

# Bayesian Inference for Functional Response in a Stochastic Predator–Prey System

Gianni Gilioli<sup>a</sup>, Sara Pasquali<sup>b,\*</sup>, Fabrizio Ruggeri<sup>b</sup>

<sup>a</sup> *Dipartimento GESAF, Università di Reggio Calabria, Piazza S. Francesco di Sales 4,  
89061 Gallina di Reggio Calabria, Italy*

<sup>b</sup> *CNR-IMATI, Via Bassini 15, 20133 Milano, Italy*

Received: 24 May 2006 / Accepted: 22 June 2007 / Published online: 16 August 2007  
© Society for Mathematical Biology 2007

**Abstract** We present a Bayesian method for functional response parameter estimation starting from time series of field data on predator–prey dynamics. Population dynamics is described by a system of stochastic differential equations in which behavioral stochasticities are represented by noise terms affecting each population as well as their interaction. We focus on the estimation of a behavioral parameter appearing in the functional response of predator to prey abundance when a small number of observations is available. To deal with small sample sizes, latent data are introduced between each pair of field observations and are considered as missing data. The method is applied to both simulated and observational data. The results obtained using different numbers of latent data are compared with those achieved following a frequentist approach. As a case study, we consider an acarine predator–prey system relevant to biological control problems.

**Keywords** Stochastic predator–prey system · Functional response · Bayesian inference · MCMC algorithms · Latent data

## 1. Introduction

Parameter estimation for the functional response of predator to prey abundance is a critical methodological problem in population ecology (McCallum, 2000). The estimation procedure can rely on classical experimental approaches performed in simple arenas using different densities of prey (McCallum, 2000). These approaches have been criticized for the oversimplification of the environment in which predator behavior occurs (Kareiva, 1990). More complex experimental setups can be considered by manipulating environmental variables (Gilioli et al., 2005) and spatial organization of the two species, as well as elements of experimental environment (e.g. spatial arrangement of the plants that represent the resource for the prey) (Kareiva, 1982). Furthermore, attempts have been performed to solve the problem of scaling up the results from the laboratory to the field.

---

\*Corresponding author.

E-mail address: sara@mi.imati.cnr.it (Sara Pasquali).

A common approach is to break down predator and prey behaviors into basic components and to determine how each of them responds to factors that might be important in the field (Casas et al., 2004). However, experimental approaches encounter important limitations due to artificial conditions needed to follow the behavioral component of single individuals (Xia et al., 2003), and methodological and theoretical problems arise concerning the possibility of defining the effects of functional response at a population level combining models of single behavioral traits. The problems of scaling up may be even more complex if the goal is to produce realistic simulations of predator impacts on prey populations under field conditions: in this case, behavior interacts with demography of the two species (Xia et al., 2003). To avoid such constraints, different approaches relying on field observations have been proposed. They have been based mostly on the analysis of change in equilibrium or mean abundance over a gradient of equilibrium (Jost and Ellner, 2000). However, the assumption of equilibrium for prey–predator population systems is difficult to validate. In fact, many systems are characterized by limited persistence, often due to the availability of plant resources, or are under the effect of environmental variation and influenced by time dependent patterns of forcing variables. To overcome the limitations of the equilibrium assumption, it is possible to use time series data on prey and predator abundances, thereby exploiting the advantages offered by natural conditions to generate more realistic parameter estimates. To deal with the problem of inferring feeding rate indirectly from sampling population dynamics, it is necessary to consider two general assumptions. First, it is necessary to assume a general model of the population dynamics from which the functional response is estimated. This model has to be able to capture the pattern of the real population dynamics components (prey and predator numerical response and predator functional response). Second, the defined biological system is assumed to be closed in terms of fluxes of individuals and protected against the interference of factors not represented in the model (i.e. the presence of other species such as alternative preys for the predator or the presence of competitive predators).

The problem of parameter estimation from time series of predator–prey abundance data can be brought back to parametric inference for discretely observed diffusion processes which has been widely studied through the likelihood function by many authors (Pedersen, 1995a, 1995b; Bibby and Sørensen, 1995; Kessler and Sørensen, 1999; Sørensen, 1999; Kessler and Parades, 2002; Durham and Gallant, 2002; Aït-Sahalia, 2002, 2006; Beskos et al., 2006; Stramer and Yan, 2007). Prakasa Rao (1999) and Sørensen (2004) summarize several methods for parameter estimation in diffusion processes with data derived from both discrete sample data or continuous sample paths. Unfortunately, very often, the estimator obtained in the case of discrete data is consistent and asymptotically normal only when the time intervals between consecutive sampling are small with respect to the total duration of observations. Given that in many field surveys the number of observations available is small, particularly in seasonal or transient environment, the estimator obtained using the likelihood function may be unreliable. The Bayesian approach allows this problem to be overcome within a Markov Chain Monte Carlo (MCMC) framework in which latent observations are inserted between two consecutive observations and treated as missing data or nuisance parameters (Tanner and Wong, 1987; Eraker, 2001; Elerian et al., 2001; Golightly and Wilkinson, 2005, 2006). In particular, Elerian et al. (2001) apply a random-block size Metropolis–Hastings algorithm to generate latent data for univariate, one-factor diffusion processes, while Eraker (2001) and Golightly and Wilkinson (2005) update latent observations of multivariate models in fixed blocks using a

single site Gibbs sampler. Finally, Golightly and Wilkinson (2006) use sequential MCMC methods to sample the posterior distribution of latent data and parameters adapting the ‘modified bridge’ approach proposed by Durham and Gallant (2002). Relying on the advantages offered by the Bayesian approach, in this paper, we present a new approach for the parameter estimation in the functional response of predator to prey abundance and apply it to field data. In particular, we focus on the formulation of a stochastic predator–prey model and propose a Bayesian approach when few sampling data are available. Latent data are generated following Eraker (2001) and Golightly and Wilkinson (2005, 2006), but here we update simultaneously latent data between two consecutive real observations. In Section 2, we present the predator–prey population dynamics model used in the estimation of functional response which depends linearly on a single parameter. This condition allows us to find a classical estimator in closed form. In Section 3, we present classical and Bayesian approaches to estimate the parameter in the case of a discrete number of observations. To deal with a small number of observations, in Section 4, we introduce latent data between each pair of observations to improve the results, and we analyze the consequences of the use of different numbers of latent data. In Section 5, we present an application of the proposed method to the acarine predator–prey system *Phytoseiulus persimilis*—*Tetranychus urticae*, both in the case of simulated data and in the case of data collected in a strawberry crop in Ispica (Ragusa, Italy).

## 2. The model

Indirect estimation of functional response from population dynamics data relies on the assumption that the population system follows a well-defined pattern that can be represented by a predator–prey mathematical model. Due to the lack of information on model components, we select a simple model formulation that involves minimum assumptions on numerical and functional responses. As a consequence, we adopt the basic functional response in the Lotka–Volterra model that provides the advantage of linearity (Turchin, 2003) at the cost of the simplified biological assumption of an unlimited predator per-capita consumption rate (Royama, 1971). We are planning to investigate the capability of the method considering other functional responses in a forthcoming paper.

We consider, on the time interval  $[0, t_{\text{fin}}]$ , a predator–prey system described by the following equations

$$\begin{cases} dx_t = [rx_t(1 - x_t) - qx_t y_t] dt, & x(0) = x_0, \\ dy_t = [cq x_t y_t - uy_t] dt, & y(0) = y_0, \end{cases} \quad (1)$$

where  $x_t$  and  $y_t$  are the biomass of prey and predator at time  $t$  per spatial habitat unit normalized with respect to the carrying capacity;  $r$  is the specific growth rate of the prey,  $c$  is the maximum specific production rate of the predator,  $u$  is the specific loss rate of the predator due to natural mortality and  $q$  is a positive constant representing the rate of effective search per predator, that is the proportion of the universe searched per predator per unit time (Gutierrez, 1996). The lumped parameters  $r$ ,  $c$ ,  $u$  (thoroughly described in Buffoni and Gilioli, 2003) are known, as the initial condition  $(x_0, y_0)$ . The behavioral parameter  $q$  in the linear functional response  $qx_t y_t$  is unknown and has to be estimated. System (1) is a modified Lotka–Volterra model which considers logistic growth of the prey

to account for intra-specific competition. Its solution is in the compact set  $[0, 1] \times [0, 1]$  for the initial condition  $(x_0, y_0)$  in the same compact.

We now formulate a stochastic extension of model (1). Following Shaffer (1981, 1987), uncertainty in ecology can be classified in phenotypic variation, demographic variation, environmental variation, and catastrophes. This taxonomy of uncertainty deals mainly with population models (Akçakaya, 2000; Regan et al., 2002) and epidemic models (Rand and Wilson, 1991). Chesson (1978) recognized that variability is also fundamental at behavior level since individual predators and preys are variable and they must respond to highly variable circumstances. However, despite its importance, uncertainty affecting behavioral responses is less considered in predator–prey models (Bonsall and Hastings, 2004). In predator–prey models, behavioral aspects are encompassed in the functional response that represents the outcome of stochastic search and capture processes (Chesson, 1978). In our stochastic extension to model (1), we assign a primary role to behavior and introduce the term  $qx_t y_t$  as source of variability, modifying the functional response parameter  $q$  as follows

$$q_t = q_0 + \sigma \xi_t, \quad \xi_t \text{ Gaussian white noise,} \quad (2)$$

so that  $q$  fluctuates around its mean value  $q_0$ , supposed unknown.

The Gaussian assumption simplifies computations remarkably but it has the drawback of allowing for negative values of  $q$ , which are in contrast with the biological interpretation of the parameter. To satisfy such biological constraint we suppose, as proposed in many statistical works, that mean and variance keep the random variable far away from negative values. Such condition is actually satisfied in the case study of Section 5. Substituting (2) in (1), we obtain the stochastic system

$$\begin{cases} dx_t = [rx_t(1 - x_t) - q_0 x_t y_t] dt - \sigma x_t y_t dw_t^{(1)}, & x(0) = x_0, \\ dy_t = [cq_0 x_t y_t - uy_t] dt + c\sigma x_t y_t dw_t^{(1)}, & y(0) = y_0. \end{cases} \quad (3)$$

We consider a second component of stochasticity affecting both the predator and prey populations. These variations can be interpreted as environmental stochasticity deriving from environmental variables fluctuation (Pascual and Kareiva, 1996). In our formulation, the variations do not affect vital rates or carrying capacity but only population abundance, hence they can be assigned to catastrophes (Akçakaya, 2000). The consideration of an additive noise for the two species depending on population density can be also interpreted as observation errors affecting population abundance estimates (Carpenter et al., 1994; Pascual and Kareiva, 1996). We obtain the system

$$\begin{cases} dx_t = [rx_t(1 - x_t) - q_0 x_t y_t] dt - \sigma x_t y_t dw_t^{(1)} + \varepsilon x_t dw_t^{(2)}, & x(0) = x_0, \\ dy_t = [cq_0 x_t y_t - uy_t] dt + c\sigma x_t y_t dw_t^{(1)} + \eta y_t dw_t^{(2)}, & y(0) = y_0, \end{cases} \quad (4)$$

where  $w^{(2)}$  is a Wiener process independent from  $w^{(1)}$  while  $\varepsilon$  and  $\eta$  are two positive parameters.

The solution to system (4) is not necessarily constrained in the compact  $[0, 1] \times [0, 1]$ , because of the Wiener process. Since normalized biomass is supposed to be in unit interval, we introduce a function  $\chi$ , continuously differentiable and Lipschitz, taking value 1 in the compact  $[\theta, 1 - \theta]$ , increasing in the interval  $(-\infty, \theta)$ , decreasing in the interval

$(1 - \theta, \infty)$  and such that  $\lim_{z \rightarrow -\infty} \chi(z) = \lim_{z \rightarrow \infty} \chi(z) = 0$  and  $\chi(0) = \chi(1) = \theta$ . It converges to the characteristic function of  $[0, 1]$  when  $\theta$  goes to zero and forces the solution of (3) to lie in  $(0, 1) \times (0, 1)$  for the initial condition  $(x_0, y_0)$  in the same space. System (4) becomes

$$\begin{cases} dx_t = [rx_t(1 - x_t) - q_0x_t y_t] \chi(x_t) dt \\ \quad - \sigma x_t y_t \chi(x_t) dw_t^{(1)} + \varepsilon x_t \chi(x_t) dw_t^{(2)}, & x(0) = x_0, \\ dy_t = [cq_0x_t y_t - uy_t] \chi(y_t) dt \\ \quad + c\sigma x_t y_t \chi(y_t) dw_t^{(1)} + \eta y_t \chi(y_t) dw_t^{(2)}, & y(0) = y_0, \end{cases} \quad (5)$$

The parameters  $q_0, \sigma, \varepsilon, \eta$  are unknown and have to be estimated. However, the main goal of the paper is the illustration of an efficient method to estimate the parameter  $q_0$  because it allows to interpret and quantify how fundamental biological properties (strategies and efficiencies of search and capture) of the predator are expressed in a natural context in relation to a given prey, and how such properties can be translated into the biological control potentialities. A prior distribution will be specified on  $q_0$  and inference will be performed based upon its posterior distribution. The estimation of the remaining parameters  $\sigma, \varepsilon, \eta$  is also of interest and prior distributions could be specified on them. The resulting inference procedure would be considerable more complex than the one entertained here, where we are interested in the relatively simple problem of estimating  $q_0$ . We plan to consider joint estimation of all model parameters in a forthcoming paper. Here we take an empirical Bayes approach (see e.g. Carlin and Louis, 2000), since we look for “optimal” point estimates of the parameters  $\sigma, \varepsilon, \eta$  and then plug them into the model.

For simplicity, we write system (5) in vectorial form

$$dX_t = \mu(X_t, q_0) dt + \beta(X_t) dW_t, \quad (6)$$

where  $X_t = [x_t, y_t]^T$ ,

$$\begin{aligned} \mu(X_t, q_0) &= \begin{bmatrix} [rx_t(1 - x_t) - q_0x_t y_t] \chi(x_t) \\ [cq_0x_t y_t - uy_t] \chi(y_t) \end{bmatrix}, \\ \beta(X_t) &= \begin{bmatrix} -\sigma x_t y_t \chi(x_t) & \varepsilon x_t \chi(x_t) \\ c\sigma x_t y_t \chi(y_t) & \eta y_t \chi(y_t) \end{bmatrix}. \end{aligned}$$

The coefficients  $\mu$  and  $\beta$  are bounded and continuously differentiable with respect to  $X$  and  $q_0$ . Moreover, they satisfy the conditions for the existence and uniqueness of a strong solution of a stochastic differential equation (see, for example, Øksendal, 1998).

### 3. Classical and Bayesian inference

Under the classical paradigm, an estimate for  $q_0$  is found equaling to zero the continuous score function obtained deriving the log-likelihood (Liptser and Shirayayev, 1977)

$$\begin{aligned} l(q_0) = \exp \Big\{ & - \int_0^{t_{\text{fin}}} \mu^T(X_t; q_0) [\beta(X_t) \beta^T(X_t)]^{-1} dX_t \\ & + \frac{1}{2} \int_0^{t_{\text{fin}}} \mu^T(X_t; q_0) [\beta(X_t) \beta^T(X_t)]^{-1} \mu(X_t; q_0) dt \Big\} \end{aligned}$$

with respect to  $q_0$ . In system (5), the parameter  $q_0$  appears linearly and the coefficient  $\mu$  in Eq. (6) can be decomposed in the form  $\mu(X_t; q_0) = a(X_t)q_0 + b(X_t)$  with

$$a(X_t) = \begin{bmatrix} -x_t y_t \chi(x_t) \\ c x_t y_t \chi(y_t) \end{bmatrix}; \quad b(X_t) = \begin{bmatrix} r x_t (1 - x_t) \chi(x_t) \\ -u y_t \chi(y_t) \end{bmatrix}.$$

It follows that, in this special case, the score function is

$$S(q_0) = - \int_0^{t_{\text{fin}}} a^T(X_t) [\beta(X_t) \beta^T(X_t)]^{-1} dX_t \\ + \int_0^{t_{\text{fin}}} a^T(X_t) [\beta(X_t) \beta^T(X_t)]^{-1} \mu(X_t; q_0) dt.$$

Since we have only a discrete number of observations available, we approximate the score function replacing Lebesgue and Itô integrals by Riemann and Itô sums (see, for example, Bibby and Sørensen, 1995). The discretized score function takes the form

$$S_p(q_0) = \sum_{i=1}^p -a^T(X_{i-1}) [\beta(X_{i-1}) \beta^T(X_{i-1})]^{-1} [X_i - X_{i-1} - \mu(X_{i-1}) \Delta_i], \quad (7)$$

where  $X_0, X_1, \dots, X_p$  are the observations of the process at times  $t_0, t_1, \dots, t_p$  and  $\Delta_i = t_i - t_{i-1}$ .

This discretization well approximates the continuous score function when the intervals between data points are small (Bibby and Sørensen, 1995; Kloeden and Platen, 1992, p. 241).

Noting that each term of the sum in (7) can be written as

$$\frac{1}{\sigma^2(\eta y_{i-1} + c \varepsilon x_{i-1})} \\ \times \left\{ \frac{-\eta[x_i - x_{i-1} - r x_{i-1}(1 - x_{i-1}) \chi(x_{i-1}) \Delta_i + q_0 x_{i-1} y_{i-1} \chi(x_{i-1}) \Delta_i]}{x_{i-1} \chi(x_{i-1})} \right. \\ \left. + \frac{\varepsilon[y_i - y_{i-1} - c q_0 x_{i-1} y_{i-1} \chi(y_{i-1}) \Delta_i + u y_{i-1} \chi(y_{i-1}) \Delta_i]}{y_{i-1} \chi(y_{i-1})} \right\},$$

with simple calculations, equaling  $S_p(q_0)$  to zero, we obtain an approximation of the classical estimator in closed form:

$$\hat{q}_{0,p} = \frac{1}{t_p} \sum_{i=1}^p \frac{1}{\eta y_{i-1} + c \varepsilon x_{i-1}} \\ \times \left[ -\frac{\eta(x_i - x_{i-1})}{x_{i-1} \chi(x_{i-1})} + r \eta (1 - x_{i-1}) \Delta_i + \frac{\varepsilon(y_i - y_{i-1})}{y_{i-1} \chi(y_{i-1})} + \varepsilon u \Delta_i \right]. \quad (8)$$

This estimator is consistent and asymptotically normal under suitable conditions, for example, when the process  $\{X_t\}$  is ergodic (see Prakasa Rao, 1999, and references therein), but ergodicity is very hard to prove. Since often in biological applications only few data

are available, we consider a data augmentation scheme and then follow a Bayesian approach.

In the Bayesian setting, we introduce a prior distribution  $\pi(q_0)$  for the parameter  $q_0$ . Denote by  $\hat{X} = (X_0, X_1, \dots, X_p)$  the vector of observations; then the posterior distribution of  $q_0$  is given by

$$\pi(q_0|\hat{X}) \propto \pi(q_0) \prod_{i=1}^p f(X_i|X_{i-1}, q_0), \quad (9)$$

where

$$\begin{aligned} f(X_i|X_{i-1}, q_0) &\propto \left| [\beta(X_{i-1})\beta^T(X_{i-1})]^{-1} \right|^{\frac{1}{2}} \\ &\times \exp \left\{ -\frac{1}{2} [X_i - X_{i-1} - \mu(X_{i-1}, q_0)\Delta_i]^T \right. \\ &\times [\Delta_i \beta(X_{i-1})\beta^T(X_{i-1})]^{-1} [X_i - X_{i-1} - \mu(X_{i-1}, q_0)\Delta_i] \left. \right\} \end{aligned}$$

is the conditional distribution of  $X_i$  given the observation at previous time and the value of the parameter. It follows that, as a function of  $q_0$ ,  $\prod_{i=1}^p f(X_i|X_{i-1}, q_0)$  is proportional to the normal density  $\mathcal{N}(\mu_p, \sigma_p^2)$  where

$$\mu_p = \hat{q}_{0,p} \quad \text{and} \quad \sigma_p^2 = \frac{\sigma^2}{t_p}. \quad (10)$$

Because of its biological meaning, the parameter  $q_0$  is positive since it measures the efficiency of the predation process; therefore, a normal prior should not be chosen. As mentioned earlier, many statistical works consider normal distributions in similar cases, provided that their mean and variance make negative values very unlikely. The choice of a normal prior  $\mathcal{N}(\mu_I, \sigma_I^2)$  leads to the posterior normal distribution

$$\mathcal{N} \left( \frac{\sigma_I^2 \mu_p + \mu_I \sigma_p^2}{\sigma_I^2 + \sigma_p^2}, \frac{\sigma_p^2 \sigma_I^2}{\sigma_I^2 + \sigma_p^2} \right). \quad (11)$$

Taking, as usual, the posterior mean as Bayesian estimator of  $q_0$ , it can be seen that it is a weighted average of the prior mean and the classical estimator and, in the case of an improper prior ( $\sigma_I \rightarrow \infty$ ), it converges to the classical estimator. Although the choice of a normal prior leads to an estimator of  $q_0$  in closed form and avoids simulations to get posterior density and estimator, we prefer to choose a prior compatible with the physical meaning of the parameter  $q_0$ . We choose two simple, commonly used distributions depending on the available prior information. When information on the parameter  $q_0$  is available, a gamma prior is chosen for its flexibility, whereas an improper prior  $\pi(q_0) \propto I_{(0,\infty)}(q_0)$  is chosen when we have no information about  $q_0$ . In the former case, an MCMC algorithm is used to obtain a sample from the posterior distribution, whereas the latter choice leads to a truncated normal posterior distribution

$$\pi(q_0|\hat{X}) = \frac{1}{\sqrt{2\pi\sigma_p^2[1 - \Phi(-\frac{\mu_p}{\sigma_p})]}} e^{-\frac{1}{2\sigma_p^2}(q_0 - \mu_p)^2} I_{(0,\infty)}(q_0) \quad (12)$$

from which sampling is straightforward. In (12),  $\Phi$  denotes the cumulative distribution function of the standard normal distribution.

#### 4. MCMC estimation

The aim of this section is to obtain a sequence of Monte Carlo samples for  $q_0$  which will be used to approximate the posterior distribution of  $q_0$ .

##### 4.1. An overview of the proposed MCMC method

Since we consider discrete observations and work with the Euler approximation of (5), it is important to employ a sufficiently large number of (real and latent) data to ensure that the discretization bias is arbitrarily small (see Elerian et al., 2001, for a detailed discussion).

In many biological surveys on population dynamics, we have at disposal only few observations and the intervals between observations are too large to guarantee a good approximation of the maximum likelihood. A possible solution to overcome this problem is to generate additional data between two consecutive real observations and treat them as missing data. In such a way, we dispose of a greater number of data to approximate the posterior density of  $q_0$ . The MCMC method used to estimate  $q_0$  is summarized in the following steps:

1. Draw a value for  $q_0$  from the prior distribution  $\pi(q_0)$ ;
2. Generate latent data between each pair of real observations using the posterior density of missing data given the current value of  $q_0$ ;
3. Sample a value for  $q_0$  from its posterior density.

One iteration of the Markov chain is completed once we have updated the whole vector of latent observations and  $q_0$ . Applying recursively steps 2 and 3, after an initial burn-in period, we obtain a sequence of  $q_0$ -values giving an approximation of the posterior density of  $q_0$ .

At each step of the MCMC algorithm, latent data have to be generated. To do this, we follow the method proposed by Eraker (2001) and then applied by Golightly and Wilkinson (2005). They consider a diffusion bridge between data at times  $t_{i-1}$  and  $t_{i+1}$  to generate latent observations at time  $t_i$ . Here, we update observations in blocks; more precisely, we generate simultaneously latent data between two consecutive real observations.

Since the distribution from which latent data have to be drawn is not of a standard form, we will apply a Metropolis–Hastings (M–H) algorithm to sample from this distribution. In the M–H algorithm, a proposal density has to be specified. This density is used to supply a possible value for the next iteration of the chain. The value is accepted or rejected on the basis of the “acceptance probability” based on the proposal density. If the value is accepted, the chain moves to the sampled value, otherwise it stays where it is (Gilks et al., 1996, p. 5). We consider as proposal distribution a truncated normal distribution analogous to that used in Eraker (2001). This approach is also similar to the Brownian bridge sampler proposed by Durham and Gallant (2002), as noted by Eraker (2002). Golightly and Wilkinson (2006) follow the approach by Durham and Gallant (2002) to obtain a proposal distribution for their block Gibbs algorithm. Here we use a different proposal distribution.

If the posterior distribution of  $q_0$  has not a standard form too, we again use a M–H algorithm to sample from it. In Section 4.3, we will discuss this problem.



#### 4.2. Generation of latent data

We generate  $M$  equidistant latent data between each pair of real observations. The problem of parameter estimation will be studied for several values of  $M$ .

Let

$$Y = \begin{pmatrix} x(t_0) & x^*(t_1) & \cdots & x^*(t_M) & x(t_{M+1}) & \cdots & x^*(t_{n-1}) & x(t_n) \\ y(t_0) & y^*(t_1) & \cdots & y^*(t_M) & y(t_{M+1}) & \cdots & y^*(t_{n-1}) & y(t_n) \end{pmatrix}$$

be the matrix of all (real and latent) data. Here  $n = (p - 1)M + p$  is the total number of observations. We shall denote by  $X_i = (x(t_i), y(t_i))$  the real datum at time  $t_i$ , by  $X_i^* = (x^*(t_i), y^*(t_i))$  the latent datum at time  $t_i$  and, generically, by  $Y_i$  a real or a latent datum at time  $t_i$ , depending on the context.

At the first step, we consider a “guess” value for  $q_0$  drawn from the prior distribution  $\pi(q_0)$  and generate latent data by means of a linear interpolation between two consecutive real observations.

At this point, we generate a new value for  $q_0$  as shown in the next subsection. Then, we update latent observations on the whole time interval  $[0, t_{\text{fin}}]$ . As already said in the previous subsection, updating process can be done for each latent datum separately (Eraker, 2001; Golightly and Wilkinson, 2005) or in blocks of latent data (Elerian et al., 2001; Golightly and Wilkinson, 2006). In the following, we will describe both the schemes.

– *Sampling scheme for a single latent observation* We generate the latent datum  $X_i^*$  from the conditional distribution of  $X_i^*$  given  $Y_{i-1}$ ,  $Y_{i+1}$  and  $q_0$ . We know that

$$\begin{aligned} & \pi(X_i^* | Y_{i-1}, Y_{i+1}; q_0) \\ & \propto P(Y_{i+1} | X_i^*, q_0) P(X_i^* | Y_{i-1}; q_0) \\ & \propto \left| [\beta(Y_{i-1}) \beta^T(Y_{i-1})]^{-1} \right|^{\frac{1}{2}} \left| [\beta(X_i^*) \beta^T(X_i^*)]^{-1} \right|^{\frac{1}{2}} \\ & \quad \times \exp \left\{ -\frac{1}{2} [X_i^* - Y_{i-1} - \mu(Y_{i-1}; q_0) \Delta_i]^T \right. \\ & \quad \times (\Delta_i \beta(Y_{i-1}) \beta^T(Y_{i-1}))^{-1} [X_i^* - Y_{i-1} - \mu(Y_{i-1}; q_0) \Delta_i] \left. \right\} \\ & \quad \times \exp \left\{ -\frac{1}{2} [Y_{i+1} - X_i^* - \mu(X_i^*; q_0) \Delta_{i+1}]^T \right. \\ & \quad \times (\Delta_{i+1} \beta(X_i^*) \beta^T(X_i^*))^{-1} [Y_{i+1} - X_i^* - \mu(X_i^*; q_0) \Delta_{i+1}] \left. \right\}, \end{aligned} \quad (13)$$

where  $\Delta_i = t_i - t_{i-1}$ . In general, this is not a density of a standard form, so we apply a M–H algorithm to simulate from  $\pi(X_i^* | Y_{i-1}, Y_{i+1}; q_0)$ . As proposal, we choose a distribution analogous to those proposed by Eraker (2001) and then used by Golightly and Wilkinson (2005); i.e. we choose a truncated bivariate normal distribution

$$g(X_i^*|Y_{i-1}, Y_{i+1}; q_0) \propto \mathcal{N}\left(\frac{1}{2}(Y_{i-1} + Y_{i+1}), \frac{1}{2}\Delta_i \cdot \beta(Y_{i-1})\beta^T(Y_{i-1})\right) I_{(0,\infty) \times (0,\infty)}(X_i^*).$$

Denote by  $X_i^{*(s)}$  the value of  $X_i^*$  at the  $s$ th iteration of the algorithm used to estimate the posterior distribution of  $q_0$  described in the previous subsection. We have to calculate the value of  $X_i^*$  at the  $(s+1)$ th iteration. Let  $w$  be a candidate vector generated from the proposal density  $g(X_i^*|Y_{i-1}, Y_{i+1}; q_0)$ . The acceptance probability is

$$\alpha(X_i^*, w|Y_{i-1}, Y_{i+1}; q_0) = \min\left\{1, \frac{\pi(w|Y_{i-1}, Y_{i+1}; q_0)g(X_i^{*(s)}|Y_{i-1}, Y_{i+1}; q_0)}{\pi(X_i^{*(s)}|Y_{i-1}, Y_{i+1}; q_0)g(w|Y_{i-1}, Y_{i+1}; q_0)}\right\}.$$

We generate a uniform random number  $a$ ; if  $a \leq \alpha(X_i^*, w|Y_{i-1}, Y_{i+1}; q_0)$  then set  $X_i^{*(s+1)} = w$ , otherwise set  $X_i^{*(s+1)} = X_i^{*(s)}$ .

– *Sampling scheme for blocks* Latent data can be updated in blocks of length  $m$ , where  $m$  is either fix or random. In particular, at a generic step of the algorithm for the estimation of the posterior of  $q_0$ , if we have updated the observations up to  $Y_{k-1}$ , we update the  $m$ -block  $X_{(k,m)}^* = (X_k^*, \dots, X_{k+m-1}^*)$ . The conditional density of  $X_{(k,m)}^*$  is given by

$$f(X_{(k,m)}^*|Y_{k-1}, Y_{k+m}; q_0) \propto \prod_{j=k}^{k+m-1} \pi(X_j^*|Y_{j-1}, Y_{k+m}; q_0),$$

where

$$\begin{aligned} & \pi(X_j^*|Y_{j-1}, Y_{k+m}; q_0) \\ & \propto P(Y_{k+m}|X_j^*; q_0) P(X_j^*|Y_{j-1}; q_0) \\ & \propto \left[ \beta(Y_{j-1})\beta^T(Y_{j-1}) \right]^{-1} \left| \beta(X_j^*)\beta^T(X_j^*) \right|^{-\frac{1}{2}} \\ & \quad \times \exp\left\{-\frac{1}{2}[X_j^* - Y_{j-1} - \mu(Y_{j-1}; q_0)\Delta_j]^T \right. \\ & \quad \times (\Delta_j \beta(Y_{j-1})\beta^T(Y_{j-1}))^{-1}[X_j^* - Y_{j-1} - \mu(Y_{j-1}; q_0)\Delta_j] \Big\} \\ & \quad \times \exp\left\{-\frac{1}{2}[Y_{k+m} - X_j^* - \mu(X_j^*; q_0)(t_{k+m} - t_j)]^T \right. \\ & \quad \times ((t_{k+m} - t_j) \beta(X_j^*)\beta^T(X_j^*))^{-1}[Y_{k+m} - X_j^* - \mu(X_j^*; q_0)(t_{k+m} - t_j)] \Big\}. \end{aligned} \quad (14)$$

In general, this is not a density of a standard form, so we apply a Metropolis–Hastings algorithm to simulate from  $f(X_{(k,m)}^*|Y_{k-1}, Y_{k+m}; q_0)$ . Since the  $Y_j$ 's must be non-negative, we choose a truncated bivariate normal distribution as proposal for the M–H algorithm (as

done for the sampling scheme for a single latent datum)

$$g(X_{(k,m)}^* | Y_{k-1}, Y_{k+m}; q_0) \propto \prod_{j=k}^{k+m-1} \mathcal{N}\left(\frac{1}{2}(Y_{j-1} + Y_{k+m}), \frac{1}{2}\Delta_j \cdot \beta(Y_{j-1})\beta^T(Y_{j-1})\right) I_{(0,\infty) \times (0,\infty)}(X_j^*).$$

This proposal distribution is in the spirit of Eraker (2001); an alternative choice could be based on the modified diffusion bridge construction of Durham and Gallant (2002, p. 305).

Denote by  $X_{(k,m)}^{*(s)}$  the value of  $X_{(k,m)}^*$  at the  $s$ th iteration of the M–H algorithm used to generate latent data. We have to calculate the value of  $X_{(k,m)}^*$  at the  $(s+1)$ th iteration. Let  $w$  be a candidate vector generated from the proposal density.

The acceptance probability is

$$\alpha(X_{(k,m)}^*, w | Y_{k-1}, Y_{k+m}; q_0) = \min \left\{ 1, \frac{f(w | Y_{k-1}, Y_{k+m}; q_0) g(X_{(k,m)}^{*(s)} | Y_{k-1}, Y_{k+m}; q_0)}{f(X_{(k,m)}^{*(s)} | Y_{k-1}, Y_{k+m}; q_0) g(w | Y_{k-1}, Y_{k+m}; q_0)} \right\}.$$

We generate a uniform random number  $a$ ; if  $a \leq \alpha(X_{(k,m)}^*, w | Y_{k-1}, Y_{k+m}; q_0)$  then set  $X_{(k,m)}^{*(s+1)} = w$ , otherwise set  $X_{(k,m)}^{*(s+1)} = X_{(k,m)}^{*(s)}$ .

In this paper, we apply the second sampling scheme presented with  $m = M$ .

The algorithm is summarized in the following steps.

1. Initialize  $q_0$  and  $Y$ . Generate  $q_0$  from the initial distribution  $\pi(q_0)$  and the latent data through a linear interpolation of two consecutive observations.
2. Set  $K_1 = 1$ .
3. Update the data between  $X_{K_1+1}^*$  and  $X_{K_1+M+1}^*$ .
4. Set  $K_1 = K_1 + M + 2$ . Repeat step 3 to update all the latent data. When  $K_1 > n$ , then update  $q_0$  from  $\pi(q_0 | Y)$ , store this value (which will be used to construct the posterior distribution of  $q_0$ ), and repeat the steps 2–4.

#### 4.3. Simulation of $q_0$ from the posterior distribution

Denote by  $q_0^{(s)}$  the value of  $q_0$  at the  $s$ th iteration of the algorithm for the estimation of the posterior distribution of  $q_0$ . We consider step  $s+1$ . Once we have simulated the matrix  $Y$  of all the data, the parameter  $q_0^{(s+1)}$  must be generated. If the prior distribution of  $q_0$  is a gamma distribution, then the posterior is not of a standard form, as mentioned in Section 3. In this case, a Metropolis–Hastings step can be used to generate  $q_0^{(s+1)}$  from  $\pi(q_0 | Y) \propto \pi(q_0) \prod_{i=1}^n \pi(Y_i | Y_{i-1}, q_0)$ . We consider, as proposal density  $\bar{g}(q_0 | Y)$  for the M–H algorithm, a gamma distribution with mean and variance equal to those of the true density  $\pi(q_0 | Y)$ , approximated numerically. Let  $q_0^*$  be a value drawn from  $\bar{g}(q_0 | Y)$ . The acceptance probability is

$$\bar{\alpha}(q_0^* | Y) = \min \left\{ 1, \frac{\pi(q_0^* | Y) \bar{g}(q_0 | Y)}{\pi(q_0 | Y) \bar{g}(q_0^* | Y)} \right\}.$$

We generate a uniform random number  $a$ . If  $a \leq \bar{\alpha}(q_0^*|Y)$ , then set  $q_0^{(s+1)} = q_0^*$ ; otherwise set  $q_0^{(s+1)} = q_0$ .

If the prior distribution is  $I_{(0,\infty)}(q_0)$  (improper distribution), then the posterior distribution is a truncated normal distribution (see Section 3). In this case, it suffices to draw at each step a value from this distribution.

## 5. Application

We consider the functional response parameter estimation in one of the most important predator–prey system relevant to biological control problems (Sabelis, 1981; Helle and Sabelis, 1985): the prey mite *Tetranychus urticae* and the predator mite *Phytoseiulus persimilis*. The population dynamics is described by system (5) where the behavioral parameter  $q_0$  is unknown. Two cases are considered. Parameter  $q_0$  will be estimated in case of data generated by model (5) to test the method proposed here. Then, the method will be applied to parameter estimation in case of observational data. Median and mean of the posterior distribution are typical estimates in the Bayesian approach. The choice of the median is usually motivated by the goal of getting a more robust estimator than the posterior mean. In our case, we will consider the posterior median as the estimate of  $q_0$ , although we found out that mean and median are roughly the same. Different numbers of latent observations between two consecutive field data will be considered and results will be compared on the basis of a set of biological indexes that interpret the effect of predator control on prey population dynamics.

### 5.1. Datasets and model parameters

**Datasets.** To estimate the parameter  $q_0$  in the functional response two datasets of predator–prey dynamics have been purposely arranged (Gilioli and Vacante, 2001). The first derives from an extensive biological control survey of the pest mite *Tetranychus urticae* (prey) by the artificially introduced biological control agent *Phytoseiulus persimilis* (predator). A field annual crop of strawberry was settled in Ispica (Ragusa, Italy). Within the field, eight separated experimental plots were identified and for each the local predator–prey dynamics were followed by means of an extensive sampling protocol. To minimize the sampling effort, only adult mites were observed. In this way, sampling procedure does not require leaf collection and laboratory observation of immature stages at the stereomicroscope. Under the allowable assumption of stable stage-structure in both predator and prey populations (Buffoni and Gilioli, 2003; Knapp et al., 2006), adult sampling can lead to an estimate of the total population abundance. Data from the extensive survey have been used to estimate parameters  $\sigma$ ,  $\varepsilon$ ,  $\eta$  and  $q_0$  in system (5) applying the least square method (see below). The second dataset comes from an intensive sampling of twelve separated local predator–prey dynamics in an experimental setup like the one described before. Time evolution of prey and predator abundance were followed during the period in which the system develops on the host plant. The sampling procedure was based on a high effort to obtain high precision in the estimates (a large number of sample units were collected and observed in laboratory using stereomicroscope to count individuals belonging to each biological stage). Data

**Table 1** Estimate of  $q_0$  (median of the posterior distribution) in the case of 10 data simulated from Eq. (5)

Latent data	1	5	10	15	20	30
Estimate	1.6358	1.5828	1.5729	1.5689	1.5678	1.5659

collected in the intensive survey have been used to improve, through the MCMC approach described in Section 4, the estimate of  $q_0$  obtained using data from the extensive survey.

*Biodemographic parameters.* The biodemographic lumped parameters  $r$  and  $c$  in system (5) have been estimated by Buffoni and Gilioli (2003) starting from experimental data obtained at individual level, while  $u$  comes from Nachmann (1996). Therefore, we consider

$$r = 0.11; \quad c = 0.35; \quad u = 0.09.$$

*Other parameters.* To reduce variability among different trends in the eight predator–prey dynamics, at each sampling date, the predator and prey average abundance per habitat unit have been considered. Applying the least square method to the normalized abundance (with respect to carrying capacity) and the simulated curves solution to (5), we obtain the following estimates

$$\hat{\sigma} = 0.321; \quad \hat{\varepsilon} = 0.079; \quad \hat{\eta} = 0.106; \quad \hat{q}_0 = 2.4767.$$

The values of  $\hat{\sigma}$ ,  $\hat{\varepsilon}$ ,  $\hat{\eta}$ , obtained as previously mentioned from the extensive survey, have been plugged into system (5), while  $\hat{q}_0$  has been considered as mean of a prior distribution in the Bayesian approach. In particular, we consider for  $q_0$  a Gamma prior with mean 2.4767 and variance 2.

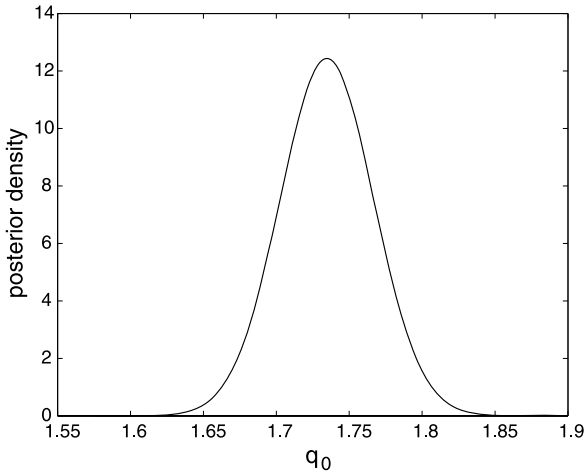
5.2. Simulated data

The initial conditions are given by the values of average abundance of the field case at the first sampling

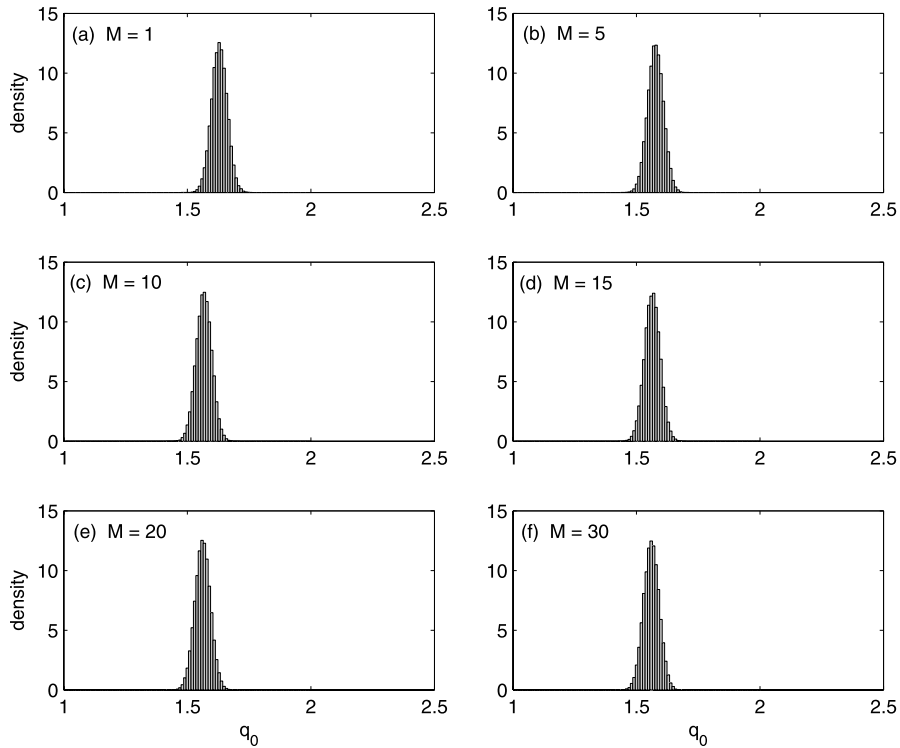
$$x_0 = 0.1; \quad y_0 = 0.007$$

and the final time is  $t_{\text{fin}} = 100$  days. We fix a value  $q_0 = 1.5$  and generate 10 observations, at equidistant time, from system (5).

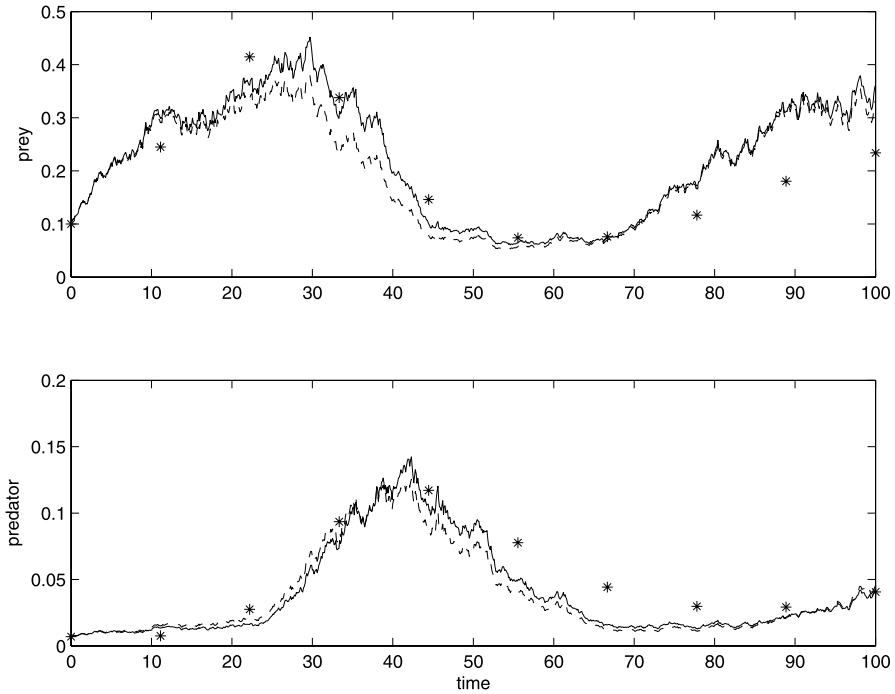
The classical estimate for 10 observations is  $\hat{q}_{0,10} = 1.7357$ . In the Bayesian setting, the posterior distribution obtained for a gamma prior with mean 1.5 and variance 1 (Fig. 1) has median equal to 1.7347 and it does not substantially modify the estimate obtained following the frequentist approach. Application of the MCMC method improves the estimate of  $q_0$ . Looking at Table 1, it seems that the estimate converges to the true value as the number of latent data increases. There is a considerable improvement in the estimate up to a certain value of  $M$ , while for greater number of latent observations the improvement is less pronounced. Figure 2 shows the histograms of the posterior distributions obtained for 1, 5, 10, 15, 20 and 30 latent data simulated between two consecutive observations in the case of a Gamma prior. Histograms are obtained from 100,000 simulations with a burn-in



**Fig. 1** Posterior density with median 1.7347, obtained using a gamma prior.



**Fig. 2** Histograms obtained applying the MCMC algorithm with 100,000 simulations, with burn-in of 10,000 iterations, for 10 observations generated from model (5). Case of gamma prior for  $q_0$ .  $M$  denotes the number of latent data between two consecutive observations.

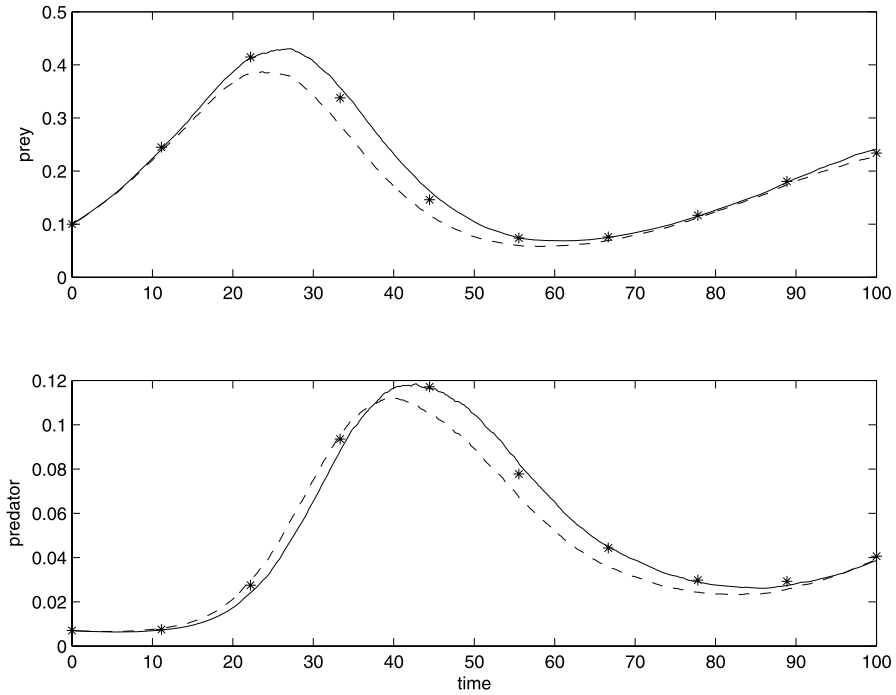


**Fig. 3** One simulated trajectory for prey and predator biomass as function of time, for initial conditions  $x_0 = 0.1$ ,  $y_0 = 0.007$ , obtained with the classical estimate  $q_0 = 1.7357$  (dashed line) and with the Bayesian estimate  $q_0 = 1.5689$  (continuous line). Asterisks denote simulated data.

of 10,000 simulations. Also, these histograms suggest the convergence of the algorithm towards a limit distribution.

The best value for  $M$  should be as large as possible but this implies a huge computational cost. Consequently, we choose a value of  $q_0$  which grants a good balance between computational cost and goodness of fit. In this case, a value of  $M$  equal to 15 seems to be a good choice.

Two simulated trajectories, one for the case  $q_0 = 1.7357$  (classical estimate) and the second for  $q_0 = 1.5689$  (Bayesian estimate with 15 latent data), are compared with the simulated data in Fig. 3, while in Fig. 4, the mean trajectories of prey and predator abundance are shown for classical and Bayesian estimates. In the classical setting, the mean is over 500 simulations, while in the Bayesian case, we consider the mean over 500 simulation and 100 different values of  $q_0$  drawn from its posterior distribution obtained for 15 latent data. The Bayesian estimate gives a better fit of simulated data than the classical estimate. In fact, trajectories obtained in the classical framework underestimate the maximum values of prey and predator and display advanced cycles (Fig. 4). Some well known convergence diagnostics (ergodic mean, autocorrelation function, Gelman–Rubin and Yu–Mykland diagnostics) have been applied to support the goodness of the estimates and the fairly good mixing properties of the Markov chain. For a comparative review on MCMC diagnostics, see Cowles and Carlin (1996).



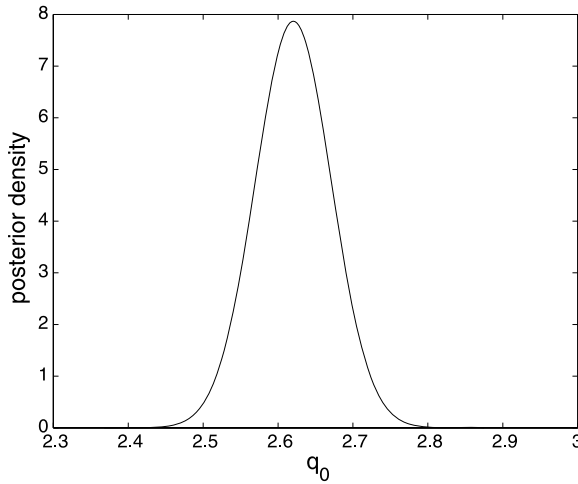
**Fig. 4** Prey and predator biomass as function of time, for initial conditions  $x_0 = 0.1$ ,  $y_0 = 0.007$ . Dashed line: classical estimate  $q_0 = 1.7357$ ; continuous line: Bayesian estimate. Trajectories in the classical framework are obtained considering the mean over 500 simulations, while in the Bayesian setting we consider the mean over 500 simulations and 100 values of  $q_0$  drawn from the posterior distribution obtained for 15 latent data. Asterisks denote simulated data.

In the case of an improper prior, it was found (results not shown) that the posterior distribution of  $q_0$  is similar to the one obtained for a gamma prior, so it will be disregarded in the following.

### 5.3. Field data

In this subsection, we consider data from the intensive survey of the twelve separated local predator–prey dynamics described in Section 5.1. As stated above, to reduce variability among different trends in the twelve predator and prey dynamics, at each sampling date, the predator and prey average abundance per habitat unit have been considered. The estimating procedure has been applied only to the first population cycle of the prey considering only six consecutive observations (from the 2nd to the 7th in the time series). Several reasons justify such a restriction. First, we have to consider periods in which homogeneous biological and ecological processes operate within the system, avoiding phenomena such as rarity conditions of one or both the populations that give rise to problems in the predator functional response and in the growth of the prey (i.e. undercrowding effects), and may require modification of model (5). Secondly, model (5) assumes that





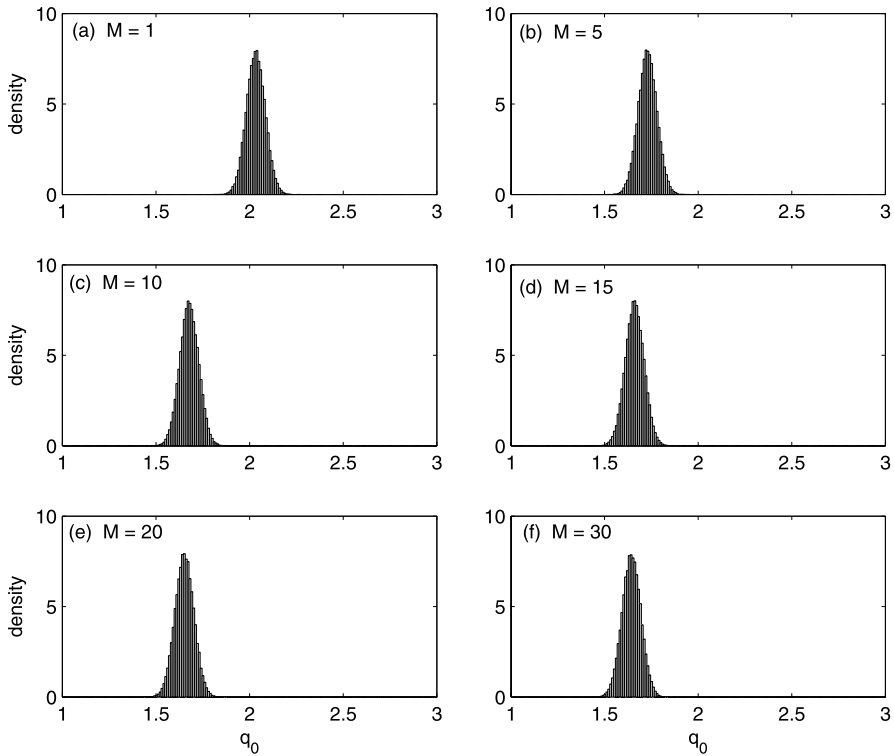
**Fig. 5** Posterior density with median 2.6204, obtained using a gamma prior.

predator–prey systems are in a relatively constant environment both in terms of environmental forcing variables and plant resource condition. For the latter, in particular during the late phase of population dynamics, it can be hypothesized that the plant withering may switch the system to other dynamical patterns at which point model (5) is likely to break down.

The classical estimate is  $\hat{q}_0 = 2.6218$ . The Bayesian estimate obtained from the 6 data using a gamma prior with mean 2.4767 and variance 2, without the introduction of latent data, does not substantially modify the classical estimate (see Fig. 5). Introducing latent data, the obtained posterior distributions (Fig. 6) have medians lower than those with no latent data and seem to converge towards a limit distribution. Histograms in Fig. 6 are obtained from 100,000 simulations with a burn-in of 10,000 simulations.

In contrast to the use of simulated data, in the field case a true value of the parameter is not available. It would be better to choose  $M$  as large as possible but the search for good balance between computational cost and goodness of fit lead us to consider relatively small values of  $M$ . To this end, we compare the mean of 1,000 dynamics of prey and predator with field observations. Comparisons between different estimates are performed accordingly to some indexes which summarize important information on the behavior of population dynamics. Such indexes are defined as follows:

1. The maximum size of the population  $S_{\max}$ ; for the prey it is related to the population impact (damage) on the plant.
2. The time  $t_{\max}$  required to reach the maximum population size; it represents an indicator of the capability of the predator to restrain the prey population exponential growth.
3. The time  $t_{0.5}$  required to halve the maximum size of the population and the time  $t_{0.1}$  required to reduce the population to one tenth of its maximum size; these are indicators of the capability of the predator to cause a rapid decrease of the prey population.
4. The integral  $I$  of the abundance up to  $t_{0.1}$  (relative to field data—Tables 2 and 3); a measure of population pressure on the resource.



**Fig. 6** Histograms obtained applying the MCMC algorithm with 100,000 simulations, with burn-in of 10,000 iterations, for real observations. Gamma prior for  $q_0$ .  $M$  denotes the number of latent data between two consecutive real observations.

These indexes can be calculated for prey and predator both in case of field data and simulated trajectories (Tables 2 and 3). They do not have all the same importance. In fact, in biological control problems, applied entomologists and biological control operators are mainly interested in the value of the maximum size of the prey and in the integral of the prey abundance. To this extent, we analyze the results summarized in Tables 2 and 3.

The best match between estimated values and field data for the prey integral and maximum size are obtained for 8 and 9 latent data. For the predator, the field-case maximum size of the population is never reached, but it is always well approximated; furthermore, the integral of the population is similar in all the Bayesian cases with more than 5 latent data. For the classical estimate, as in the case of a simple posterior distribution with no latent data, there is a good approximation of the time required to reach the maximum value of the prey (about a 1 day delay), but the maximum size and the integral of the prey density underestimate the corresponding field values. The integral of the predator is better approximated than in all the Bayesian cases, but we have a greater difference in the time required to reach the maximum of the predator (about 11 days). Consequently, we can choose a value  $M = 9$  to obtain a satisfactory approximation of both prey and predator behaviors limiting the computational cost. This conclusion is also confirmed by

**Table 2** Indexes for prey for different values of latent data, classical estimate and field observations. 3rd column: maximum prey abundance in the simulated curve. 4th column: time required to reach  $S_{\max}$ . 5th column: half of the maximum prey abundance. 6th column: time required to reach  $S_{0.5}$ . 7th column: tenth part of the maximum prey abundance. 8th column: time required to reach  $S_{0.1}$ . 9th column: integral of the prey curve up to 48.58 days ( $t_{0.1}$  of field case); in the case of field data the integral of the curve is obtained linking field data with segments. 10th column: mean square error between simulated curve and real data

Latent data	Median of posterior distribution	$S_{\max}$	$t_{\max}$ (days)	$S_{0.5}$	$t_{0.5}$ (days)	$S_{0.1}$	$t_{0.1}$ (days)	Integral	Residual
1	2.0386	0.6059	22.93	0.3084	30.61	0.0610	39.17	14.0505	0.3693
2	1.8732	0.6155	23.72	0.3078	31.69	0.0616	41.49	14.8856	0.3428
3	1.7992	0.6184	24.07	0.3092	32.24	0.0618	42.83	15.3190	0.3310
4	1.7588	0.6202	24.07	0.3101	32.57	0.0620	43.64	15.5738	0.3249
5	1.7332	0.6213	24.07	0.3107	32.80	0.0621	44.23	15.7425	0.3213
6	1.7161	0.6221	24.07	0.3110	32.96	0.0622	44.62	15.8585	0.3190
7	1.7038	0.6227	24.40	0.3114	33.06	0.0623	44.88	15.9436	0.3174
8	1.6942	0.6232	24.40	0.3116	33.16	0.0623	45.11	16.0110	0.3163
9	1.6869	0.6236	24.40	0.3118	33.22	0.0624	45.31	16.0628	0.3154
10	1.6809	0.6239	24.40	0.3119	33.29	0.0624	45.44	16.1058	0.3147
15	1.6639	0.6248	24.40	0.3124	33.45	0.0625	45.86	16.2296	0.3129
20	1.6554	0.6252	24.40	0.3126	33.55	0.0625	46.09	16.2926	0.3120
30	1.6467	0.6256	24.40	0.3128	33.61	0.0626	46.32	16.3579	0.3112
Classical estimate	2.6218	0.6007	21.89	0.3004	28.03	0.0601	33.87	12.0858	0.4475
Field data		0.6393	21	0.3179	37.8	0.0639	48.58	16.0472	

the mean square errors between data and estimated curves reported in Tables 2 and 3. In fact, after 8 and 9 latent data, the residual does not decrease significantly as the number of latent data increases. It is worth mentioning the robustness of the procedure since the medians for different numbers of latent data are similar. As for the simulated data, the same convergence diagnostics support the goodness of our results.

The choice of the number of latent data needs some tuning since our experience suggests that no optimal number exists for all possible cases. It would be better to choose  $M$  large, but this requires a large computational effort. Consequently, we performed a sensitivity analysis by changing the number of latent data between 1 and 30, searching for the smallest value of  $M$  which allows to obtain a good approximation of the introduced indexes. Of course, we recommend reporting not only the results for the chosen number of latent data, but also for other values. We remark that the time required to generate the posterior distribution for 9 latent data and 100,000 iterations of the algorithm is about 160 seconds with a FORTRAN program on a AMD Athlon MP 2800+, 2.25 GHz processor.

Figure 7 shows prey and predator behaviors in classical and Bayesian (with 9 latent data) cases. Trajectories are obtained by considering the mean over 500 simulations and, in the Bayesian case, also over 100 different values of  $q_0$  drawn from its posterior distribution for 9 latent observations. In the classical framework (dashed line), the simulated prey nearly reaches the field prey maximum abundance, but the decrease after this maximum is quicker than in the field case (asterisks). The predator maximum abundance is well approximated, but strongly in advance with respect to the field case. For the MCMC estimates (continuous line), the slope of the prey curve is less pronounced than in the

**Table 3** Indexes for predator for different values of latent data, classical estimate and field observations. 3rd column: maximum predator abundance in the simulated curve. 4th column: time required to reach  $S_{\max}$ . 5th column: half of the maximum predator abundance. 6th column: time required to reach  $S_{0.5}$ . 7th column: tenth part of the maximum predator abundance. 8th column: time required to reach  $S_{0.1}$ . 9th column: integral of the predator curve up to 56.84 days ( $t_{0.1}$  of field case); in the case of field data the integral of the curve is obtained linking field data with segments. 10th column: mean square error between simulated curve and real data

Latent data	Median of posterior distribution	$S_{\max}$	$t_{\max}$ (days)	$S_{0.5}$	$t_{0.5}$ (days)	$S_{0.1}$	$t_{0.1}$ (days)	Integral	Residual
1	2.0386	0.1685	34.69	0.0842	47.27	0.0168	70.13	3.4652	0.2031
2	1.8732	0.1673	34.57	0.0836	49.03	0.0167	74.74	3.5383	0.1998
3	1.7992	0.1668	35.42	0.0834	49.98	0.0167	77.84	3.5702	0.1989
4	1.7588	0.1666	35.42	0.0833	50.47	0.0167	81.83	3.5871	0.1985
5	1.7332	0.1665	36.75	0.0832	50.83	0.0166	82.00	3.5974	0.1984
6	1.7161	0.1664	36.85	0.0832	51.06	0.0166	82.00	3.6042	0.1983
7	1.7038	0.1663	37.24	0.0832	51.22	0.0166	82.00	3.6090	0.1983
8	1.6942	0.1663	37.24	0.0831	51.38	0.0166	82.00	3.6127	0.1983
9	1.6869	0.1662	37.24	0.0831	51.48	0.0166	82.00	3.6154	0.1983
10	1.6809	0.1662	37.24	0.0831	51.55	0.0166	82.00	3.6176	0.1983
15	1.6639	0.1661	37.24	0.0830	51.74	0.0166	82.00	3.6239	0.1983
20	1.6554	0.1660	37.24	0.0830	51.87	0.0166	82.00	3.6269	0.1984
30	1.6467	0.1659	36.57	0.0830	51.97	0.0166	82.00	3.6300	0.1984
Classical estimate	2.6218	0.1738	31.36	0.0869	43.02	0.0174	62.72	3.2338	0.2217
Field obs.		0.1763	42	0.0879	50.84	0.0167	56.84	3.0145	

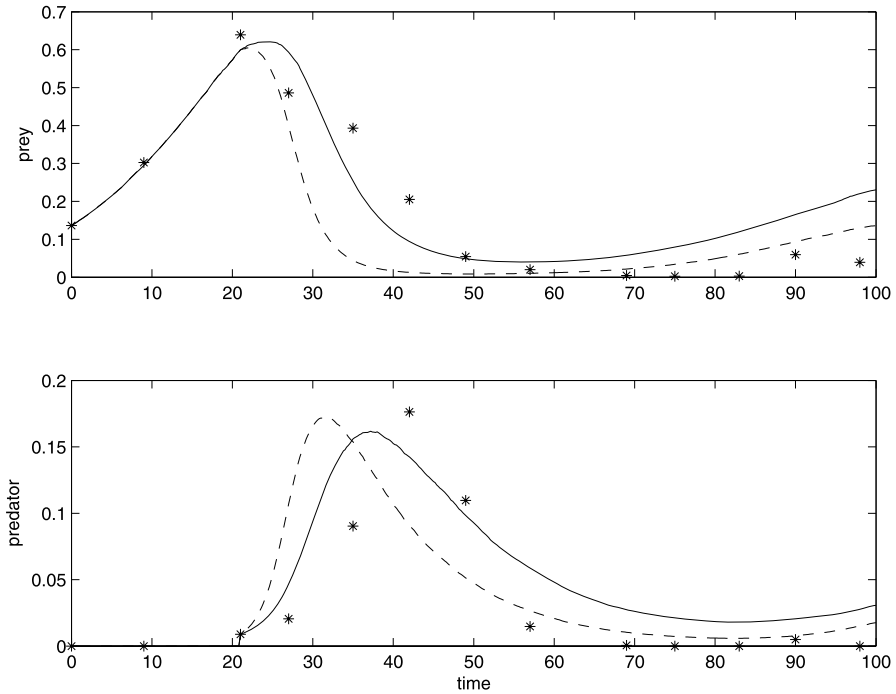
classical case and the time required to reach the maximum predator abundance is closer to the field case than in the classical framework.

The simulated curves show an increase in the last part of the dynamics, that is after 70–80 days when the second population cycle is starting. The increase for the prey is greater than the one shown by field data but, as outlined at the beginning of this subsection, we have limited the study to the first cycle of the prey since when the population reaches low levels, many new factors come into play, and since these are not accounted for in Eq. (5), the model proposed is likely to be unreliable under such circumstances.

Finally, we remark that the predator and prey mean trajectories over 500 realizations for 100 values of  $q_0$  extracted from its posterior distribution are very similar to those obtained using the median of the posterior distribution and considering the mean only over different simulations. This suggests that we cannot find significant changes in the behavior of both prey and predator for small variations of the parameter  $q_0$ . This is a further confirmation that it is not necessary to generate a very large number of latent data because, as pointed out above, the difference between estimates for large numbers of latent data is negligible (see Tables 2 and 3).

6. Concluding remarks

In the paper, we have presented a Bayesian approach to estimate the behavioral parameter of a predator–prey system. The novelty of the approach here proposed mainly relies:



**Fig. 7** Prey and predator biomass as function of time, for initial conditions  $x_0 = 0.1$ ,  $y_0 = 0.007$ . Dashed line: classical estimate; continuous line: Bayesian estimate. Trajectories in the classical framework are obtained considering the mean over 500 simulations, while in the Bayesian setting we consider the mean over 500 simulations and 100 values of  $q_0$  drawn from the posterior distribution obtained for 9 latent data. Asterisks denote field data.

(i) on the formulation of the system (5) in which we consider the unexplored contribution of variability in predator behavioral strategies in functional response modeling, (ii) on the simultaneous update of all  $M$  ( $M > 1$ ) latent data between two real consecutive observations in a multidimensional stochastic system combining the papers by Eraker (2001), Gollightly and Wilkinson (2005, 2006) and Durham and Gallant (2002), and (iii) on its application to field data, introducing some indexes to choose a convenient number of latent data to be generated between two consecutive real observations to reach a good balance between computational cost and goodness of fit.

We have proposed different choices of prior distributions on the parameter of interest, showing the relation between Bayesian and classical estimators when a normal distribution is chosen. The normal prior avoids simulation to get posterior density, but it is not realistic because it allows negative values of the behavioral parameter  $q_0$ . Consequently, we prefer to choose a Gamma or an improper (on the positive real set) prior distribution. We found that these distributions lead to similar results. The choice of the number of latent data is crucial and we suggested both choosing the value which leads to the best predicted fit, and reporting the findings resulting from a range of latent data, as well. Convergence diagnostics have confirmed the goodness of the results. It is worth mentioning

that the same convergence diagnostics have shown a slow convergence and poor mixing properties when using a random block size algorithm.

The MCMC fit obtained, presented in Fig. 7 (continuous line), can be considered satisfactory taking into account the difficulties in collecting data and measurement errors. Moreover, in comparison with the classical estimate (Fig. 7, dashed line) the Bayesian estimate improves the fit in the sense that it reduces residuals and well approximates indexes introduced in Section 5.3 (see Tables 2 and 3).

Future research will be devoted to further improving this fit by considering different forms of functional response in addition to the estimation of the remaining parameters within the Bayesian framework presented here.

## Acknowledgements

The authors wish to thank Antonio Pievatolo for useful discussions and suggestions and two anonymous referees for their valuable comments.

## References

- Akçakaya, H.R., 2000. Viability analyses with habitat-based metapopulation model. *Popul. Ecol.* 42, 45–53.
- Ait-Sahalia, Y., 2002. Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach. *Econometrica* 70(1), 223–262.
- Ait-Sahalia, Y., 2006. Likelihood inference for diffusion: a survey. In: Fan, J., Koul, H.L. (Eds.) *Frontiers in Statistics: in Honor of Peter J. Bickel's 65th Birthday*. Imperial College Press.
- Beskos, A., Papaspiliopoulos, O., Roberts, G.O., Fearnhead, P., 2006. Exact and computationally efficient likelihood-based estimation for discretely observed diffusion processes. *J. Roy. Stat. Soc. Ser. B* 68(3), 333–382.
- Bibby, B.M., Sørensen, M., 1995. Martingale estimation functions for discretely observed diffusion processes. *Bernoulli* 1, 17–39.
- Bonsall, M.B., Hastings, A., 2004. Demographic and environmental stochasticity in predator–prey metapopulation dynamics. *J. Animal Ecol.* 73, 1043–1055.
- Buffoni, G., Gilioli, G., 2003. A lumped parameter model for acarine predator–prey population interactions. *Ecol. Modell.* 170, 155–171.
- Carlin, B.P., Louis, T.A., 2000. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman & Hall, London.
- Carpenter, S.R., Cottingham, K.L., Stow, C.A., 1994. Fitting predator–prey models to time series with observation errors. *Ecology* 75(5), 1254–1264.
- Casas, J., Swarbrick, S., Murdoch, W.W., 2004. Parasitoid behavior: predicting flies from laboratory. *Ecol. Entomol.* 29, 657–665.
- Chesson, P., 1978. Predator–prey theory and variability. *Ann. Rev. Ecol. Syst.* 9, 323–347.
- Cowles, M.K., Carlin, B.P., 1996. Markov Chain Monte Carlo convergence diagnostics: a comparative review. *J. Am. Stat. Assoc.* 91(434), 883–904.
- Durham, G.B., Gallant, A.R., 2002. Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes. *J. Bus. Econ. Stat.* 20(3), 297–316.
- Elerian, O., Chib, S., Shephard, N., 2001. Likelihood inference for discretely observed nonlinear diffusions. *Econometrica* 69(4), 959–993.
- Eraker, B., 2001. MCMC analysis of diffusion models with application to finance. *J. Bus. Econ. Stat.* 19(2), 177–191.
- Eraker, B., 2002. Comment to ‘Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes’. *J. Bus. Econ. Stat.* 20(3), 327–329.

- Gilioli, G., Vacante, V., 2001. Aspetti della dinamica di popolazione del sistema *Tetranychus urticae*—*Phytoseiulus persimilis* in pieno campo: implicazioni per le strategie di lotta biologica. In: Atti del Convegno “La difesa delle colture in agricoltura biologica” Grugliasco-Torino, 5–6 Settembre 2001, Notiziario sulla protezione delle piante, 13 (nuova serie), pp. 95–99.
- Gilioli, G., Baumgärtner, J., Vacante, V., 2005. Temperature influences on the functional response of *Coenosia attenuata* (Diptera Muscidae) individuals. *J. Econ. Entomol.* 98(5), 1524–1530.
- Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in Practice. Chapman & Hall, London.
- Golightly, A., Wilkinson, D.J., 2005. Bayesian inference for stochastic kinetic models using a diffusion approximations. *Biometrics* 61(3), 781–788.
- Golightly, A., Wilkinson, D.J., 2006. Bayesian sequential inference for nonlinear multivariate diffusions. *Stat. Comput.* 16(4), 323–338.
- Gutierrez, A.P., 1996. Applied Population Ecology. A Supply-Demand Approach. Wiley, New York.
- Helle, W., Sabelis, M.W., 1985. Spider Mites. Their Biology, Natural Enemies and Control. Elsevier, Amsterdam.
- Jost, C., Ellner, S.P., 2000. Testing for predator dependence in predator–prey dynamics: a non-parametric approach. *Proc. Roy. Soc. Lond. B* 267, 1611–1620.
- Kareiva, P., 1982. Experimental and mathematical analysis of herbivore movement: quantifying the influence of plant spacing and quality on foraging discrimination. *Ecol. Monogr.* 52(3), 261–282.
- Kareiva, P., 1990. Population dynamics in spatially complex environments: theory and data. *Philos. Trans. Roy. Soc. Lond.* 330, 175–190.
- Kessler, M., Parades, S., 2002. Computational aspects related to martingale estimating functions for a discretely observed diffusion. *Scand. J. Stat.* 29, 425–440.
- Kessler, M., Sørensen, M., 1999. Estimating equations based on eigenfunctions for a discretely observed diffusion process. *Bernoulli* 5(2), 299–314.
- Kloeden, P.E., Platen, E., 1992. Numerical Solution of Stochastic Differential Equations. Springer, Berlin.
- Knapp, M., Sarr, I., Baumgärtner, J., Gilioli, G., 2006. Temporal dynamics of *tetranychus urticae* populations in small-scale Kenyan farmers tomato fields. *Exp. Appl. Acarol.* 39, 195–212.
- Liptser, R.S., Shirayev, A.N., 1977. Statistics of Random Processes I—General Theory. Springer, New York.
- McCallum, H., 2000. Population Parameters. Estimation for Ecological Models. Blackwell, Oxford.
- Nachmann, G., 1996. Within- and between-system variability in an acarine predator–prey metapopulation. In: Di Cola, G., Gilioli, G. (Eds.) Computer Science and Mathematical Methods in Plant Protection. Quaderni del Dipartimento di Matematica, Università di Parma, n. 135, pp. 110–132.
- Øksendal, B., 1998. Stochastic Differential Equations: An Introduction with Applications, 5th edn. Springer, Berlin.
- Pascual, M.A., Kareiva, K., 1996. Predicting the outcome of competition using experimental data: maximum likelihood and Bayesian approaches. *Ecology* 77(2), 337–349.
- Pedersen, A.R., 1995a. A new approach to maximum likelihood estimation for stochastic differential equations based on discrete observations. *Scand. J. Stat.* 22, 55–71.
- Pedersen, A.R., 1995b. Consistency and asymptotic normality of an approximate maximum likelihood estimator for discretely observed diffusion processes. *Bernoulli* 1, 257–279.
- Prakasa Rao, B.L.S., 1999. Statistical Inference for Diffusion Type Processes. Arnold, London.
- Rand, D., Wilson, H.B., 1991. Chaotic stochasticity: a ubiquitous source of unpredictability in epidemics. *Proc. Roy. Soc. Lond. B* 246, 179–184.
- Regan, H.M., Colyvan, M., Burgman, M.A., 2002. A taxonomy and treatment of uncertainty for ecology and conservation biology. *Ecol. Appl.* 12(2), 618–628.
- Royama, T., 1971. A comparative study of models for predation and parasitism. *Res. Popul. Ecol. Suppl.* 1, 1–91.
- Sabelis, M.W., 1981. Biological Control of Two Spotted Spider Mites Using Phytoseiid Predators. Part I. Modelling the Predator-Prey Interaction at the Individual Level. Pudoc, Wageningen.
- Shaffer, G., 1981. Minimum population size for species conservation. *Bioscience* 31, 131–134.
- Shaffer, G., 1987. Minimum viable populations: coping with uncertainty. In: Soulé, M.E. (Ed.), Viable Populations for Conservation. Cambridge University Press, Cambridge.
- Sørensen, M., 1999. On asymptotics of estimating functions. *Braz. J. Probab. Stat.* 13, 111–136.
- Sørensen, H., 2004. Parametric inference for diffusion processes observed at discrete points in time: a survey. *Int. Stat. Rev.* 72(3), 337–354.
- Stramer, O., Yan, J., 2007. On simulated likelihood of discretely observed diffusion processes and comparison to closed-form approximation. *J. Comput. Graph. Stat.*, to appear.

- Tanner, M.A., Wong, W.H., 1987. The calculation of posterior distributions by data augmentation. *J. Am. Stat. Assoc.* 82, 528–541.
- Turchin, P., 2003. *Complex Population Dynamics. A Theoretical/Empirical Synthesis*. Princeton University Press, Princeton.
- Xia, J.Y., Rabbinge, R., van der Werf, W., 2003. Multistage functional responses in a ladybeetle-aphid system: scaling up from the laboratory to the field. *Environ. Entomol.* 32(1), 151–162.