bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fitting deterministic population models to stochastic data: from trajectory matching to state-space models

Benjamin Rosenbaum[1,2] and Emanuel A. Fronhofer[3]

1. German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, Germany

2. Institute of Biodiversity, Friedrich Schiller University Jena, Jena, Germany

3. ISEM, Université de Montpellier, CNRS, IRD, EPHE, Montpellier, France

Corresponsence: benjamin.rosenbaum@idiv.de, emanuel.fronhofer@umontpellier.fr

Running title: Fitting population dynamics

Keywords: logistic population growth, Lotka-Volterra, predator-prey dynamics, Bayesian inference, time series, microcosms

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Abstract

Population and community ecology traditionally has a very strong theoretical foundation with well-known models, such as the logistic and its many variations, and many modification of the classical Lotka-Volterra predator-prey and interspecific competition models. More and more, these classical models are confronted to data via fitting to empirical time-series, from the field or from the laboratory, for purposes of projections or for estimating model parameters of interest. However, the interface between mathematical population or community models and data, provided by a statistical model, is far from trivial. In order to help empiricists make informed decisions, we here ask which error structure one should use when fitting classical deterministic ODE models to empirical data, from single species to community dynamics and trophic interactions. We use both realistically simulated data and empirical data from microcosms to answer this question in a Bayesian framework. We find that pure observation error models mostly perform adequately overall. However, state-space models clearly outperform simpler approaches when observation errors are sufficiently large or biological models sufficiently complex. Finally, we provide a comprehensive tutorial for fitting these models in R.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Introduction

Studying biotic interactions, measuring their strength, correlations with other traits and understanding consequences of these interactions for ecological systems is at the heart of scientific ecology. Biotic interactions are at the centre of classical questions in population ecology, such as density regulation (e.g., Sibly et al., 2005), but also in community ecology, including modern coexistence theory (Chesson, 2000; Godwin et al., 2020) and beyond, such as host-parasite (epidemiological parameters) and trophic interactions. Importantly, interaction strengths also provide a bridge between ecology and evolution as biotic interactions directly or indirectly determine fitness. Biotic interactions are therefore central to eco-evolutionary dynamics and feedbacks (Yoshida et al., 2003; Hiltunen et al., 2014).

As a consequence, correct estimations of biotic interaction strengths are of great importance, be it using times series from the field (e.g., Sibly et al., 2005) or from laboratory systems (Rosenbaum et al., 2019). The strength of ecology in this context is its important and solid body of theory and the availability of mechanistic models. As a consequence, researchers would like to fit these models to time series data to extract the relevant parameters (e.g., Godwin et al., 2020). This exercise is not always straightforward as models often consist of (coupled) differential equations, such as classical models including logistic or other limited local population growth, Lotka-Volterra-type models for interspecific competition and consumer-resource dynamics or SI-type epidemiological models.

Fitting such models to data is possible via multiple approaches, ranging from "naive" trajectory matching (e.g. Fronhofer and Altermatt, 2015) using nonlinear least-squares or Bayesian approaches (e.g. Nørgaard et al., 2021) to state-space models that allow to explicitly take into account observation and process errors (for a recent overview, see Auger-Méthé et al., 2021). Ecologists and evolutionary biologists faced with these choices may often wonder what the pros and cons or these different approaches are, especially given the varying degrees of complexity of these approaches and technical skills they may require. These decisions are not made easier as the questions include multiple dimensions of complexity, such as frequentist versus Bayesian approaches, discrete-time versus continuous-time models, stochastic versus deterministic models and, finally, multiple error structures.

Importantly, these choices can impact scientific results and conclusions. For instance, Sibly et al. (2005) used a rather simple likelihood approach to fit the θ-logistic model to population census data. The authors concluded that a large array of taxa exhibit similarly shaped density-regulation functions which has important applied consequences for conservation and management. However, Clark et al. (2010) could show that these results are likely flawed due to likelihood ridges (parameter combinations of similar likelihood), a problem which could have been mitigated, for instance, using a Bayesian approach with informed priors.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

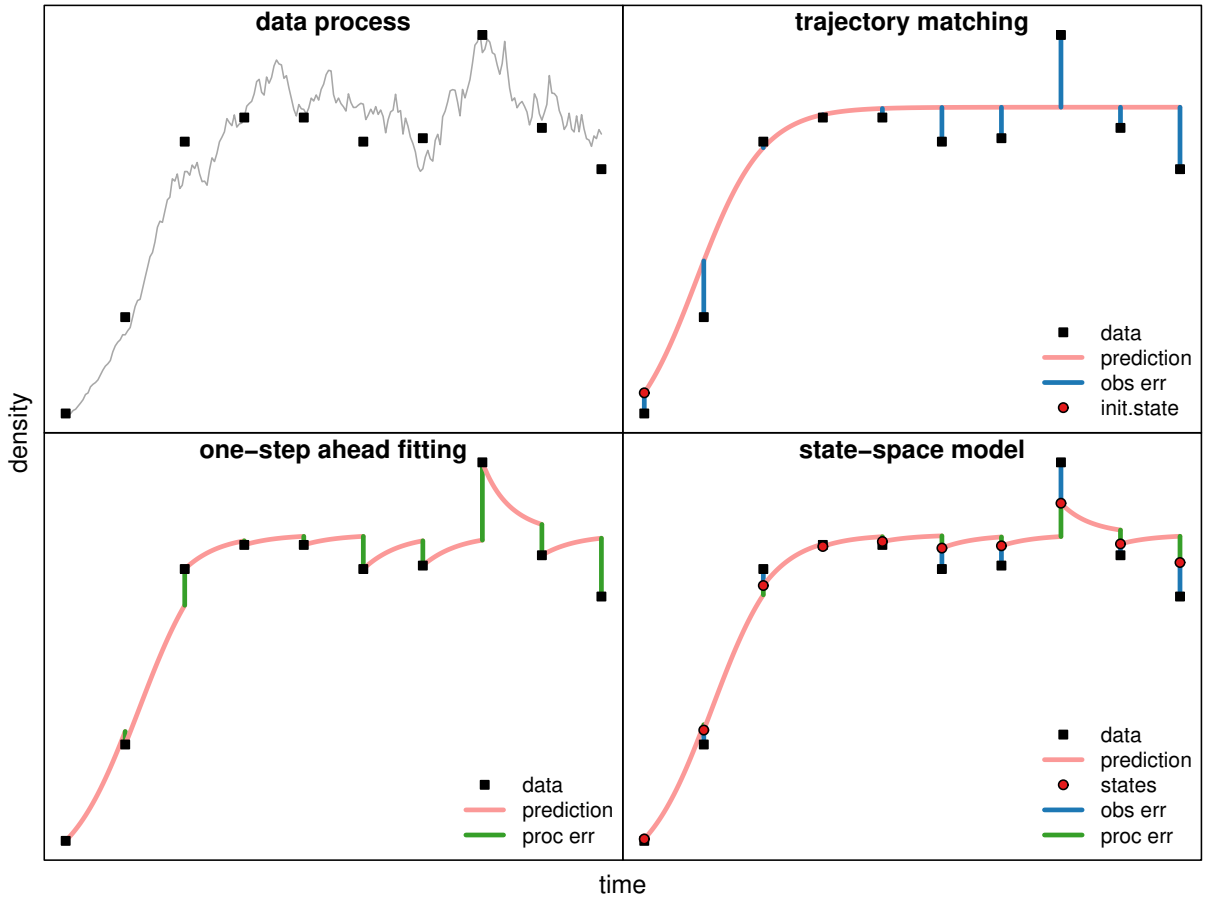

Figure 1: Possible statistical approaches for fitting a population model to a time series of population densities. The upper left panel depicts the underlying stochastic birth-death process (logistic growth) and the black squares show the sampled data with corresponding errors. The simplest statistical model assumes only observation errors and matches the trajectory of the ODE directly to the data (top right). Assuming process error ("one step ahead fitting"; bottom left) takes into account the non-independence of the subsequent data points in the time-series. Finally, the state-space model (SSM; bottom right) takes both sources of error (observation and process) into account.

Here, we will explore with a Bayesian approach the use of statistical models accounting for observation or process errors only, or full state-space models including both sources of errors to fit dynamical population and community models to time-series data (Fig. 1).

Including only observation error reduces model fitting to a nonlinear regression problem, matching the whole trajectory to the time series. Process error only models, however, are nonlinear autoregressive models, where each observation in time is predicted from the previous one. State-space models cannot be classified as standard regression problems, since they require the simultaneous estimation of latent state variables (true population abundances). For benchmarking we will rely on realistically simulated data sets in which we can vary process and observation errors independently, but we will also apply our models to empirical time-series from microbial microcosms. We will consider both, single and multispecies ODE models, with and without non-linearities in their density-regulation functions (Allee effects, Beverton-Holt

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

model). Multispecies models include horizontal communities as well as trophic interactions.

Our analysis builds on previous work, for example by Clark and Bjørnstad (2004), who focus on discrete-time models, including exponential growth and the Ricker model. While these authors also use data to compare state-space approaches to simpler statistical models (e.g., pure observation error models), we explore biological models in continuous-time and more systematically using simulated data as well as more biological scenarios. While our approach relies on Bayesian inference for reasons discussed above, de Valpine and Hastings (2002) compare state-space models to observation and process error models using a likelihood framework. They show that state-space models overall have lower bias and variance than the other methods when fitting Ricker and Beverton-Holt density-regulation models.

We here only take into account observation and process errors and not other sources of stochasticity, such as environmental or demographic stochasticity (see Shoemaker et al., 2020). We don't account for detection error, but see Hefley et al. (2013) for a recent treatment of this question. Finally, we don't include model errors (Xu et al., 2019).

Our work addresses specifically the questions: Which error structure should one use when fitting classical deterministic ODE models to empirical data, from single species to community dynamics and trophic interactions. Does the presence of observation and process error in observed time series influence the identifiability of these models? Using a Bayesian approach, we find that pure observation error models do not perform too badly overall. However, state-space models clearly outperform simpler approaches when observation errors are sufficiently large or biological models sufficiently complex. Additionally, we provide a comprehensive tutorial for fitting these models in R.

Material and methods

In order to understand under which conditions observation errors, process errors or both have to be modelled when analysing population and community dynamics data we consider ecological scenarios of increasing complexity, ranging from single species dynamics up to predator-prey systems. Briefly, we use stochastic individual-based models to generate observations with known underlying processes and sampling regimes following a "virtual ecologist" approach (Zurell et al., 2010). This allows us to include ecologically sound levels of demographic stochasticity, for example. We subsequently fit the appropriate dynamical equations in R with Markov Chain Monte Carlo (MCMC) sampling using Stan (Stan Development Team, 2018). Finally, we complement our work with an analysis of real population dynamic data from microbial laboratory systems using published data (Fronhofer et al., 2020).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Mathematical models

We assume that the following ecological models generate the observed data. Example time series are depicted in Fig. S1.

Single species dynamics: Logistic growth model

As the simplest single species density regulation model we take the Verhulst (1838) model, that is, the $r - \alpha$ formulation of the logistic growth equation (for a detailed discussion of the advantages of this formulation see Mallet, 2012)

$$\frac{dN}{dt} = (r_0 - \alpha N)N \tag{1}$$

with N as the size of the focal population, r_0 as the intrinsic rate of increase and α as the intraspecific competition coefficient. The equilibrium population size can be calculated as $K = \frac{r_0}{\alpha}$.

We extend this model to include an Allee effect by adding a density-dependent mortality term as described in Thieme (2003) which can be derived mechanistically, for example, for mate-finding Allee effects or satiating generalist predators:

$$\frac{dN}{dt} = \left(r_0 - \alpha N - \frac{\eta}{1 + \gamma N}\right) N \tag{2}$$

with η as the amount by which growth is reduced at $N = 0$ and γ controlling the consequences of the Allee effect for higher densities. Calculations the two equilibrium population densities A and K ($\frac{dN}{dt} < 0$ for $N < A$ and for $N > K$) are given in the Supplementary Information.

Single species dynamics: Beverton-Holt model

As a more mechanistic single species density-regulation function (Thieme, 2003; Fronhofer et al., 2020) we explore the continuous-time Beverton-Holt model which follows

$$\frac{dN}{dt} = \left(\frac{b}{1 + \beta N} - d\right) N \tag{3}$$

with b as the birth rate and d as the mortality rate. The intrinsic rate of increase can be calculated as $r_0 = b - d$. The equilibrium population size is $K = \frac{r_0}{\beta d}$.

In analogy to Eq. 2 we can expand the Beverton-Holt model to include an Allee effect by adding a mortality terms (Thieme, 2003), which yields

$$\frac{dN}{dt} = \left(\frac{b}{1 + \beta N} - d - \frac{\eta}{1 + \gamma N}\right) N \tag{4}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Again we refer to the Supplementary Information for the calculation of the equilibrium population densities A and K.

Interspecific competition: n-species Lotka-Volterra competition model

To capture interspecific competition and the dynamics of a horizontal community, we will first use an expansion of the logistic model (Eq. 1):

$$\frac{dN_i}{dt} = (r_{0,i} - \sum_{j=1}^{n} \alpha_{i,j} N_j) N_i \tag{5}$$

where $\alpha_{i,j}$ represent the inter- and intraspecific competition coefficients and form the community matrix. We investigate two separate two-species scenarios, which are different in their interspecific competition coefficients. In the first scenario, the system reaches a stable equilibrium (coexistence), while in the second one species is outcompeted by the other and goes extinct (competitive exclusion).

Predator-prey interactions

For predator-prey interactions, we use the following general model

$$\frac{dR}{dt} = g(R) - f(R)N \tag{6a}$$

$$\frac{dN}{dt} = ef(R)N - dN. \tag{6b}$$

where $g(R)$ is the growth of the resources which can follow any of the above introduced single species population growth functions (Eqs. 1 – 4). For simplicity we will assume that $g(R)$ follows Eq. 1. $f(R)$ is the consumer's functional response which can be linear, saturating or sigmoid. e captures the consumer's assimilation efficiency. In our analyses we will focus on a Holling type II (Holling, 1959), that is, saturating functional response of the form $\frac{a}{1+ahR}$ with a as the predator's search efficiency and h as the handling time. This combination of growth and functional response is also known as the Rosenzweig-MacArthur model and we investigate scenarios that feature a limit cycle (Rosenzweig and MacArthur, 1963).

Individual-based simulations

In order to simulate the ecological models introduced above, we use an individual-based modelling approach that relies on a modified Gillespie algorithm (Allen and Dytham, 2009). The main difference to a classical Gillespie algorithm consists in calculating maximum rate constants which speeds up the simulation as updating occurs less frequently. More generally, our approach assumes that birth and death events happen stochastically. Via increasing or decreasing birth and death rates but keeping the resulting

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

intrinsic rates of increase constant ($d \in [0, 1]$, $b = r_0 + d$, $r_0 = 0.1$), we can simulate biologically relevant increases or decreases in demographic stochasticity and therefore process error σ_{proc}. This approach has the disadvantage that the system's frequency is changed at the same time, which is especially disturbing for predator-prey dynamics. We therefore additionally include the possibility of a constant, additive process error in our predator-prey simulations.

We initially tested the coeffient of variation $\frac{\sigma}{\mu}$ of a logarithmic growth model in its equilibrium density $K = 10^4$ and found that it increases approximately linearly with d in that range, experiencing a maximum of 0.05. Hence we identify σ_{proc} with d. See Fig. S1 for examples of stochastic time series and their deterministic counterparts.

Finally, we sample from the generated time series at varying time intervals to include observation error (Fig. 1). The first three days, sampling occurs every 12 hours, then every 24 hours until 14 days. We assume a binomial sampling process with a fixed sampling fraction p, e.g. the fraction of space monitored or the faction of volume sampled. Observation error σ_{obs} is quantified as variation in equilibrium density K, which reads $\sigma_{\text{obs}} = \frac{\sigma}{\mu} = \sqrt{\frac{1-p}{Kp}}$ for the binomial distribution. Sampling fraction p is randomly chosen from the interval $[0.01, 1.0]$ (on a logarithmic scale), which for $K = 10^4$ results in observation errors between 0 and ≈ 0.1.

For each scenario, that is each population or community model, we ran 10,000 simulations with varying process and observation error. In each simulation we ran 10 independent replicates. In the two-species scenarios, additional 5 independent replicates for the two single-species systems, each, were carried out. They serve as control data that inform model parameters in addition to the two-species treatments.

Empirical data example

In order to confront our statistical approach with empirical data, we complemented the above described simulations by using population times series data from microbial laboratory systems. More precisely, we used the data collected by Fronhofer et al. (2020) from microcosms of the freshwater protist *Tetrahymena thermophila*. These cultures were grown from low density in volumes of 20 mL. Data was collected using a computer-vision and video-analysis pipeline on sample volumes of 31 µL. For details see Fronhofer et al. (2020).

Statistical approach

Let Y_i denote the observations at times t_i ($i = 1, \ldots, n$) of one time series replicate of an experiment. Each observation is univariate for single species systems ($Y_i \in \mathbb{R}$), or bivariate for two-species systems ($Y_i \in \mathbb{R}^2$). Generally, we use m time series replicates denoted by Y_{ij} ($i = 1, \ldots, n$, $j = 1, \ldots, m$). In the

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Table 1: Important model parameters, their meaning and tested values.

Parameter	**Values**	**Meaning**
r_0	0.1	resource intrinsic growth rate
d_R	$\in [0, 1]$	resource death rate (also determines process error)
b	$r_0 + d_R$	resource birth rate
K	10^4	resource carrying capacity
	$2 \cdot 10^4$	resource carrying capacity (predator-prey model)
α	$\frac{r_0}{K}$	resource intraspecific competition coefficient (logistic model)
	$\alpha(r_0, \eta, \gamma, K)$	resource intraspecific competition coefficient (logistic Allee model)
β	$\frac{r_0}{d_R K}$	resource intraspecific competition coefficient (Beverton-Holt model)
	$\beta(r_0, d_R, \eta, \gamma, K)$	resource intraspecific competition coefficient (Beverton-Holt Allee model)
η	0.125	Allee effect strength at low densities
γ	0.0005	Allee effect strength at high densities
$\alpha_{i,j}$	$\in [0.5 \cdot 10^{-5}, 1.5 \cdot 10^{-5}]$	resource interspecific competition coefficient
a	$3.333 \cdot 10^{-4}$	consumer search efficiency
h	0.5	consumer handling time for type II functional response
e	0.05	consumer assimilation efficiency
d_C	$0.05 + c$	consumer death rate
c	$0.1 \cdot d_R$	additive constant for consumer birth and death rate (determines process error)
T_{max}	300	maximal simulation time
R_0	$\in [10^2, 10^3]$	resource starting density
C_0	400	consumer starting density

following, we describe the model fitting of single time series replicates for simplicity and briefly explain extensions to multiple replicates where needed.

$U(t)$ is a deterministic prediction, or process model, for the population time series Y. Here we use $U(t) = U(t|t_1, U_1, \theta)$, the numerical solution of the continuous-time ODE $\frac{\partial U}{\partial t} = f(U(t), \theta)$ with initial value $U(t_1) = U_1$ and model parameters θ. A discrete-time process model $U(t)$ would work analogously.

We combine the deterministic prediction model with a statistical model to fit it to the observations and to estimate model parameters θ. The following three statistical models vary in their treatment of observation and process error, i.e. their variance structure. Mathematically, they differ in the way $U_i = U(t_i)$ is computed from information at the previous time point t_{i-1}: either from the previous prediction U_{i-1}, the previous observation Y_{i-1}, or the underlying previous true state Z_{i-1}. We use a Bayesian approach for parameter estimation, but the models can generally be fitted with maximum likelihood estimation as well (e.g. Xu et al., 2019; DeLong and Lyon, 2020). Both methodologies require the evaluation of the likelihood function $\mathcal{L}(\theta) = p(Y|\theta) = \prod_{i=1}^{n} p(Y_i|\theta)$, where $p(Y|\theta)$ denotes the probability density function of the observed data Y given the model parameters θ.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Observation error model

When ignoring process error, the whole trajectory

$$U(t|t_1, U_1, \theta), \quad t \in [t_1, t_n] \tag{7}$$

is computed with predictions $U_1, \ldots, U_n$. These are confronted with the data by evaluating the likelihood $\mathcal{L}(\theta, U_1)$. The initial abundance $U(t_1) = U_1$ is also a free parameter to be estimated. Eq. 7 is identical to

$$U_i = U(t_i|t_{i-1}, U_{i-1}, \theta), \quad i = 2, \ldots, n \tag{8}$$

and predictions U_i are (iteratively) defined by initial state U_1 and model parameters θ (Fig. 1b). The likelihood $\mathcal{L}(\theta, U_1)$ is defined by assuming pure observation error

$$Y_i \sim \mathcal{D}_{\text{obs}}(U_i), \quad i = 1, \ldots, n \tag{9}$$

with some distribution $\mathcal{D}_{\text{obs}}$. E.g., in case of a normal distribution, Eq. 9 would read $Y_i \sim \mathcal{N}(U_i, \sigma)$ or, equivalently, $Y_i = U_i + \epsilon_i$, with residual errors $\epsilon_i \sim \mathcal{N}(0, \sigma)$ and standard deviation σ. Here, we chose a negative binomial distribution which works with our integer observations Y_i (including zeros) and can account for overdispersion (O'Hara and Kotze, 2010). For time series that did not include zeros as our empirical data example, we found that a lognormal distribution works as well. By neglecting process error and assuming that the process is sufficiently described by a deterministic trajectory, parameter estimation reduces to a nonlinear regression problem, fitting $U(t)$ (Eq. 7) to observations Y_i with independent residuals (Eq. 9).

In case of multiple time series replicates (observations Y_{ij}), we fitted m trajectories $U(t|t_1, U_{1j}, \theta)$ using a joint set of model parameters θ, but allowing individual initial values U_{1j} $(j = 1, \ldots, m)$ in one statistical model, i.e. using a joint likelihood function $\mathcal{L}(\theta, U_{11}, \ldots, U_{1m})$.

Process error model

When observation error is ignored, we assume that the observations Y_i are sufficiently close to the true states. Predictions are generated one-step-ahead

$$U_i = U(t_i|t_{i-1}, Y_{i-1}, \theta), \quad i = 2, \ldots, n \tag{10}$$

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

i.e. predicting U_i from the previous observation Y_{i-1} only (Fig. 1c). The deviations from this piecewise deterministic process

$$Y_i \sim \mathcal{D}_{\text{proc}}(U_i), \quad i = 2, \ldots, n \tag{11}$$

with some distribution for the process error $\mathcal{D}_{\text{proc}}$ define the likelihood $\mathcal{L}(\theta)$. Again, we used a negative-binomial distribution. Thus, by neglecting observation error, parameter estimation reduces to a nonlinear autoregressive problem with independent residuals.

For multiple time series replicates, we straightforwardly iterate Eqs. 10,11 for Y_{ij} and U_{ij} ($j = 1, \ldots, m$) with a joint set of model parameters θ, to evaluate the joint likelihood $\mathcal{L}(\theta)$ in one statistical model.

State-space model

This approach assumes both observation error and process error are present. It requires explicitly modelling the time series of true states Z_i ($i = 1, \ldots, n$). These parameters are unknown a-priori, but can be estimated together with the model parameters θ from the data during the model fitting process. For each "guess" of Z and θ, predictions are generated one-step-ahead

$$U_i = U(t_i | t_{i-1}, Z_{i-1}, \theta), \quad i = 2, \ldots, n \tag{12}$$

from the previous states Z_{i-1} (Fig. 1d). The process error in that prediction is quantified via

$$U_i \sim \mathcal{D}_{\text{proc}}(Z_i), \quad i = 2, \ldots, n \tag{13}$$

with some distribution for the process error $\mathcal{D}_{\text{proc}}$. Here, we used a lognormal distribution allowing continuous parameters Z_i. Additionally, the observations deviate from the true states defined by

$$Y_i \sim \mathcal{D}_{\text{obs}}(Z_i), \quad i = 1, \ldots, n \tag{14}$$

with some distribution for the observation error $\mathcal{D}_{\text{obs}}$. Again, we chose a negative-binomial distribution due to the integer observations. Eqs. 13,14 define the likelihood function $\mathcal{L}(\theta, Z_1, \ldots, Z_n)$ for each θ and Z.

When dealing with multiple time series replicates Y_{ij} ($j = 1, \ldots, m$), we fit m individual time series of true states $Z_{1j}, \ldots, Z_{nj}$ with a joint set of θ, using the likelihood function $\mathcal{L}(\theta, Z_{11}, \ldots, Z_{nm})$ in one statistical model.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Parameter estimation

We used MCMC to sample from the posterior probability distribution of the model parameters given the observations $p(\theta|Y) \sim p(Y|\theta) \cdot p(\theta)$, where $p(Y|\theta) = \mathcal{L}(\theta)$ is the likelihood function and $p(\theta)$ denotes some prior distribution for the model parameters θ. We coded the models using the "rstan" package (Stan Development Team, 2018) and used the built-in Runge-Kutta method for numerical solutions of ODE predictions $U(t)$, and the no-u-turn sampler for computing the posterior. Vague or uninformative prior distributions $p(\theta)$ were chosen for all model parameters to guarantee that the measured model performance was not confounded with prior information (Table S1).

Each model fit was computed by 2000 warmup steps and 2000 samples in three chains, adding up to 6000 posterior samples. We discarded every dataset from subsequent analysis if either of the three fitting methods did non converge (more than 100 divergent iterations).

Evaluation

For each model fit, we computed Bias and root mean squared error (RMSE) from the posterior distribution $p(\widehat{\theta}|Y)$ to evaluate accuracy, where $\widehat{\theta} = (\widehat{\theta}_1, \ldots, \widehat{\theta}_k)$ and $\theta = (\theta_1, \ldots, \theta_k)$ denote the k estimated and true model parameters, respectively. Relative Bias

$$\mathrm{Bias}(\widehat{\theta}_i) = \frac{\mathrm{E}(\widehat{\theta}_i) - \theta_i}{\theta_i} \tag{15}$$

is a measure of point estimates (posterior means $\mathrm{E}(\widehat{\theta}_i)$), while the relative RMSE

$$\mathrm{RMSE}(\widehat{\theta}_i) = \frac{\sqrt{(\mathrm{E}(\widehat{\theta}_i) - \theta_i)^2 + \mathrm{Var}(\widehat{\theta}_i)}}{\theta_i} \tag{16}$$

also accounts for the uncertainty of the estimation (by including posterior variances $\mathrm{Var}(\theta_i)$). Additionally, we computed generalized additive models (GAMs) for geometric means over all k parameters of Bias and RMSE, respectively, versus process and observation error with the "mgcv" package (Wood, 2011).

Results

Fitting simulated data

In most scenarios, SMM performed best in terms of the model parameters' bias (Fig. 2) and RMSE (Fig. 3). In cases where models were identifiable via SSM, OBS performed comparably or just slightly less accurately, while PROC generally produced the least accurate parameter estimates. The dependency of accuracy on the data's process and observation error is presented in Figs. S2–S13.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

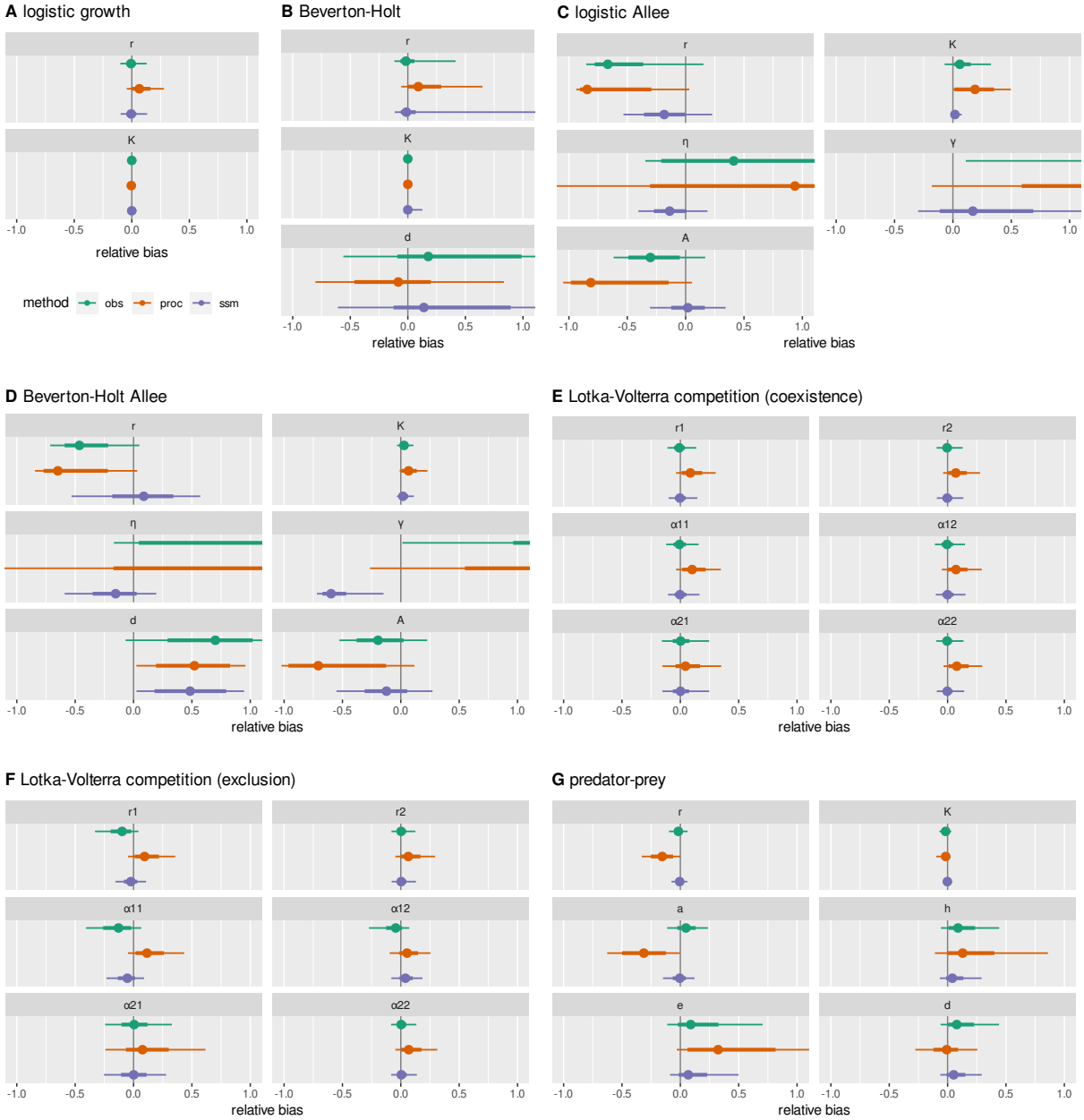

Figure 2: Distribution of relative bias, model parameters vs. statistical models for all population dynamics models (10,000 simulated time series each). Dots are mean, thin lines are 95%, and bold lines are 66% quantiles. For Allee models, A is computed from other model parameters and not a free parameter itself.

For the **logistic growth** model (Eq. 1), both SSM and OBS produced unbiased estimates. PROC overstimated the growth rate r slightly but systematically, especially for high observation error (Fig. S2).

In the **Beverton-Holt** model (Eq. 3), parameter d defines the mortality rate as well as the level of process error in the individual-based simulations. Further, with increasing d, the density-regulation function converges to the linear version of logistic growth and time series resemble those of the logistic growth model. In our parametrization, a nonlinear effect in the density-regulation function on the stochastic time series was only visible for $d < 0.5$ (approximately). Therefore, the exact values of large d

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

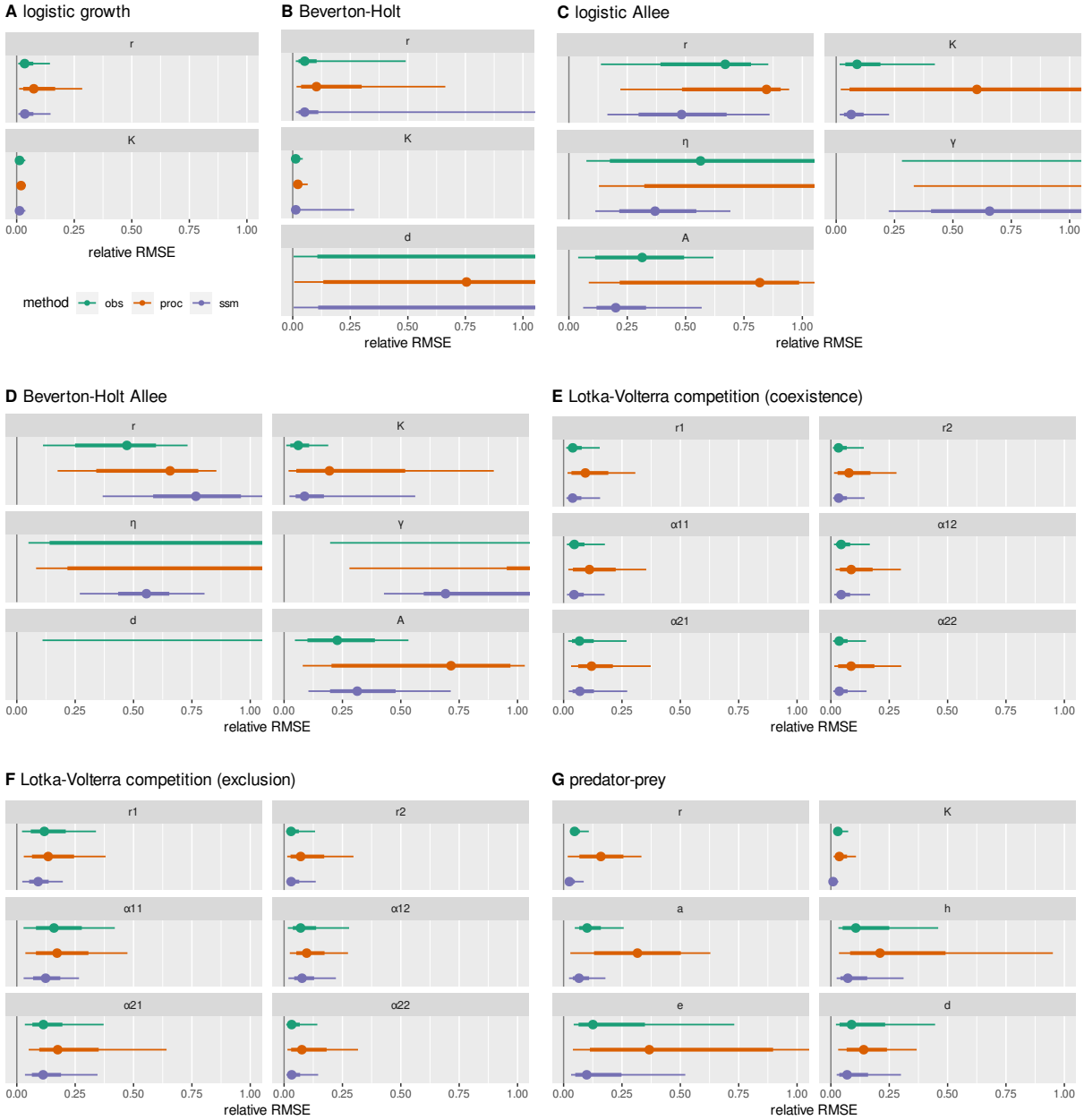

Figure 3: Distribution of relative RMSE, model parameters vs. statistical models for all population dynamics models (10,000 simulated time series each). Dots are mean, thin lines are 95%, and bold lines are 66% quantiles. For Allee models, A is computed from other model parameters and not a free parameter itself.

were not identifiable, which caused high variation in the estimates and their uncertainty. But if the data provided evidence for the Beverton-Holt model vs. logistic growth (d small), then all three statistical models provided accurate estimates (Fig. S3).

The parameter estimates of the **logistic growth with Allee effect** model (Eq. 2) were generally biased. SSM still produced more accurate estimates than both other approaches, while OBS and PROC estimates were heavily biased if not process and observation error in the data were both low (Fig. S4).

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Their inaccuracy was most noticeable for high observation error. All statistical models suffered from highly correlated parameters r, η and γ in the posterior distribution (of a single fit). We assume that this four parameter model is not accurately identifiable under the presence of process and observation error. However, the critical point $A(r, K, \eta, \gamma)$ (see Supplementary Information), which separates positive and negative population growth, was estimated comparably accurately by the SSM.

With one additional parameter d for mortality, the **Beverton-Holt with Allee effect** model (Eq. 4) featured problems similar to the Beverton-Holt model and the logistic Allee effect model. While an accurate estimation of d was not possible unless d was small (and therefore process error was low), other parameters' estimation accuracy was generally low, too. Again, we assume that this five parameter model is not identifiable with process or observation error present in the data (Fig. S5).

We tested a two-species **Lotka-Volterra competition** (Eq. 5) model in a **coexistence** scenario (both species reached a positive steady state) using additional control data (single species time series, each growing to their carrying capacity). The accuracy level was generally high. While OBS and SSM mostly produced unbiased estimates, PROC slightly overestimated some parameters especially under the presence of observation error (Fig. S6), comparable to the logistic growth model. Without control data, results were similar albeit slightly less accurate (Fig. S9).

In a **competitive exclusion** scenario (one species was outcompeted and went extinct, while the other one grew to its carrying capacity) using additional control data, accuracy levels were widely high and just slightly lower than in the coexistence scenario. SSM produced marginally better estimates than the other two fitting approaches. Estimation errors grew with observation error (Fig. S7), especially for parameters r_1, α_{11} and α_{12} of the outcompeted species 1. Without control data however, estimation accuracy decreased rapidly with observation error (Fig. S10), leading to biased estimates for all three fitting approaches. Without data on the carrying capacity $K_1 = \frac{r_1}{\alpha_{11}}$ of the outcompeted species, model identifiability was highly sensitive towards observation error.

We tested a **predator-prey** model (Eq. 6a, 6b) with cyclic dynamics. The parameters were chosen such that the system experienced approximately two full cycles in the observed time. First, we used additional single-species control time series (resource growing to its carrying capacity, consumer going extinct). SSM and OBS provided mostly unbiased estimates, with slight overestimation of parameters e, h and d_C. Here, the posterior distributions of the individual fittings showed correlations in these parameters. PROC generally featured biased estimates for the species interaction parameters e, a and h. Surprisingly, error levels of all three approaches were sensitive to observation error, but not so much to process error (Fig. S8). Second, we fitted the model without any control data (Fig. S11). While accuracy decreased for parameters estimated with OBS, SSM results surprisingly were comparable to the estimates with control data.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Additional scenarios

The OBS and the SSM fitting approach performed comparably well in some two-species scenarios, even under the presence of process and observation error. We repeated the analysis of the Lotka-Volterra competition model and the predator-prey model to validate the methods under more challenging conditions.

We fitted the **Lotka-Volterra competition** model using **fewer replicates**, i.e. only two instead of ten time series of the two-species mixtures, and only one instead of five single-species control time series each. Our assumption was that, even though process error can speed up or slow down dynamics, these effects would average out over ten replicates and would inform parameters (e.g. the average growth rates) correctly. We found that this generally holds even with fewer replicates, and results were only slightly less accurate (Fig. S12). In particular, OBS estimates were as accurate as SSM estimates across all levels of observation and process error.

We also fitted **predator-prey** models to **longer time series** over 35 days instead of 14 days, such that the system featured approximately four full cycles instead of two. Additional single-species control time series were used as before. Our assumptions were that process error did not affect the estimation quality significantly in the shorter time series, since the stationarity of the process (regularity of the cycles) was not seriously disturbed, and that regularity would only decrease with longer time series, leading to less accurate results especially for the OBS model. Fitting longer time series (Fig. S13), results were however comparable to the results with shorter time series. Notably, OBS estimated parameters only marginally less precise than for the shorter time series. We conclude that OBS and SSM are robust against the level of process variation and the resulting non-stationarity in the investigated simulations.

Fitting an empirical data set

We fitted the Beverton-Holt model (Eq. 3) with all three statistical approaches to seven empirical datasets, comprised of six time series replicates each (Fig. 4).

Since all observed densities were positive, we used lognormally distributed residuals for process and observation errors, both. Nonlinear effects on the density-regulation function were detected in datasets 4, 6 and 7 by all three statistical models, indicated by low values of mortality rate d. For the remaining datasets, larger values of d were estimated with a high uncertainty in the posterior distributions, which suggested no clear evidence for the Beverton-Holt model over the logistic growth model. Estimates of the three model parameters r, K and d were similar across the three statistical approaches, but posterior uncertainty was generally higher in the PROC estimates. The SSM indicated that process error was smaller than observation error in all datasets. OBS and PROC however have only one error term that has to account for the total variation, so estimates of observation error and process error were higher

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

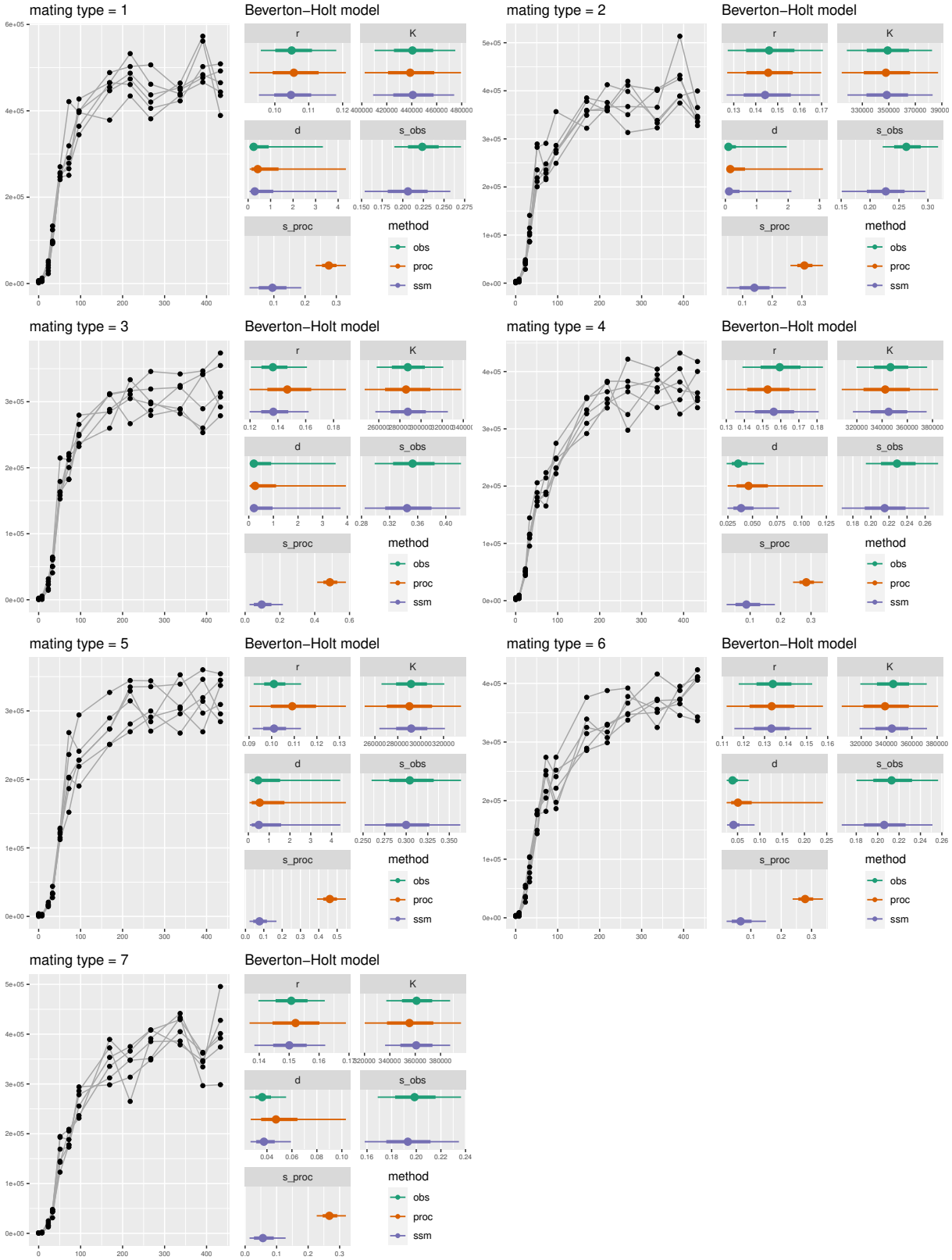

Figure 4: Empirical time series and posterior distributions of fitted Beverton-Holt models, including model parameters r, K, d and standard deviation parameters $s_{\text{obs}}, s_{\text{proc}}$ of the lognormal residuals distributions.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

than the respective SSM estimates.

Discussion

Our work shows that using deterministic ODEs and Bayesian inference, it is possible to accurately estimate parameters from time series data of stochastic IBMs including process and observation error. Practical identifiability, which is based both on the model and data quality (Raue et al., 2009), was validated for several widely-used population models with respect to the different statistical models used here. Importantly, our work covers multiple dimensions of complexity, both in terms of population ecology and statistical models, from simple single species models, such as the logistic model, to multispecies community models, and, in terms of statistical models, from pure observation (OBS), via process error (PROC) to state-space models (SSM). Overall, we can show that OBS and SSM can be generally preferred over PROC models.

More precisely, OBS can compete with the more complex SSM, especially when a true steady state > 0, such as an equilibrium density, is reached (logistic model, Beverton-Holt model (d small, so not logistic data), Lotka-Volterra model with coexistence) or when the biological models lead to regular cycles, such as in some predator-prey dynamics. The good performance of OBS for predator-prey models is in line with recent studies (Rosenbaum et al., 2019; DeLong and Lyon, 2020). Interestingly, even for 2-species Lotka-Volterra competition models with exclusion of one of the players, we can show that OBS performs only slightly worse than SSM. Therefore, OBS may be a viable alternative to SSM, especially if SSM suffer from, e.g., extensive model complexity or convergence problems (Bolker, 2008; Auger-Méthé et al., 2016; Best and Punt, 2020).

Our work also shows that estimating Allee effect strengths and more generally fitting population growth models with Allee effects is challenging. For these models, the 4- or even 5-parameter density-regulation functions may be overparameterised and nonidentifiable especially in the presence of process and observation error. As a consequence, these models may only be useful for data from highly controlled experiments in combination with SSMs. It it important to note that these models can be derived mechanistically (Thieme, 2003), which may allow to inform priors of one or multiple model parameters which could help make the SSM more accurate. Of course, as an alternative, non-mechanistic formulations such as $\frac{dN}{dt} = rN(1 - \frac{N}{K})(\frac{N}{A} - 1)$ with less parameters could also be used.

In the case of two-species systems we highly recommend the use of control time series from single species settings (Figs. S6 & S9, S7 & S10, S8 & S11). If these are not available SSMs should be used. Biological settings that lead to the exclusion of one species, such as some of the Lotka-Volterra competition models we have used here, make the statistical models non-identifiable without such control data, unless

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

the observation error is very small.

While we here explored model identifiability based on time series data alone, a Bayesian approach easily allows including additional sources of information to improve parameter estimation. This information could enter the model via informed priors or hierarchical models (Kindsvater et al., 2018), or by using multiple, potentially heterogeneous data sources (Barraquand and Gimenez, 2019). For example, feeding experiments (Rosenbaum and Rall, 2018) additionally inform functional response parameters in predator-prey models (Barraquand and Gimenez, 2021).

In conclusion, we have explored multiple dimensions of complexity, both in terms of biological complexity as well as in terms of statical model complexity, in order to pinpoint which error structure one should use when fitting classical deterministic ODE models to empirical data, from single species to community dynamics and trophic interactions. Our results show that, overall, observation and state-space models outperform process error models. Importantly, our continuous-time models allow to include uneven sampling intervals (and therefore missing values), because the model is not linearised within a time-step as in discrete-time models. More generally, our work shows that deterministic models seem to be sufficient to describe the stochastic dynamics emerging from process and observation errors.

Code and data availability

Code for stochastic individual-based simulations is available from https://doi.org/10.5281/zenodo.5500442. A tutorial for fitting ODE models to time series data in R is presented in the Supplementary Information and is also available online https://github.com/benjamin-rosenbaum/fitting_deterministic_population_models.

Acknowledgements

This is publication ISEM-YYYY-XXX of the Institut des Sciences de l'Evolution – Montpellier. This work was supported by a grant from the Agence Nationale de la Recherche (No.: ANR-19-CE02-0015) to EAF. BR gratefully acknowledges the support of iDiv funded by the German Research Foundation (DFG-FZT 118, 202548816). The scientific results have in part been computed at the High-Performance Computing Cluster EVE of the Helmholtz Centre for Environmental Research (UFZ) and iDiv, and we thank Christian Krause for technical support.

Authors contributions

BR and EAF conceived the ideas and designed methodology. EAF developed the simulation model and BR developed the statistical analysis framework. BR analysed the data. Both authors contributed to

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

writing the manuscript.

Conflict of interest

The authors declare no conflict of interest.

References

Allen, G. E. and Dytham, C. 2009. An efficient method for stochastic simulation of biological populations in continuous time. – Biosystems 98(1): 37–42.

Auger-Méthé, M., Field, C., Albertsen, C. M., Derocher, A. E., Lewis, M. A., Jonsen, I. D. and Mills Flemming, J. 2016. State-space models' dirty little secrets: even simple linear Gaussian models can have estimation problems. – Scientific reports 6: 26677.

Auger-Méthé, M., Newman, K., Cole, D., Empacher, F., Gryba, R., King, A. A., Leos-Barajas, V., Mills Flemming, J., Nielsen, A., Petris, G. and Thomas, L. 2021. A guide to state–space modeling of ecological time series. – Ecological Monographs n/a(n/a).

Barraquand, F. and Gimenez, O. 2019. Integrating multiple data sources to fit matrix population models for interacting species. – Ecological modelling 411: 108713.

— 2021. Fitting stochastic predator-prey models using both population density and kill rate data. – Theoretical population biology 138: 1–27.

Best, J. K. and Punt, A. E. 2020. Parameterizations for Bayesian state-space surplus production models. – Fisheries research 222: 105411.

Bolker, B. M. 2008. Ecological Models and Data in R. – Princeton University Press.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. – Annu. Rev. Ecol. Syst. 31: 343–366.

Clark, F., Brook, B. W., Delean, S., Reşit Akçakaya, H. and Bradshaw, C. J. A. 2010. The theta-logistic is unreliable for modelling most census data. – Methods Ecol. Evol. 1(3): 253–262.

Clark, J. S. and Bjørnstad, O. N. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. – Ecology 85(11): 3140–3150.

DeLong, J. P. and Lyon, S. 2020. Temperature alters the shape of predator–prey cycles through effects on underlying mechanisms. – PeerJ 8: e9377.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Fronhofer, E. A. and Altermatt, F. 2015. Eco-evolutionary feedbacks during experimental range expansions. – Nat. Commun. 6: 6844.

Fronhofer, E. A., Govaert, L., O'Connor, M. I., Schreiber, S. J. and Altermatt, F. 2020. The shape of density dependence and the relationship between population growth, intraspecific competition and equilibrium population density. – bioRxiv .

Godwin, C. M., Chang, F.-H. and Cardinale, B. J. 2020. An empiricist's guide to modern coexistence theory for competitive communities. – Oikos 129(8): 1109–1127.

Hefley, T. J., Tyre, A. J. and Blankenship, E. E. 2013. Fitting population growth models in the presence of measurement and detection error. – Ecol. Model. 263: 244–250.

Hiltunen, T., Hairston, N. G., Hooker, G., Jones, L. E. and Ellner, S. P. 2014. A newly discovered role of evolution in previously published consumer-resource dynamics. – Ecol. Lett. 17(8): 915–923.

Holling, C. S. 1959. The components of predation as revealed by a study of small-mammal predation of the European pine sawfly. – Can. Entomol. 91(05): 293–320.

Kindsvater, H. K., Dulvy, N. K., Horswill, C., Juan-Jordá, M.-J., Mangel, M. and Matthiopoulos, J. 2018. Overcoming the Data Crisis in Biodiversity Conservation. – Trends in ecology & evolution 0(0).

Mallet, J. 2012. The struggle for existence: how the notion of carrying capacity, K, obscures the links between demography, Darwinian evolution, and speciation. – Evol. Ecol. Res. 14(5): 627–665.

Nørgaard, L. S., Zilio, G., Saade, C., Gougat-Barbera, C., Hall, M. D., Fronhofer, E. A. and Kaltz, O. 2021. An evolutionary trade-off between parasite virulence and dispersal at experimental invasion fronts. – Ecol. Lett. 24(4): 739–750.

O'Hara, R. B. and Kotze, D. J. 2010. Do not log-transform count data. – Methods in ecology and evolution / British Ecological Society 1(2): 118–122.

Raue, A., Kreutz, C., Maiwald, T., Bachmann, J., Schilling, M., Klingmüller, U. and Timmer, J. 2009. Structural and practical identifiability analysis of partially observed dynamical models by exploiting the profile likelihood. – Bioinformatics 25(15): 1923–1929.

Rosenbaum, B., Raatz, M., Weithoff, G., Fussmann, G. F. and Gaedke, U. 2019. Estimating parameters from multiple time series of population dynamics using Bayesian inference. – Front. Ecol. Evol. 6: 234.

Rosenbaum, B. and Rall, B. C. 2018. Fitting functional responses: Direct parameter estimation by simulating differential equations. – Methods in ecology and evolution / British Ecological Society 9(10): 2076–2090.

bioRxiv preprint doi: https://doi.org/10.1101/2021.09.13.460028; this version posted September 16, 2021. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC-ND 4.0 International license.

Rosenzweig, M. L. and MacArthur, R. H. 1963. Graphical Representation and Stability Conditions of Predator-Prey Interactions. – Am. Nat. 97: 209–223.

Shoemaker, L. G., Sullivan, L. L., Donohue, I., Cabral, J. S., Williams, R. J., Mayfield, M. M., Chase, J. M., Chu, C., Harpole, W. S., Huth, A., HilleRisLambers, J., James, A. R. M., Kraft, N. J. B., May, F., Muthukrishnan, R., Satterlee, S., Taubert, F., Wang, X., Wiegand, T., Yang, Q. and Abbott, K. C. 2020. Integrating the underlying structure of stochasticity into community ecology. – Ecology 101(2): e02922.

Sibly, R. M., Barker, D., Denham, M. C., Hone, J. and Pagel, M. 2005. On the Regulation of Populations of Mammals, Birds, Fish, and Insects. – Science 309(5734): 607–610.

Stan Development Team 2018. RStan: the R interface to Stan. – R package version 2.18.2.

Thieme, H. R. 2003. Mathematics in Population Biology. – Princeton University Press.

de Valpine, P. and Hastings, A. 2002. Fitting population models incorporating process noise and observation error. – Ecological Monographs 72(1): 57–76.

Verhulst, P.-F. 1838. Notice sur la loi que la population suit dans son accroissement. – Correspondance Mathématique et Physique 10: 113–121.

Wood, S. N. 2011. Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. – Journal of the Royal Statistical Society. Series B, Statistical methodology 73(1): 3–36.

Xu, L., Li, B., Chen, X. and Chen, Y. 2019. A comparative study of observation-error estimators and state-space production models in fisheries assessment and management. – Fish. Res. 219: 105322.

Yoshida, T., Jones, L. E., Ellner, S. P., Fussmann, G. F. and Hairston, N. G. 2003. Rapid evolution drives ecological dynamics in a predator–prey system. – Nature 424(6946): 303–306.

Zurell, D., Berger, U., Cabral, J. S., Jeltsch, F., Meynard, C. N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. and Grimm, V. 2010. The virtual ecologist approach: simulating data and observers. – Oikos 119: 622–635.