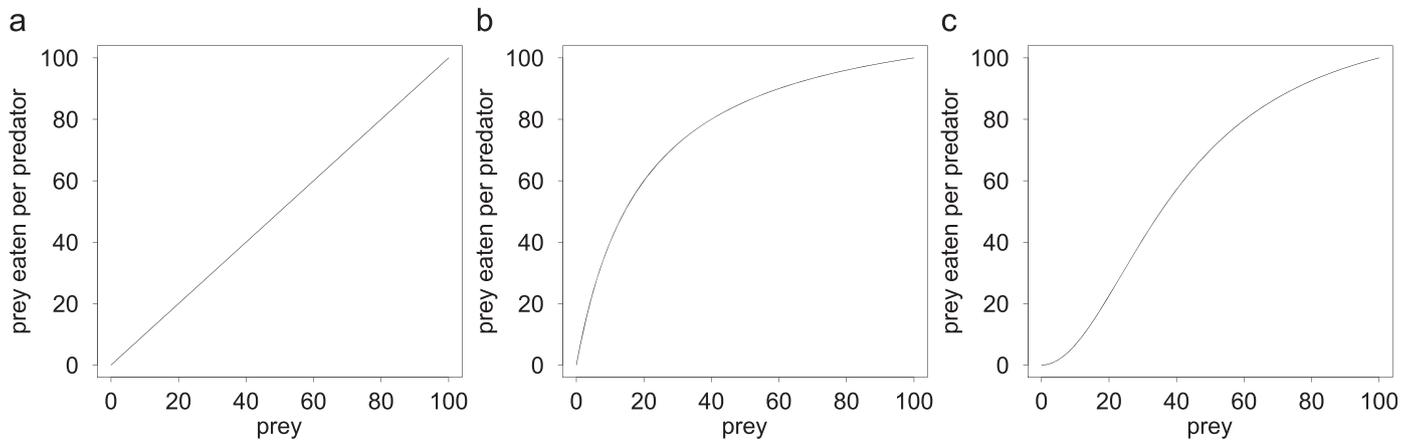


Fig. 1. The three types of functional responses in Holling's model: (a) type I or linear, (b) type II or hyperbolic and (c) type III or sigmoid.

the monotone increase condition is not sufficient. As explained in Appendix A, the supplementary condition $a'(0)/[a(0)]^2 > \tau$ assures type III functional response in the general model (Eq. (1)).

Many standard functional response models can be derived from the general model (Eq. (1)) with a Juliano searching rate function (Eq. (2)). The model proposed by Hassell (1978, p. 45) arises when $d = 0$. When $c = 0$ and $b = 0$, the functional response becomes a standard form of Holling type II functional response, whereas when $c = 0$ and $d = 0$, it reduces to a simple form of Holling type III functional response (Gotelli, 2001, p. 137). In the extreme case where every parameter is zero except d , the functional response is said to be linear.

Functional response modelling is itself exploited in the broader field of prey–predator system modelling. The classical models are represented using systems of ordinary differential equations (ODEs) such as

$$\begin{aligned} \frac{\partial N(t)}{\partial t} &= f_1(N(t)) + g_1(N(t), P(t)), \\ \frac{\partial P(t)}{\partial t} &= f_2(P(t)) + g_2(N(t), P(t)), \end{aligned} \quad (3)$$

where the functions $f_1(N(t))$ and $f_2(P(t))$ stand for the intrinsic dynamics of each respective population, and the functions $g_1(N(t), P(t))$ and $g_2(N(t), P(t))$ represent the functional response and the numerical response (Gotelli, 2001, pp. 127–128). As the numerical response is usually considered to be simply a fraction of the functional response and the intrinsic functions are easy to model, research has mainly focused on the functional response.

Almost all the (bitrophic) models developed so far in population dynamics share this common structure.

Although ODE systems have given valuable insight into population dynamics, they have many drawbacks. The variables $N(t)$ and $P(t)$ are taken to be continuous but represent discrete numbers of predators and prey. When the abundances are low, however, the distinction may be important. Moreover, the trajectories of populations under ODE models are deterministic: if the initial conditions are given, the dynamics are entirely known. Nature is not like this, so stochastic models may better account for the random nature of secretive processes like predation. In their book, Curry and Feldman (1987, p. 5) wrote: “In general, a stochastic model is more realistic and facilitates a proper validation analysis, but [...] is at present a greatly needed but much underdeveloped research area”. More than 20 years later, the situation has changed overall, but although many biologists use stochastic models, most areas of ecological modelling remain deterministic. In the modelling of functional response, Fenlon and Faddy (2006) reported only two early contributions

resting on stochastic theory (Curry and DeMichele, 1977; Curry and Feldman, 1979).

The present paper integrates the classical Holling models for functional responses into a stochastic prey depletion model, thus enabling a stochastic reappraisal of data first analysed through an ODE model. We will show it provides other interesting and useful results, all likelihood-based and thus inaccessible through ODE models.

2. Data

Many classical analyses of the functional response have used the data produced by Hassell et al. (1977) in laboratory conditions. The data from the experiment conducted by Schenk and Bacher (2002) that we shall use are of another kind. They studied the depletion of the shield beetle *Cassida rubiginosa* due to its exclusive predator, the paper wasp *Polistes dominulus*, in natural conditions. For the field biologist, this prey–predator system has two advantages: nearly 100% of the predation events were committed by *P. dominulus*, meaning that every removal can be imputed to this predator, a situation rare in nature; the beetles, which feed on creeping thistle *Cirsium arvense*, always remain on the same plant, which simplifies counting them. As the type III functional response occurs when the predator is able to switch prey, experiments intended to discriminate between types II and III functional responses must be performed in natural conditions.

Between July and August 2000, and again between July and August 2001, Schenk and Bacher (2002) observed prey depletion under constant predation for around 6 h each day. They added prey when abundances became too low. These observations were performed on two different sites: a natural patch of *C. arvense* in a larger uncultivated meadow in 2000 and an ecological compensation area within arable farmland in 2001. The estimated numbers of predators present on the sites were 3.05 and 2.46 for the years 2000 and 2001, respectively. Fig. 2 shows the prey dynamics for these experiments, more details of which may be found in the original paper.

3. Modelling and methods

3.1. Classical approach based on ODE

Schenk and Bacher (2002) used the functional response model proposed by Juliano (2001), which improves Holling's (1959) original model by taking into account the continuous depletion of

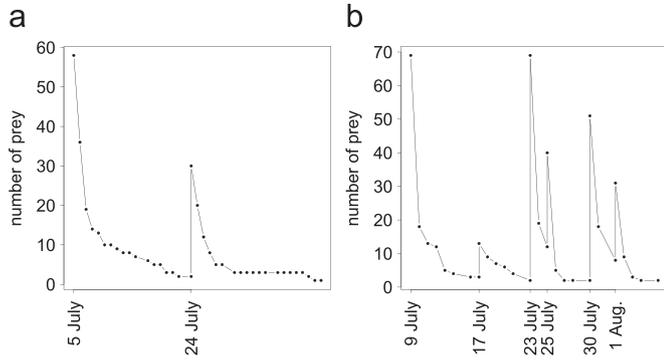


Fig. 2. Depletion of *C. rubiginosa* due to *P. dominulus* in the experiment conducted by Schenk and Bacher (2002) (a) in 2000 and (b) in 2001. The dates when the experimenters added further prey are shown on the horizontal axis.

prey throughout the experiment (suggested by Rogers, 1972) and by using a very general function for modelling the searching rate (see Eq. (2)). The number of prey eaten per predator N_e is

$$N_e = N_0 \left(1 - \exp \left[\frac{(d + bN_0)(\tau N_e - T)}{1 + cN_0} \right] \right), \quad (4)$$

where N_0 stands for the initial number of prey and T stands for the duration of the experiment. Discrimination between types II and III relies on logistic regression and maximum likelihood estimation (Juliano, 2001). Statistical analysis revealed that the attack rates were significantly correlated with the prey density for both 2000 and 2001. The logistic regression proposed by Juliano (2001) did not deliver clear-cut results, predicting a type II functional response for the year 2000 and a type III functional response for the year 2001. Schenk and Bacher (2002) overcame this using a standard nonlinear regression, which predicted a significantly positive b and negligible c and d in both cases (see Eq. (2)), indicating type III functional response for both years. However, this is a poor approach for the study of prey depletion because it assumes that correlated discrete responses with unequal variances are independent continuous homoscedastic random variables.

3.2. Markov model

The prey depletion process studied by Schenk and Bacher (2002) is treated as a discretely observed continuous time Markov chain. In the theory of stochastic processes, the continuous decrease of a population is commonly called a pure death process.

Let $N(t)$ be a random variable representing the number of prey alive at time t ; thus $N(0) = n_0$ is the known prey population at the beginning of the experiment. The number of predators $P(t) = p$ is supposed to be constant. As the process is assumed stationary (or time-homogeneous), the transition probability function for $t > 0$,

$$Pr_{m,n}(t) = Pr(N(t+u) = n | N(u) = m),$$

is independent of $u \geq 0$. The expression $Pr_{m,n}(t)$ should be read as “the probability that the prey population decreases from m to n individuals in a time t ”.

In the case of prey depletion due to predation, we make the following assumption about the infinitesimal probability transitions: for a small unit of time h ,

$$\begin{cases} Pr_{n,n-1}(h) = g(n,p)ph + o(h), \\ Pr_{n,n}(h) = 1 - g(n,p)ph + o(h), \end{cases} \quad (5)$$

where $g(n,p)$ is the functional response (see Eq. (1)) and $h^{-1}o(h) \rightarrow 0$ as $h \rightarrow 0$, so the error term $o(h)$ is negligible. The rationale behind these assumptions is outlined in Appendix B. Using the Chapman–Kolmogorov equations, one is able to derive the forward Kolmogorov differential equations (see, e.g. Taylor and Karlin, 1998, pp. 359–362):

$$\begin{cases} \frac{\partial}{\partial t} Pr_{n_0,n_0}(t) = -g(n_0,p)p Pr_{n_0,n_0}(t), \\ \frac{\partial}{\partial t} Pr_{n_0,n}(t) = g(n+1,p)p Pr_{n_0,n+1}(t) - g(n,p)p Pr_{n_0,n}(t). \end{cases} \quad (6)$$

Eq. (6) expresses the transition probability changes with respect to time (derivatives) in terms of functional responses and transition probabilities.

3.3. Functional response

One can replace the function $g(n,p)$ by any functional response, though the analytical solution of Eq. (6) might become quite complicated. As this paper is intended to reappraise the Schenk and Bacher (2002) results, we have considered the general Holling model for the functional response $g(n,p)$ (Eq. (1)), with Juliano’s (2001) function (Eq. (2)) for the searching rate.

3.4. Analytical solutions

In the framework of pure death processes, it is often possible to find an analytical solution to the differential system represented by Eq. (6). The rationale leading to this solution and a sketch of the proof are given in Appendix C.

Theorem C.1 makes it sometimes possible to obtain a simple formula for the transition probability, provided the functional response is not too complicated. Unfortunately, even with the simple form $a(n) = bn$, i.e. $g(n,p) = bn^2/(1 + b\tau n^2)$, the transition probability,

$$\begin{aligned} Pr_{n_0,k}(t) = & (1 + b\tau k^2) \sum_{j=k}^{n_0} (-1)^{j-k} \frac{\binom{n_0}{j} \binom{j}{k}^2}{\binom{2j-1}{j-k} \binom{n_0+j}{n_0-j}} \\ & \times (1 + b\tau j^2)^{n_0-1-k} \exp\left(\frac{-bj^2 p}{1 + b\tau j^2} t\right), \end{aligned} \quad (7)$$

is quite complex.

3.5. Matrix exponential method

Even with simple functional responses, analytical expressions for the transition probabilities do not yield stable numerical results. Indeed, in the above case (Eq. (7)), the binomial coefficients are such an enormous source of numerical errors that any computation relying on the analytical formula for the transition probabilities is very imprecise. As parameter estimation requires this kind of computation, an alternative method is needed. In any case, in Markov chains, analytical expressions for the transition probabilities are often unavailable, or not useful, as just mentioned. Thus, much effort has been devoted to numerical methods. Faddy and Fenlon (1999) gave a general method to compute the transition probabilities for a class of pure birth processes (extended Poisson processes); we briefly outline it, with some adaptations for pure death processes (see also Ross et al., 2006 for a more general example).

Let $\Pr_{n_0}(t) = (\Pr_{n_0,0}(t), \Pr_{n_0,1}(t), \dots, \Pr_{n_0,n_0}(t))$ denote the transition probability vector of the pure death process. We define

$$\frac{\partial}{\partial t} \Pr_{n_0}(t) = \left(\frac{\partial}{\partial t} \Pr_{n_0,0}(t), \dots, \frac{\partial}{\partial t} \Pr_{n_0,n_0}(t) \right).$$

We also define the matrix of transition rates

$$Q := \begin{pmatrix} -\lambda_0 & 0 & \dots & \dots & 0 \\ \lambda_1 & -\lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & -\lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & \lambda_{n_0} & -\lambda_{n_0} \end{pmatrix},$$

where $\lambda_j = g(j, p)p$. We can now rewrite the Kolmogorov forward equations (Eq. (6)) as matrices:

$$\frac{\partial}{\partial t} \Pr_{n_0}(t) = \Pr_{n_0}(t)Q, \quad (8)$$

whose solution is (Allen, 2003, p. 189)

$$\Pr_{n_0}(t) = \Pr_{n_0}(0) \exp(Qt).$$

In our case, $\Pr_{n_0}(0) = (\Pr_{n_0,0}(0), \dots, \Pr_{n_0,n_0}(0)) = (0, \dots, 0, 1)$, since $\Pr_{ij}(0) = \delta_{ij}$ (Kronecker's delta). Thus

$$\Pr_{n_0}(t) = (0, \dots, 0, 1) \exp(Qt), \quad (9)$$

so the transition probabilities are in the last row of $\exp(Qt)$.

The computation of matrix exponentials is not an easy task (Moler and Loan, 2003). Fortunately, a software package designed for Markov chain models, expokit (Sidje, 1998), has been implemented in MatLab. Many authors advocate its use in this context (Podlich et al., 1999; Ross et al., 2006). In most cases, computing the transition probabilities via the exponential matrix method is far more accurate and faster than computing them directly from expressions such as Eq. (7). The computational precision of these probabilities is crucial, because errors in small probabilities may result in negative numerical values.

3.6. Inference

Let $\theta = (\theta_1, \dots, \theta_k)$ represent the parameters of a model. If this is a Markov chain then the log-likelihood function is

$$\ell(\theta) = \sum_{k=1}^m \log \Pr_{n_{k-1}, n_k}(t_k - t_{k-1}; \theta) \quad (10)$$

for a given set of m observations $N(t_k) = n_k$ ($k = 1, \dots, m$), and we can use the expokit package to compute these transition probabilities. We treat the initial value n_0 for each downward trajectory in Fig. 2 as a constant, so the overall log-likelihood is a sum of expressions of the form (Eq. (10)); this is justified because the experimenters chose the numbers of prey to be added. We can then calculate the maximum likelihood estimator (MLE) $\hat{\theta}$ of the model parameters using the fmincon routine, a basic constrained optimization routine implemented in MatLab. Under mild conditions and in large samples, the MLE follows a k -dimensional normal distribution (see, e.g. Davison, 2003, p. 119),

$$\hat{\theta} \sim N_k(\theta^0, V),$$

where $V = I(\hat{\theta})^{-1}$ and $I(\theta)$ stands for the Fisher information matrix

$$I(\theta) = E \left(- \frac{\partial^2 \ell(\theta)}{\partial \theta \partial \theta^T} \right). \quad (11)$$

When θ is a vector, the i -th element of the MLE, $\hat{\theta}_i$ has approximately the $N(\theta_i^0, V_{ii})$ distribution, where θ_i^0 is the i -th element of the true parameter vector and V_{ii} is the (i, i) element of $I(\hat{\theta})^{-1}$, and confidence intervals are readily obtained. In practice it

is convenient and can provide better intervals if $I(\hat{\theta})$ is replaced by the Hessian matrix of second-order partial derivatives of the negative log-likelihood, which does not require the expectation that appears in Eq. (11). However when information about a parameter is limited it is better to base confidence intervals on the profile log-likelihood. Let ψ denote a scalar component of θ for which a confidence interval is required, let λ denote the remaining elements of θ , and let $\ell_p(\psi) = \max_{\lambda} \ell(\theta)$ denote the profile log-likelihood for ψ . Then standard theory (see, e.g. Davison, 2003, p. 128) shows that a $(1 - 2\alpha) \times 100\%$ confidence interval for ψ is the set

$$\{\psi : \ell_p(\psi) \geq \ell_p(\hat{\psi}) - \frac{1}{2}c_1(1 - 2\alpha)\},$$

where $c_1(1 - 2\alpha)$ is the $(1 - 2\alpha)$ -quantile of the χ_1^2 -distribution. This interval has the advantage of accommodating asymmetry in the log-likelihood, and of being invariant to changes in the parametrization of the model. This approach also allows the testing of nested hypotheses concerning model parameters, using the likelihood ratio statistic.

Eq. (11) involves second derivatives of the likelihood function and thus of the transition probabilities, but these can be extremely cumbersome. In Section 3.5, we have seen that direct computation of transition probabilities (Eq. (7)) can be very imprecise. As the derivatives become more complicated, we can anticipate more severe numerical problems. Podlich et al. (1999) point out that these derivatives of the transition probabilities can also be calculated through matrix exponentials. An adaptation of their method and some extensions concerning the log-likelihood function are outlined in Appendix D.

4. Results

4.1. Model comparison and parameter estimates

The full, 4-parameter, Holling-Juliano model results in maximum log-likelihoods of -38.15 and of -54.16 when fitted to the data for the years 2000 and 2001, respectively. The traditional Holling type II model, where the parameters b and c are zero, gives a far poorer fit, with maximum log-likelihoods equal to -54.58 and -82.95 (likelihood ratio test, $p < 10^{-3}$ and $p < 10^{-3}$). Removing the parameters c and d does not change the fit significantly, since the maximum log-likelihoods are -38.61 and -54.58 ($p = 0.63$ and 0.69). The model with $c = d = 0$ is usually presented as a minimalist version of the Holling type III model (Gotelli, 2001, p. 137). Dropping the parameter τ from the latter model, which corresponds to a linear functional response (sometimes termed a Holling type I functional response), gives fits as poor as in the Holling type II case, with -54.58 and -82.95 as respective maximum log-likelihoods ($p < 10^{-3}$ and $p < 10^{-3}$). Thus the Holling type III model is retained for the functional response in this experiment.

Under the deterministic Holling-Juliano model, Schenk and Bacher (2002) found that both b and τ were required, while c and d were not. Our stochastic model confirms their conclusion. The simple Holling model obtained when the searching rate function is $a(N) = bN$ turns out to be the best model. This model is readily checked to be type III; see Appendix A. As mentioned above, the type III model fits both data sets much better than does the type II.

Table 1 gives the parameter estimates for the deterministic Holling-Juliano model considered by Schenk and Bacher (2002) and for the reduced stochastic Holling-Juliano (H-J) model. Using video recording, Schenk and Bacher (2002) were able to observe directly a number of predation events, and compute the time needed to handle the prey and to eat it. Note the high variability in observation. Interestingly, the estimates for the handling time are

much higher than those directly observed, a result similar to that of Schenk and Bacher (2002). This, along with the quite wide confidence intervals, reflects the difficulty of estimating τ . The stochastic model provides even wider confidence intervals for τ than the deterministic one, but narrower confidence intervals for

b . Both models give estimates of the same overall orders of magnitude.

Fig. 3 reproduces the profile log-likelihoods for the parameters of the reduced Holling–Juliano stochastic model. The confidence intervals for the handling time τ are highly asymmetric,

Table 1

Comparison of parameter estimates, with 95% confidence intervals indicated by square brackets. Units of b are h^{-1} and units of τ are h .

	Year 2000		Year 2001	
<i>Schenk and Bacher (observed)</i>				
b	–		–	
τ	Between 0.0031 and 0.24 (median 0.015)		Between 0.0042 and 0.19 (median 0.02)	
<i>Schenk and Bacher (computed)</i>				
b	0.0039	[0.0029; 0.0049]	0.0074	[0.0052; 0.0096]
τ	0.17	[0.13; 0.21]	0.068	[0.046; 0.09]
<i>Stochastic H–J model (reduced)</i>				
b	0.0014	[0.00095; 0.0019]	0.004	[0.003; 0.0052]
τ	0.4	[0.1; 0.79]	0.09	[0; 0.18]

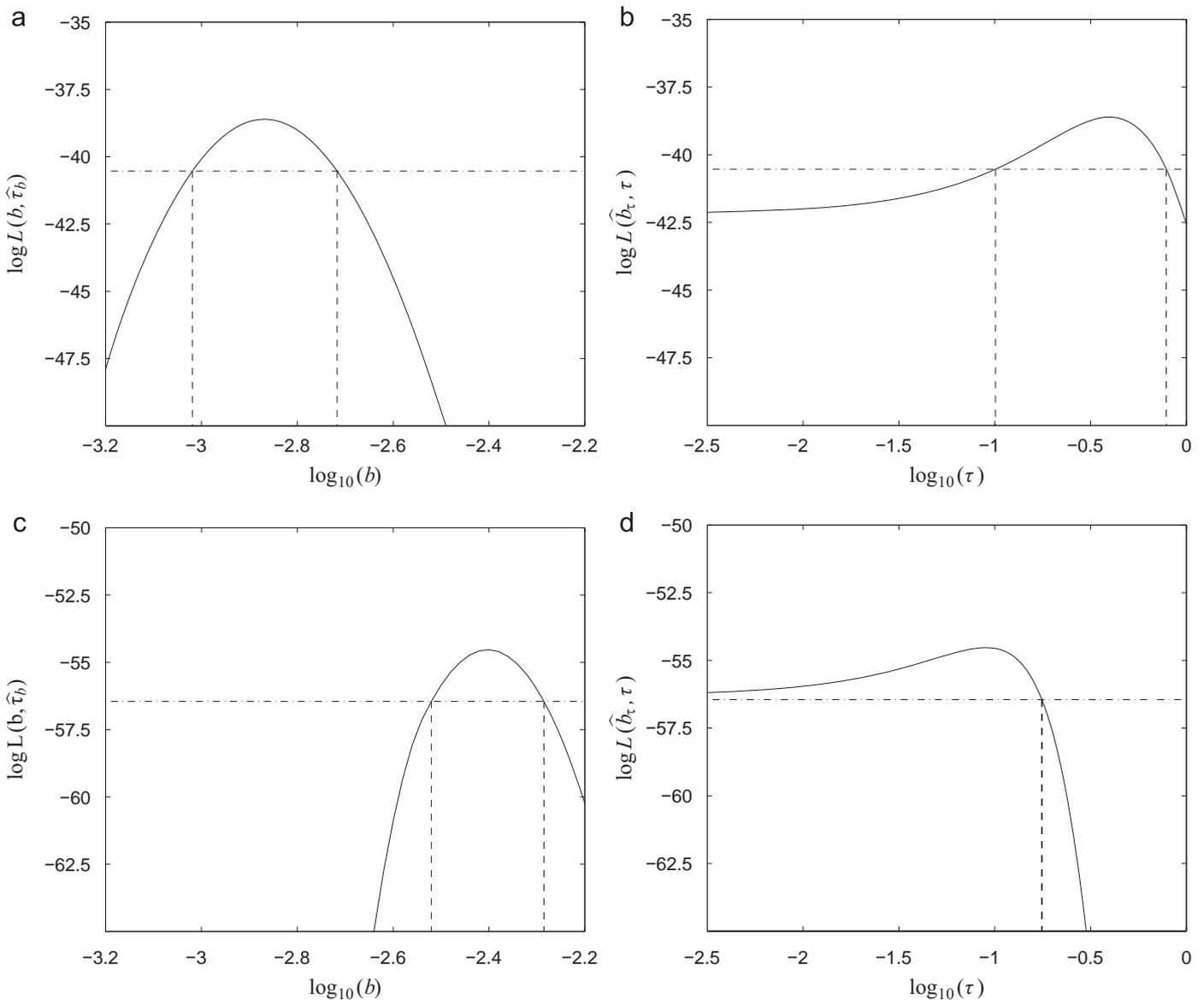


Fig. 3. Profile plots under the reduced Holling–Juliano model computed (a) for the parameters b , (b) τ with the 2000 data set, (c) for the parameters b and (d) τ with the 2001 data set. Dashed lines represent the 95% confidence interval limits.

particularly for 2001. For that year, the confidence interval has lower limit 0, which was set as lower bound in the optimization routine; of course the handling time cannot be negative. This means that for the 2001 data set, the handling time might well be zero. It also seems that the values differ sharply for the two years; we discuss this further below.

4.2. Sampling time optimum

The profile plots for the handling time τ (Figs. 3b and d) show the difficulty of estimating this parameter. A possible explanation is that the sampling time, i.e. the time between each census of the

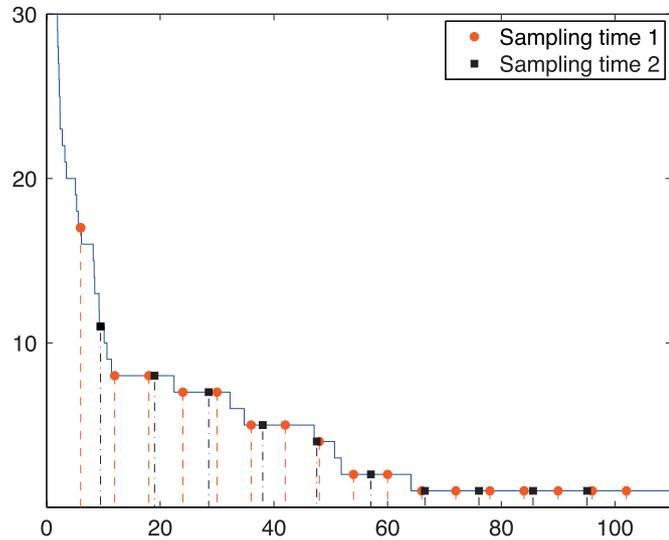


Fig. 4. Example of the construction of data sets for the different sampling times $t = 6$ and 9.51 . The depletion times are simulated with the following parameters: number of initial prey $n_0 = 60$, predators $p = 3.05$, searching rate $b = 0.0014$, handling time $\tau = 0.4$. To reflect the conditions prevailing during the experiment, the depletion is observed for 102 h of predation (corresponds to 17 days). The table on the right gives the counts.

prey population, might be so large compared to the handling time that the latter becomes essentially impossible to estimate precisely. More frequent counting would yield a bigger data set and this supplementary information could provide better estimates.

To check this, we used the diagonal elements of the inverse of Eq. (11) to compute the approximate variance of the maximum likelihood estimator $\hat{\tau}$ as a function of the sampling time t (see Appendix D for the computational method). Since we only have data sets whose sampling time is fixed we used simulated data to calculate the variance of $\hat{\tau}$ for different sampling times. In order to ensure that the results were relevant to the real data, we used the same initial numbers of prey and estimated numbers of predators as in the two experiments, and generated two data sets using the Holling type III model with the MLEs of b and τ , for a fixed observation period. In order to reduce Monte Carlo error we generated the pure death process in continuous time; this is easy because the Markov property (see Section 3.2) implies that the sojourn time T_k in a certain state k is exponentially distributed with mean $1/(g(k, p)p)$, where k is the number of prey, p the number of predators and $g(k, p)$ the functional response. For each simulated data set we then applied a range of sampling times, thus obtaining data analogous to the original data, and computed the maximum likelihood estimates of the parameters b and τ . Finally we computed the variance of the estimates of τ for each sampling time of interest. Fig. 4 illustrates this depletion process and gives example data sets generated using two different sampling times.

Figs. 5a and b show the variances of $\hat{\tau}$ as functions of the sampling time t . The variances from the simulated data seem to reflect those of the real process well, as the true variance of $\hat{\tau}$, i.e. the variance computed for the real data (where sampling time $t = 6$ h), matches the variance computed with the simulated data for the same sampling time almost perfectly. This suggests that the simulated data offer a reliable basis for inferences about the optimal sampling time.

For the data supposed to reflect the predation process of the year 2000, the optimum sampling time is near the effective sampling time, namely 6 h, cf. Fig. 5a. As the predation activity has

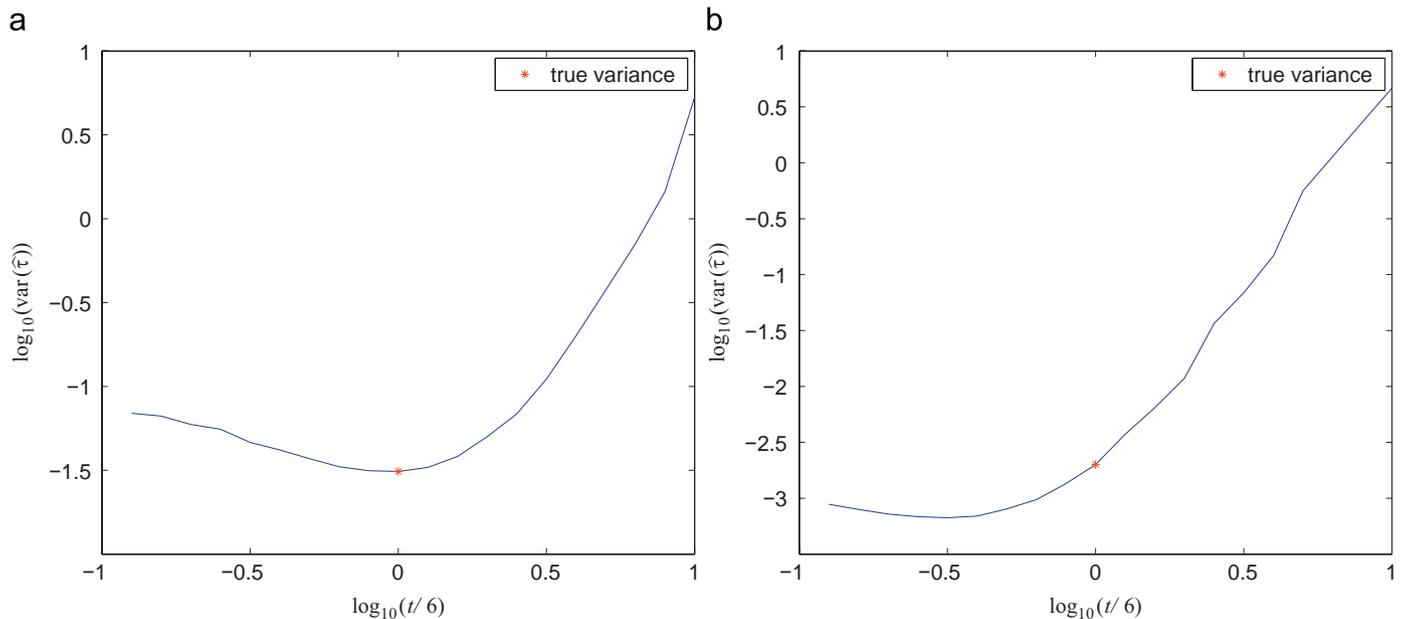


Fig. 5. Variance of the handling time estimate $\hat{\tau}$ with respect to the sampling time t evaluated (a) for a simulated data set close to the year 2000 data set and at the maximum likelihood estimates $b = 0.0014$ and $\tau = 0.4$ and (b) for a simulated data set close to the year 2001 data set and at the maximum likelihood estimates $b = 0.004$ and $\tau = 0.09$.

been estimated to last 6 h a day, a daily count of the remaining prey amounts to a count every 6 h. On the other hand, Fig. 5b suggests that the census of the remaining prey should have been done more frequently in the year 2001, for which the optimal sampling time is near 2 h.

Figs. 5a and b show that the variance initially decreases as t increases from the minimum value used. This is slightly counter-intuitive, but we believe that it stems from rounding errors that enter when the numbers of transitions become very large. Moreover, since the variation in the variance is quite small for small t , and because few transitions would be observed at small intervals, little supplementary information would be extracted by taking too small a sampling interval.

These comparisons suggest that the prey should be counted at intervals of between 2 and 6 h. The increase in work resulting from more frequent sampling is not rewarded by much lower errors on estimates, whereas less frequent sampling may lead to a large loss of information.

Although the variances look rather small, they are large in relation to the maximum likelihood estimates for τ , which are 0.4 for 2000 and 0.09 for 2001. The lowest variances are around 0.031 and 0.00067 for the two simulated data sets, so the minimal standard deviations are about 0.18 and 0.026, which represent roughly 50% and 30% of the corresponding estimates!

One explanation for the poor estimation of τ is that if small enough it has only a small influence on the prey depletion distribution, so for a given data set there are many plausible values for τ . Fig. 6 shows the probability distribution for the type III model with 30 initial prey, 2.46 predators (the estimated mean number of predators for 2001) and $b = 0.004$. The sampling time t and τ both vary. Clearly τ barely influences the distribution of the depletion of prey (Figs. 6b–d), except when τ exceeds the maximum likelihood estimate $\hat{\tau} = 0.09$ (Fig. 6a). This illustrates why τ cannot be removed from the model (see Section 4.1) and why the profile likelihood function (Figs. 3b and d) displays a dramatic decrease for values exceeding $\hat{\tau}$.

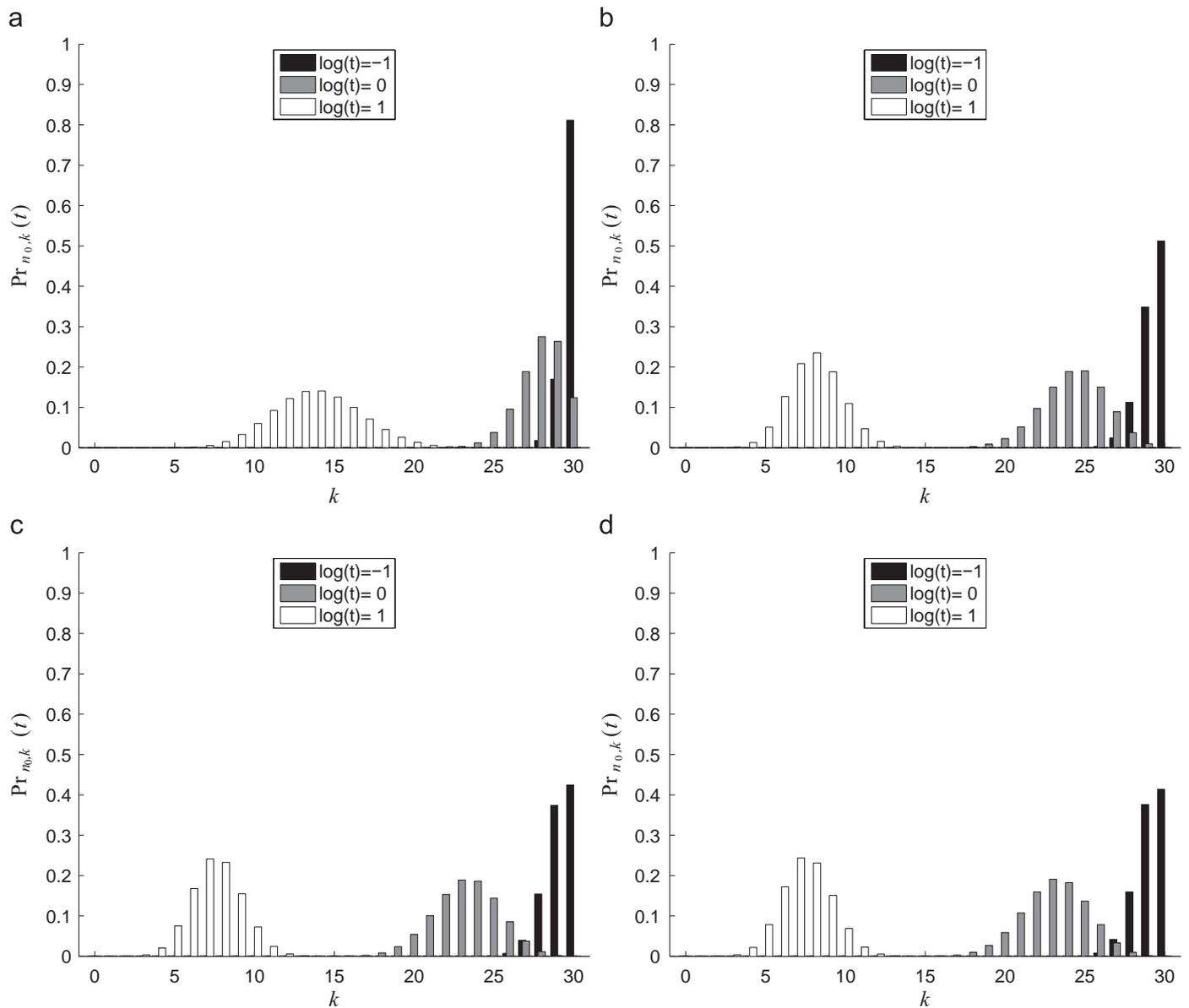


Fig. 6. Probability distributions of the prey depletion for $b = 0.004$, for given values of initial prey (30) and predator population (2.46) and various values of τ : (a) $\tau = 0.9$, (b) $\tau = 0.09$ (maximum likelihood estimate), (c) $\tau = 0.009$, (d) $\tau = 0.0009$ and t ($t = 0.1, 1, 10$). On the horizontal axis, the variable k represents the number of remaining prey.

4.3. Data set comparison

An interesting question is whether the two predation data sets of [Schenk and Bacher \(2002\)](#) have the same parameters. An affirmative answer is of course more desirable, since it would mean we could combine the two data sets. However the corresponding likelihood ratio test gives a highly significant value ($\chi^2_2 = 115.4$, p -value $p < 10^{-3}$), so we must strongly reject this hypothesis.

5. Discussion

The stochastic approach is an appealing way to model natural processes, and can be regarded as a good alternative to classical modelling based on ODEs. From a conceptual point of view, it provides a systematic approach to model biological phenomena such as the dynamics of small populations, which are naturally represented as random. From a practical point of view, the possibility of building a likelihood function enables statistical analyses that would be impossible with an ODE model.

The experiment we have considered represents a prey population depleted by predation. The main goal of the study was to determine if the predation is type II or type III. ODE modelling required [Schenk and Bacher \(2002\)](#) to use somewhat inappropriate statistical tools that do not account for the heteroscedasticity and correlatedness of the data. Discrimination between types II and III predation seems to be difficult, and several studies have concentrated on this topic ([Trexler et al., 1988](#); [Casas and Hulliger, 1994](#)). As we have shown, stochastic modelling provides a simple general framework in which all the statistical analysis may be based on the likelihood function, which is in principle easy to compute for a Markov chain.

The major drawback of the stochastic approach is numerical. We computed the probability distribution $\Pr_{n_0,k}(t)$ in the Holling type III functional response case (Eq. (7)) for different parameter values. In general, it turned out that for values of n_0 above 50, the probability distribution computation becomes erratic, with negative probabilities or probabilities bigger than 1. However, the exponential matrix method and expokit package overcome this, are in principle applicable for any discrete space Markov chain, and can be extended to produce further results about the variances of parameter estimates.

Another drawback of Markov chain is more intrinsic. The Markov assumption entails that given the number of available prey, the previous trajectory of the process can be ignored. In particular, this rules out the possibility of satiation. The number of prey could have dramatically decreased from 100 to 10, or it could have just slightly decreased from 11 to 10, this has no influence on the probability distribution of the future dynamics of the prey population, since only the present state matters! Yet, it intuitively has an influence, since the predators are more likely to be satiated in the first case and are thus less likely to consume many prey in the next time step. However, in [Schenk and Bacher's \(2002\)](#) experiment, satiation seems to play little role, as prey depletion is continuously abrupt. But the effect of satiation cannot be a priori neglected in all cases.

Concerning the experiment itself, stochastic modelling gave interesting results. First, it permitted us to assess whether the predation processes were comparable in 2000 and 2001, which is not the case. The very different parameter estimates for the two years might be explained by meteorological changes, though temperature, wind force, humidity and cloud cover were found to be non-significant in relation to predation by [Schenk and Bacher \(2002\)](#). A simpler explanation is the difference between the two sites: a partly uncultivated meadow surrounded by orchard trees in 2000, and a ecological compensation area within arable farmland in 2001. Second, our model clearly indicates a preference

for the type III and the parameter estimates are of the same order of magnitude as those of [Schenk and Bacher \(2002\)](#).

The estimate of the parameter b is rather hard to interpret. Even if this parameter refers to a phenomenological variable, namely the searching rate, it has not been measured, so it is impossible to appraise if the estimated value makes sense or not. The second parameter estimate is more readily interpreted. The estimate of the handling time τ for the year 2000, 24 min, greatly exceeds the maximum handling time measured during the experiment, of 871 s (15 min). [Schenk and Bacher \(2002\)](#) faced the same problem, as their estimates of τ systematically exceed the observed values. [Figs. 3b and d](#) reflect this estimation difficulty. The explanation that the sampling time might be too big is only partly relevant, as the variance of the estimator remains large even with the optimal sampling time.

These difficulties may stem from the models themselves. Both deterministic and stochastic models treat the handling time as constant, but the figures of [Schenk and Bacher \(2002\)](#) suggest that this is unrealistic. Bayesian or random effects models in which the handling time is treated as stochastic might be preferable, if the data set is sufficiently large. The difficulties might also originate from the choice of the functional response. We have chosen to use the Holling functional response in our depletion model in order to build a stochastic counterpart of the deterministic model considered by [Schenk and Bacher \(2002\)](#), which allowed a comparison of the parameter estimates of both approaches, but other functional responses might also have been used, and could lead to better estimates.

This work was meant as a stochastic reappraisal of classical ODE models. It has the advantage of being quite simple and very general, and it gives a likelihood-based framework that facilitates parameter estimation and model comparison. Future refinements of our approach could be undertaken in many directions. Recent developments, combining stochastic processes and Bayesian modelling, may be appropriate for reaching higher realism in predation and related biological phenomena.

Acknowledgement

This work was funded by the Swiss National Science Foundation (FNS) through the National Centre of Competence in Research "Plant Survival" (<http://www2.unine.ch/nccr>).

Appendix A

On differentiating the functional response twice, one finds

$$\frac{\partial^2 g(N, P)}{\partial N^2} = \frac{(1 + a(N)\tau N)(a''(N)N + 2a'(N)) - 2\tau(a'(N)N + a(N))^2}{(1 + a(N)\tau N)^3}.$$

Under the customary assumption that the functional response is asymptotically constant and under the usually accepted assumption that it does not have several plateaux, the sign of the second derivative of $g(N, P)$ at zero discriminates types II and III. The type III response then arises when the second derivative is positive, mathematically when $2a'(0) - 2\tau[a(0)]^2 > 0$. Provided $a(0) \neq 0$, this condition becomes

$$\frac{a'(0)}{(a(0))^2} > \tau.$$

This last condition translated in terms of the searching rate function proposed by [Juliano \(2001\)](#), $a(N) = (d + bN)/(1 + cN)$, is

$$\frac{b - cd}{d^2} > \tau. \quad (12)$$

As the first derivative is $a'(N) = (b - cd)/(1 + cN)^2$, the numerator in Eq. (12) is guaranteed to be positive because of the first assumption $a'(N) > 0$ ($a'(0) = b - cd$). However, as we see, the condition $b - cd > 0$ is not sufficient to ensure type III, since Eq. (12) entails $b - cd > \tau d^2$.

Appendix B

Let $g(n, p)$ be the functional response, i.e. the number of prey killed per predator and unit of time, where n stands for the total number of prey and p for the number of predators. Thus, the number of prey killed by p predators within a time h is $g(n, p)ph$. If we assume the prey population to be homogeneous, then the probability $\pi_n h$ for each individual prey to be killed in such a situation is just this last expression divided by n , i.e. $\pi_n h = g(n, p)ph/n$. Thus, the total number of prey killed $K_n(h)$ among n available prey within a time h when p predators are present is treated as a binomial variable with parameters n and $\pi_n(h)$. The probability that one death occurs is then

$$\begin{aligned} Pr(K_n(h) = 1) &= n\pi_n(h)(1 - \pi_n(h))^{n-1} \\ &= n\pi_n(h) \left(\sum_{j=0}^{n-1} \binom{n-1}{j} (-1)^j (\pi_n(h))^j \right) \\ &= n\pi_n(h) \left(1 + \sum_{j=1}^{n-1} \binom{n-1}{j} (-1)^j (\pi_n(h))^j \right) \\ &= n\pi_n(h) + \sum_{j=1}^{n-1} (j+1) \binom{n}{j+1} (-1)^j (\pi_n(h))^{j+1} \\ &= g(n, p)ph + \sum_{j=1}^{n-1} (j+1) \binom{n}{j+1} (-1)^j \left(\frac{g(n, p)p}{n} \right)^{j+1} h^{j+1} \\ &= g(n, p)ph + o(h). \end{aligned}$$

As the second term is a polynomial without any constant or first degree term, the limit $\lim_{h \rightarrow 0} (\sum_{j=1}^{n-1} (j+1) \binom{n}{j+1} (-1)^j (g(n, p)p/n)^{j+1} h^{j+1})/h = 0$, denoted by the little- o notation $o(h)$. Since this probability is equal to the probability that the prey population decreases from n to $n-1$ individuals, we have obtained the first equation of Eq. (5). On assuming that the elapsed time h is so small that only up to one event, one death in this case, might occur, the second equation follows.

Appendix C

It can be proved that in a Markovian process with discrete state space and continuous time parameter, the sojourn time in a certain state is an exponential random variable (see, e.g. Allen, 2003, p. 198). In the simple case of pure death process this property is very interesting. A Markov chain can be thought of as two other chains: the jump chain W_i , which consists in the stochastic process giving the state of the process before the i -th jump (or i -th state change), and the inter-event chain T_k , which gives the sojourn time in state k . Since our process is a pure death process, W_i takes all values between n_0 and 0 in decreasing order. Thanks to this division into two separate processes, it is possible to give an analytical formula for the transition probability.

Theorem C.1. *Let $N(t)$ be a Markovian pure death process, with initial state $N(0) = n_0$, where the interevent time distribution $T_k \sim \exp(1/\lambda_k)$ is exponential with mean $1/\lambda_k$. Then the transition probability is*

$$Pr_{n_0, k}(t) = \left(\prod_{i=k+1}^{n_0} \lambda_i \right) \sum_{j=k}^{n_0} \frac{\exp(-\lambda_j t)}{\prod_{\ell \neq j, \ell=k}^{n_0} (\lambda_\ell - \lambda_j)}.$$

In the modelling of a prey population depletion, the stochastic process $N(t)$ will stand for the population size at time t . The random variable T_k will be the sojourn time in state k , i.e. the time the population has k individuals. This variable is exponential with mean $1/\lambda_k$. If the pure death process models a prey depletion due to predation like in Eq. (5), the parameter λ_i is just the functional response multiplied by the number of predators, i.e. $g(n, p)p$.

In order to prove Theorem C.1, we need first to express the probability transition of $N(t)$ in terms of T_k . This can be done as follows:

Lemma C.2.

$$Pr_{n_0, k}(t) = Pr \left(\sum_{i=k+1}^{n_0} T_i \leq t \right) - Pr \left(\sum_{i=k}^{n_0} T_i \leq t \right), \quad k = 1, \dots, n_0 - 1. \quad (13)$$

The formulae for the cases $k = 0$ and n_0 are easily found.

Proof of C.1. The relation

$$Pr_{n_0, k}(t) = Pr \left(\left\{ \sum_{i=k+1}^{n_0} T_i \leq t \right\} \cap \left\{ \sum_{i=k}^{n_0} T_i > t \right\} \right)$$

expresses the fact that at time t the chain has already reached the state k but not the state $k-1$ yet. Basic manipulation of it yields Eq. (13). \square

The distribution of the two convolutions in Eq. (13) can be calculated via the following lemma:

Lemma C.3. *Let $(T_i)_{i=1, \dots, n}$, where $n \geq 2$, be independent exponential random variables with pairwise distinct parameters λ_i . Then the density of their sum is*

$$f_{\sum_{i=1}^n T_i}(t) = \left(\prod_{i=1}^n \lambda_i \right) \sum_{j=1}^n \frac{\exp(-\lambda_j t)}{\prod_{\ell \neq j, \ell=1}^n (\lambda_\ell - \lambda_j)}. \quad (14)$$

The proof is given in Ross (2006, p. 299).

Thus

$$\begin{aligned} Pr_{n_0, k}(t) &= Pr \left(\sum_{i=k+1}^{n_0} T_i \leq t \right) - Pr \left(\sum_{i=k}^{n_0} T_i \leq t \right) \\ &= \int_0^t \left(\prod_{i=k+1}^{n_0} \lambda_i \right) \sum_{j=k+1}^n \frac{\exp(-\lambda_j x)}{\prod_{\ell \neq j, \ell=k+1}^n (\lambda_\ell - \lambda_j)} dx \\ &\quad - \int_0^t \left(\prod_{i=k}^{n_0} \lambda_i \right) \sum_{j=k}^n \frac{\exp(-\lambda_j x)}{\prod_{\ell \neq j, \ell=k}^n (\lambda_\ell - \lambda_j)} dx \\ &\stackrel{\star}{=} \left(\prod_{i=k+1}^{n_0} \lambda_i \right) \sum_{j=k}^{n_0} \frac{\exp(-\lambda_j t)}{\prod_{\ell \neq j, \ell=k}^{n_0} (\lambda_\ell - \lambda_j)}. \end{aligned}$$

The equality marked with a \star involve basic algebraic manipulations. Its proof is left to algebra enthusiasts.

Appendix D

Starting from Eq. (8), Podlich et al. (1999) differentiated the equation on both sides with respect to a parameter, θ say. On reversing the order of differentiation and combining this new equation with Eq. (8), they established the elegant matrix equation:

$$\frac{\partial}{\partial t} \left(Pr_{n_0}(t), \frac{\partial Pr_{n_0}(t)}{\partial \theta} \right) = \left(Pr_{n_0}(t), \frac{\partial Pr_{n_0}(t)}{\partial \theta} \right) \begin{pmatrix} Q & \frac{\partial Q}{\partial \theta} \\ 0 & Q \end{pmatrix}. \quad (15)$$

In the context of Markov chains, the matrix Q usually involve rather simple functions of the parameters and is very sparse. Its differentiation with respect to one parameter poses therefore no problems. The authors were able to extend this formula to a complete formula, with first and second derivatives of all parameters. For two parameters θ_1 and θ_2 , this “all-in” formula is

$$\frac{\partial}{\partial t} \mathbb{P}_{n_0}(t) = \mathbb{P}_{n_0}(t) \begin{pmatrix} Q & \frac{\partial Q}{\partial \theta_1} & \frac{\partial Q}{\partial \theta_2} & \frac{\partial^2 Q}{\partial \theta_1^2} & \frac{\partial^2 Q}{\partial \theta_2^2} & \frac{\partial^2 Q}{\partial \theta_1 \partial \theta_2} \\ 0 & Q & 0 & 2 \frac{\partial Q}{\partial \theta_1} & 0 & \frac{\partial Q}{\partial \theta_2} \\ 0 & 0 & Q & 0 & 2 \frac{\partial Q}{\partial \theta_2} & \frac{\partial Q}{\partial \theta_1} \\ 0 & 0 & 0 & Q & 0 & 0 \\ 0 & 0 & 0 & 0 & Q & 0 \\ 0 & 0 & 0 & 0 & 0 & Q \end{pmatrix} = \mathbb{P}_{n_0}(t) Q^*,$$

where, say,

$$\mathbb{P}_{n_0}(t) = \left(\Pr_{n_0}(t), \frac{\partial \Pr_{n_0}(t)}{\partial \theta_1}, \frac{\partial \Pr_{n_0}(t)}{\partial \theta_2}, \frac{\partial^2 \Pr_{n_0}(t)}{\partial \theta_1^2}, \frac{\partial^2 \Pr_{n_0}(t)}{\partial \theta_2^2}, \frac{\partial^2 \Pr_{n_0}(t)}{\partial \theta_1 \partial \theta_2} \right).$$

Its solution is

$$\mathbb{P}_{n_0}(t) = \mathbb{P}_{n_0}(0)^T \exp(Q^* t), \tag{16}$$

where the vector $\mathbb{P}_{n_0}(0) = (0, \dots, 0, 1, 0, \dots, 0)$; it can be proved that only the term $\Pr_{n_0, n_0}(0) = 1$ is not equal to zero. When the number of parameters is greater than two, this matrix becomes too big and may cause computational issues. Should the case arise, Podlich et al. (1999) suggest using several smaller matrices instead. The first and second derivatives of the parameters $\theta_1, \dots, \theta_m$ are encompassed in the equations:

$$\frac{\partial}{\partial t} \begin{pmatrix} \Pr \\ \Pr^i \\ \Pr^{ii} \end{pmatrix}^T = \begin{pmatrix} \Pr \\ \Pr^i \\ \Pr^{ii} \end{pmatrix}^T \begin{pmatrix} Q & \frac{\partial Q}{\partial \theta_i} & \frac{\partial^2 Q}{\partial \theta_i^2} \\ 0 & Q & 2 \frac{\partial Q}{\partial \theta_i} \\ 0 & 0 & Q \end{pmatrix} \tag{17}$$

and the mixed second derivatives involving θ_i and θ_j can be wrapped up in the equations:

$$\frac{\partial}{\partial t} \begin{pmatrix} \Pr \\ \Pr^i \\ \Pr^j \\ \Pr^{ij} \end{pmatrix}^T = \begin{pmatrix} \Pr \\ \Pr^i \\ \Pr^j \\ \Pr^{ij} \end{pmatrix}^T \begin{pmatrix} Q & \frac{\partial Q}{\partial \theta_i} & \frac{\partial Q}{\partial \theta_j} & \frac{\partial^2 Q}{\partial \theta_i \partial \theta_j} \\ 0 & Q & 0 & \frac{\partial Q}{\partial \theta_j} \\ 0 & 0 & Q & \frac{\partial Q}{\partial \theta_i} \\ 0 & 0 & 0 & Q \end{pmatrix}, \tag{18}$$

where \Pr , \Pr^i and \Pr^{ij} are abbreviations for $\Pr_{n_0}(t)$, $\partial \Pr_{n_0}(t)/\partial \theta_i$ and $\partial^2 \Pr_{n_0}(t)/\partial \theta_i \partial \theta_j$, respectively. The solutions of Eqs. (17) and (18) are analogous to Eq. (16).

Differentiating the likelihood function (Eq. (10)) with respect to the parameter θ_i leads to

$$\frac{\partial \ell(\theta)}{\partial \theta_i} = \sum_{k=1}^m \frac{\Pr_{n_{k-1}, n_k}^i}{\Pr_{n_{k-1}, n_k}}$$

and differentiating it twice with respect to θ_i and θ_j yields

$$\frac{\partial^2 \ell(\theta)}{\partial \theta_i \partial \theta_j} = \sum_{k=1}^m \frac{\Pr_{n_{k-1}, n_k}^{ij} \Pr_{n_{k-1}, n_k} - \Pr_{n_{k-1}, n_k}^i \Pr_{n_{k-1}, n_k}^j}{(\Pr_{n_{k-1}, n_k})^2},$$

where $\Pr_{n_{k-1}, n_k}(t_k - t_{k-1} | \theta)$ its first derivative with respect to θ_i and its second derivative with respect to θ_i and θ_j have been abbreviated as \Pr_{n_{k-1}, n_k} , \Pr_{n_{k-1}, n_k}^i and \Pr_{n_{k-1}, n_k}^{ij} respectively. Eq. (11)

is finally computed as follows:

$$\begin{aligned} I(\hat{\theta}) &= E \left(- \frac{\partial^2 \ell(\hat{\theta})}{\partial \theta_i \partial \theta_j} \right) \\ &= E \left(\sum_{k=1}^m \frac{\Pr_{n_{k-1}, n_k}^i \Pr_{n_{k-1}, n_k}^j - \Pr_{n_{k-1}, n_k}^{ij} \Pr_{n_{k-1}, n_k}}{(\Pr_{n_{k-1}, n_k})^2} \right) \\ &= \sum_{k=1}^m E \left(\frac{\Pr_{n_{k-1}, n_k}^i \Pr_{n_{k-1}, n_k}^j - \Pr_{n_{k-1}, n_k}^{ij} \Pr_{n_{k-1}, n_k}}{(\Pr_{n_{k-1}, n_k})^2} \right) \\ &= \sum_{k=1}^m \sum_{x=0}^{n_k-1} \left(\frac{\Pr_{n_{k-1}, x}^i \Pr_{n_{k-1}, x}^j - \Pr_{n_{k-1}, x}^{ij} \Pr_{n_{k-1}, x}}{(\Pr_{n_{k-1}, x})^2} \right) \Pr_{n_{k-1}, x} \\ &= \sum_{k=1}^m \sum_{x=0}^{n_k-1} \left(\frac{\Pr_{n_{k-1}, x}^i \Pr_{n_{k-1}, x}^j - \Pr_{n_{k-1}, x}^{ij} \Pr_{n_{k-1}, x}}{\Pr_{n_{k-1}, x}} \right). \end{aligned}$$

Thus, it turns out that the second derivatives of the log-likelihood function which arise in Eq. (11) only consist of second derivatives of the transition probabilities, which are computable through matrix exponentials.

References

Allen, L.J.S., 2003. An Introduction to Stochastic Processes with Applications to Biology. Pearson Education Upper Saddle River, NJ.

Casas, J., Hulliger, B., 1994. Statistical analysis of functional response experiments. *Biocontrol Science and Technology* 4 (2), 133–145.

Curry, G.L., DeMichele, D.W., 1977. Stochastic analysis for the description and synthesis of predator–prey systems. *Canadian Entomologist* 109, 1167–1174.

Curry, G.L., Feldman, R.M., 1979. Stochastic predation model with depletion. *Canadian Entomologist* 111, 465–470.

Curry, G.L., Feldman, R.M., 1987. *Mathematical Foundations of Population Dynamics*. Published for the Texas Engineering Experiment Station, Texas A and M University System, Texas A and M University Press College Station, USA.

Davison, A.C., 2003. *Statistical Models*. Cambridge University Press, Cambridge.

Faddy, M.J., Fenlon, J.S., 1999. Stochastic modelling of the invasion process of nematodes in fly larvae. *Journal of the Royal Statistical Society Series C—Applied Statistics* 48 (1), 31–37.

Fenlon, J.S., Faddy, M.J., 2006. Modelling predation in functional response. *Ecological Modelling* 198 (1–2), 154–162.

Gotelli, N.J., 2001. *A Primer of Ecology*. Sinauer Associates Inc, Sunderland, MA.

Hassell, M.P., 1978. *The Dynamics of Arthropod Predator–Prey Systems*. Princeton University Press, Princeton, NJ.

Hassell, M.P., Lawton, J.H., Beddington, J.R., 1977. Sigmoid functional responses by invertebrate predators and parasitoids. *Journal of Animal Ecology* 46 (1), 249–262.

Holling, C.S., 1959. The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. *Canadian Entomologist* 91 (5), 293–320.

Juliano, S.A., 2001. Nonlinear curve fitting predation and functional response curves. In: Scheiner, S.M., Gurevitch, J. (Eds.), *Design and Analysis of Ecological Experiments*. Chapman & Hall, New York, pp. 159–182.

Moler, C., Loan, C.V., 2003. Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later. *SIAM review (Print)* 45 (1), 3–49.

Murdoch, W.W., 1969. Switching in general predators: experiments on predator specificity and stability of prey populations. *Ecological Monographs* 39 (4), 335–354.

Murdoch, W.W., Oaten, A., 1975. Predation and population stability. *Advances in Ecological Research* 9 (1).

Podlich, H.M., Faddy, M.J., Smyth, G.K., 1999. Likelihood computations for extended poisson process models. *InterStat*, September 1.

Rogers, D., 1972. Random search and insect population models. *Journal of Animal Ecology* 41, 369–383.

Ross, J.V., Taimre, T., Pollett, P.K., 2006. On parameter estimation in population models. *Theoretical Population Biology* 70 (4), 498–510.

Ross, S.M., 2006. *Introduction to Probability Models*, ninth ed. Academic Press, New York.

Schenk, D., Bacher, S., 2002. Functional response of a generalist insect predator to one of its prey species in the field. *Journal of Animal Ecology* 71 (3), 524–531.

Sidje, R.B., 1998. Expokit: a software package for computing matrix exponentials. *ACM Transactions on Mathematical Software (TOMS)* 24 (1), 130–156.

Solomon, M.E., 1949. The natural control of animal populations. *Journal of Animal Ecology* 18 (1), 1–35.

Taylor, H.M., Karlin, S., 1998. *An Introduction to Stochastic Modeling*, third ed. Academic Press, New York.

Trexler, J.C., McCulloch, C.E., Travis, J., 1988. How can the functional response best be determined? *Oecologia* 76 (2), 206–214.

Turchin, P., 2003. *Complex Population Dynamics: A Theoretical/Empirical Synthesis*. Princeton University Press, Princeton, NJ.