Progress report – Fitting stochastic predator-prey models with and without functional response data

Frederic Barraquand, Olivier Gimenez

January 19, 2018

Abstract

Asking whether parameters can be identified from a given dataset is an important step in the model development process. Most predator-prey model fitting has attempted inverse modelling, identifying parameters solely from time series. Here, we aim at identifying the potential benefits of combining data when both growth and attack processes are viewed as stochastic. We fit a stochastic predator-prey model of the Leslie type to time series and functional response data simulated from the model. Our model has both environmental stochasticity in the growth rates and interaction stochasticity, i.e. a stochastic functional response. We examine what the FR data brings to the quality of the estimates, and whether estimation is possible (for various time series length) solely with time series data. Both bayesian and frequentist estimation are performed, and in both cases we report diagnostics of identifiability of the various parameters. [stuff on results here]. Our framework to combine data sets is general may be extended to other interaction scenario when both data on interaction rates and population counts are available.

Introduction

tentative list

- Many attempts to fit predator-prey models to data assume observation stochasticity and do inverse modelling. Inverse modelling from time series can be hard due to identifiability issues, is therefore fraud with some uncertainties. Even in simple, phenomenological statistical models for two species, there can remain considerable uncertainty about the model parameters.

- One way to decrease uncertainty is to use combination of data sets to decrease that uncertainty on the parameters.

- This may be best done in a fully stochastic context: there is widespread environmental stochasticity (which may dominate). REFs stochastic predator-prey models.

- Stochasticity also affects the functional response for various reasons. Point clouds are anything but just a type II curve. Considering the FR as a stochastic object is both biologically realistic and statistically convienient.

- Here we consider a ground truth model with both stochasticity in growth rates and interactions parameters. The model is fitted time series and functional response data in both frequentist and bayesian settings.

- We consider both long time series, close to perfect information (T=1000) and time series of realistic ecological length (T= 100, T = 50 or 25). [the latter two remain to be done]. Yet to do: we vary the percentage of data points attributed to the time series vs the functional response.

Models and statistical methods

Predator-prey model in discrete time

We chose a model with a numerical response of the Leslie type (though similar analyses can be done for Lotka-Volterra or Rosenzweig MacArthur models can be done as well, see Supplement S1). A Beverton-Holt type density-dependence for the prey was chosen to avoid cycles in the prey in absence of the predator, so that the model behaviour is more reminiscent of its continuous-time counterpart (see Frank et al's article for more on connecting continous to discrete time models). Our model writes

$$N_{t+1} \quad = \quad N_t \frac{e^{r+\epsilon_{1t}}}{1+\gamma N_t} \exp\left(-G(N_t)\frac{P_t}{N_t}\right), \, \epsilon_{1t} \sim \mathcal{N}(0, \sigma_1^2) \tag{1}$$

$$P_{t+1} \quad = \quad P_t \frac{e^{s+\epsilon_{2t}}}{1+qP_t/N_t}, \, \epsilon_{1t} \sim \mathcal{N}(0, \sigma_1^2) \tag{2}$$

The roots of this model can be traced back to Leslie (1948), Leslie and Gower (1960). Parameter values were loosely inspired by rodents (e.g. Turchin & Hanski 1997, also using a Leslie-type numerical response).

⁴⁵ The division by N_t in $exp(-G(N_t)P_t/N_t)$ expresses the fact that all terms within the exponential are on the

⁴⁶ prey fitness (per capita mortality) scale. This is similar to assuming a Nicholson-Bailey type predation term

⁴⁷ (Weide et al.).

⁴⁸ Until now, we have not specified a model for the functional response $G(N_t)$. With a deterministic functional

⁴⁹ response (FR), we have a classic stochastic predator-prey model with log-normal environmental noise but an

⁵⁰ otherwise 'deterministic skeleton'. However, here we consider a data-generating process where the functional

⁵¹ response is not deterministic but itself stochastic.

$$G_t|N_t \sim \mathcal{N}(\mu_{3t}, \sigma_3^2), \; \mu_{3t} = \frac{CN_t}{D + N_t} \tag{3}$$

⁵² This corresponds to a case where there is mild Gaussian fluctuations around the functional response. Because

⁵³ there can be substantial noise on the FR (see e.g. plots in [REFs there]), we also consider cases where the

⁵⁴ parameters C and D are themselves allowed to vary in time in Appendix S2 [I was thinking to only do this in

⁵⁵ a bayesian setting for the estimation part].

General statistical methodology

⁵⁷ Although we apply here data integration to a predator-prey case, the methodology is general and can in

⁵⁸ principle be applied to any auxiliary information (interaction rate, demographic rate) which is available over

⁵⁹ time and not simply produced by the count data.

⁶⁰ Let us illustrate with the predator-prey case where log-densities for both predator and prey are gathered in a

⁶¹ vector $\mathbf{x}_t = (\ln(N_t), \ln(P_t))^T$. Auxiliary information on the functional response, or rather the observed kill

⁶² rate per predator is denoted G_t. To this can be added other demographic rates R_{t} that are stacked in a

⁶³ vector as well, $\mathbf{a}_t = (G_t, R_t)$. Currently we use $\mathbf{a}_t = G_t$ but it may useful to add more information in other

⁶⁴ applications; hence the derivation is kept general.

⁶⁵ We consider a discrete-time dynamical system (nonlinear difference equation). The population dynamics part

⁶⁶ of the model gives us the probability law of $\mathbf{x}_{t+1}|(\mathbf{x}_t, \mathbf{a}_t)$. We also know the law of $\mathbf{a}_t|\mathbf{x}_t$ (in our case, the

⁶⁷ functional response). We can therefore write down easily the joint likelihood for both data sources in quite

⁶⁸ general terms, denoting $\mathbf{X} = (\mathbf{x}_1, ..., \mathbf{x}_{t_m})$ and $\mathbf{A} = (\mathbf{a}_1, ..., \mathbf{a}_{t_m})$:

$$\mathcal{L}(\mathbf{X}, \mathbf{A}) = p(\mathbf{x}_1, \mathbf{a}_1) \prod_{t=1}^{t_m-1} p_C(\mathbf{x}_{t+1}, \mathbf{a}_{t+1}|\mathbf{x}_t, \mathbf{a}_t) \tag{4}$$

where $p(y)$ and $p_C(y)$ are continuous probability densities for the vector $(\mathbf{x}, \mathbf{a})$ and its conditional pdf, respectively. The conditional pdf can be further decomposed using the chain rule

$$p(\mathbf{x}_{t+1}, \mathbf{a}_{t+1} | \mathbf{x}_t, \mathbf{a}_t) = p_2(\mathbf{a}_{t+1} | \mathbf{x}_{t+1}, \mathbf{x}_t, \mathbf{a}_t) \times p_1(\mathbf{x}_{t+1} | \mathbf{a}_t, \mathbf{x}_t) = p_2(\mathbf{a}_{t+1} | \mathbf{x}_{t+1}) \times p_1(\mathbf{x}_{t+1} | \mathbf{a}_t, \mathbf{x}_t)$$

where $p_1(y)$ is given by the dynamical system and $p_2(y)$ is the functional response model (or a demographic model). We therefore end up with a model

$$\mathcal{L}(\mathbf{X}, \mathbf{A}) = p_1(\mathbf{a}_1 | \mathbf{x}_1) p(\mathbf{x}_1) \prod_{t=1}^{t_m - 1} \underbrace{p_1(\mathbf{x}_{t+1} | \mathbf{a}_t, \mathbf{x}_t)}_{\text{dynamical system}} \times \underbrace{p_2(\mathbf{a}_{t+1} | \mathbf{x}_{t+1})}_{\text{auxiliary information model}} \tag{5}$$

where we swapped p_1 and p_2 to get the dynamical system model first.

Application to the stochastic predator-prey model

In the simplest case highlighted by our discrete-time dynamical systems of the sections above, $p_1(y) = p_{11}(x_1) p_{12}(x_2)$ is the product of the two gaussian pdf for log-densities conditional on past densities. Using the equations 1-2, we have

$$n_{t+1} | \mathbf{x}_t = \ln(N_{t+1}) | \mathbf{x}_t \sim \mathcal{N}(\mu_{1t}, \sigma_1^2) \, , \mu_{1t} = n_t + r - G_t \frac{P_t}{N_t} - \ln(1 + \gamma N_t) \tag{6}$$

$$p_{t+1} | \mathbf{x}_t = \ln(P_{t+1}) | \mathbf{x}_t \sim \mathcal{N}(\mu_{2t}, \sigma_2^2) \, , \mu_{2t} = p_t + s - \ln(1 + q P_t / N_t) \tag{7}$$

with a functional response model (p_2) also given by a Gaussian pdf (in the simplest case where we assume near Gaussian noise on the FR, see Discussion)

$$G_t | \mathbf{x_t} \sim \mathcal{N}(\mu_{3t}, \sigma_3^2), \ \mu_{3t} = \frac{C N_t}{D + N_t}$$

Model scenarios

We considered the following parameter sets for the model [illustration needed, 4 panels: (N(t), P(t), F(N), N vs P) for each parameter, illustrate the first one in the MS, second one in Appendix?]

Parameter	Quasi-cycles	Noisy LC
C	2.5	0
D	1	
Q	10	

[83] The rest of the parameters are [fill in there].

[84] The first parameter set corresponds to a forced focus or quasi-cycles, i.e. sustained oscillations that arise the

[85] interaction between noise and dampened oscillations to the fixed point in the deterministic model. We also

[86] consider a noisy limit cycle, i.e., parameters that give rise to a limit cycle without the noise, so that the cycle

[87] is still very regular but perturbed by the noise.

[88] These two sets of parameters are crossed with several modalities of data availability: - we consider a time

[89] series length T=1000, 100 [should I add 50 and 25?] - we consider that we have functional response data over

[90] 100%, 25%, or 0% of the time series data. This is meant to emulate common scenarios in which the kill rates

[91] are not monitored over the whole time series, and quantify the benefits of adding just a little FR data.

[92] Note that in the case without FR data, we fit the model without noise in the functional response [this is

[93] something we need to discuss in depth Olivier], which is what is usually done in this case (Ellner & Turchin

[94] 2000, newer stuff).

[95] For each parameter x data availability scenario, we fit the models in both bayesian and frequentist settings.

Model fitting implementation

[97] We fitted the model by MCMC in JAGS and also derived mathematically its likelihood, which we then

[98] optimised using the BFGS algorithm [see comments below on Nelder-Mead as well] with optim() in R. Several

[99] starting values were considered.

[100] Identifiability was inspected in different ways. In a frequentist setting, we computed the Hessian matrix

[101] $H(\theta, Y)$ where $Y = (X, A)$ following eq. XX, at the estimated parameter set value $\hat{\theta}$. Non-zero eigenvalues of

[102] the Hessian - the observed Fisher Information Matrix (FIM) - typically translate into identifiable models

[103] (REFs). We also compute the expected FIM at the true parameter value. This value is relevant theoretically

[104] because $\theta \sim \mathcal{N}(\theta_{\text{true}}, \mathcal{I}(\theta_{\text{true}})^{-1})$ where $\mathcal{I}(\theta_{\text{true}})$ is the FIM. In the case where the time series length $T \to \infty$,

[105] $H(\theta, Y^T) \to \mathcal{I}(\theta_{\text{true}})$ [should I try to prove this properly btw??].

We also computed the expected pairwise correlation between the parameters: the variance-covariance matrix of the parameters $\Sigma = H^{-1}$ so that we get easily the expected pairwise parameter correlation matrix (ρ_{ij}) with each element defined as $\frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}\Sigma_{jj}}}$.

In a bayesian setting, we inspected the correlations in the Markov chains for pairs of parameters, which translates into pair posterior distributions of parameters. Parameters whose chains were too positively or negatively correlated were considered not identifiable separately.

Results

[so far this is very preliminary and I do not consider all the abovementioned modalities of the analysis. We focus for now on T=1000 and the comparison with/without FR data. In both bayesian and frequentist setting]

Bayesian analysis

Frequentist analysis

Maximum likelihood and Hessian computation

Likelihood profiles

Bayesian analysis II - reparameterization of the model

Model wFR is parameterized as often done in the literature (e.g. Weide et al.) but not exactly in terms of carrying capacities for the prey. Here we attempt a reparameterization to decrease the correlation between pairs of parameters belonging to the same function in the model.

Discussion

Identifiability

It seems that the model is globally identifiable when FR data is present. However, for the first parameterization, pairs of parameters belonging to the same functional forms of the models are highly correlated. These pairs are, respectively, (r K), (s, Q), and (C,D). Plots of the functions realized for each parameter pair

129 Reparameterizing so that K and Q are more closely related to carrying capacities improved substantially

130 pairwise identifiability

131 The absence of FR data substantially compromises identifiability. This is all the more important that we

132 consider here a fairly small noise on the functional response and relatively rich datasets. [more here based on

133 what we find]

Maximum likelihood algorithms

135 [BFGS vs Nelder-Mead]

136 [Try likelihood profiles for the case without FR data?]

Stochastic or deterministic FR in absence of FR data?

138 [try model with noise on the FR even without FR data?]

Future avenues

140 • Inserting demographic information together with functional response data.

141 • [do I mention more complex FR here or in a Supplement + before]

142 • Other kinds of interactions models. Perhaps cases with more interaction data and less time series

143   (mutualisms?).