Confronting population models with experimental microcosm 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

Correspondence: benjamin.rosenbaum@idiv.de

Manuscript type: Article

Running title: Confronting population models with microcosm data

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

Open research: 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. Data (Fronhofer et al., 2020) will be provided via GitHub and Zenodo.

Abstract

Population and community ecology traditionally has a very strong theoretical foundation with well-known dynamical models, such as the logistic and its variations, and many modification of the classical Lotka-Volterra predator-prey and interspecific competition models. More and more, these classical models are being confronted with data via fitting to empirical time series for purposes of projections or for estimating model parameters of interest. However, using statistical models to fit theoretical models to data is far from trivial, especially for time series data where subsequent measurements are not independent. This raises the question of whether statistical inferences using pure observation error models, such as simple (non-)linear regressions, are biased, and whether more elaborate process error models or state-space models have to be used to address this complexity.

In order to help empiricists, especially researchers working with experimental laboratory populations in micro- and mesocosms, make informed decisions about the statistical formalism to use, we here compare different error structures one could use when fitting classical deterministic ODE models to empirical data. We consider a large range of biological scenarios and theoretical models, from single species to community dynamics and trophic interactions. In order to compare the performance of different error structure models, we use both realistically simulated data and empirical data from microcosms in a Bayesian framework.

We find that many model parameters can be estimated precisely with an appropriate choice of error structure using pure observation error or state-space models, if observation errors are not too high. However, Allee effect models are typically hard to identify and state-space models should be preferred when model complexity increases.

Our work shows that, at least in the context of low environmental stochasticity and high quality observations, deterministic models can be used to describe stochastic population dynamics that include process variability and observation error. We discuss when more complex state-space model formulations may be required for obtaining accurate parameter estimates. Finally, we provide a comprehensive tutorial for fitting these models in R.

Introduction

Studying biotic interactions and measuring their strength in order to understand how ecological systems work 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 influence fitness. Biotic interactions are therefore components of eco-evolutionary dynamics and feedbacks (Yoshida et al., 2003; Hiltunen et al., 2014).

As a consequence, correct and unbiased 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 (ODEs), 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 of these different approaches are, especially given the varying degrees of complexity of these approaches and technical skills they may require.

Importantly, these choices are not mere technical details since they 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.

Making a decision regarding the correct statistical formalism to use is not made easier by the fact that the question includes 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.

Of course, we cannot treat all these dimensions comprehensively here. Regarding the first, we will rely on Bayesian inference because Bayesian models are capable of quantifying uncertainty of estimated parameters exactly, even for nonlinear problems, and they perform well with complex models (Clark et al., 2010). Nevertheless, de Valpine and Hastings (2002) compare state-space models to observation and process error models using a likelihood framework and show that state-space models outperform other options.

In terms of the second axis of complexity, we will focus on continuous-time models. Note that Clark and Bjørnstad (2004) study discrete-time models, including exponential growth and the Ricker model and provide a complementary view to the results we will present below.

Regrading the third axis, our study focuses on deterministic theoretical models, which makes our work most relevant when environmental stochasticity is low, such as for data collected in the context of laboratory experimental populations from microcosms or mesocosms. Applications include, for example, plant growth (Paine et al., 2012), Lotka-Volterra competition and predation (Mühlbauer et al., 2020), bacteria or protist predator-prey systems (Rosenbaum et al., 2019; DeLong and Lyon, 2020), host-pathogen interactions (Lunn et al., 2013), and even terrestrial mesocosm foodwebs (Wootton et al., 2022). Data from the field, which may be heavily impacted by environmental stochasticity (Shoemaker et al., 2020), are beyond the scope of our work, and are likely best analysed using stochastic models (see e.g., Barraquand and Gimenez, 2019, 2021). We also do not account for detection error (see Hefley et al., 2013) or model errors (see Xu et al., 2019).

We will here extensively study the fourth axis of complexity, the choice of error structure (Fig. 1). This choice seems especially important when one intends to fit dynamical models to time series data, since in time series, by default, subsequent measurements are not statistically independent. This implies that, besides observation error, process error may have to be taken into account. Nevertheless, researchers often choose a trajectory-matching approach and account for observation error only because this omission reduces model fitting to a comparatively simple nonlinear regression problem. By contrast, process error-only models are nonlinear autoregressive models, where each observation in time is predicted from the previous one. The most complete treatment of error structure involves using state-space models (Auger-Méthé et al., 2021) which take both sources of error into account. Such state-space models cannot be classified as standard regression problems, since they require the simultaneous estimation of latent states as parameters (estimated true population abundances).

In order to understand under which conditions observation error, process error or both have to be modelled when analysing population and community dynamics data with ODE models, 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). Specifically, (1) we choose an ecological model with known parameters and (2) simulate trajectories of population abundances with a stochastic algorithm (Gillespie, 2007). Each "experiment" has a duration of 14 days and is replicated 10 times. Here, process error is introduced as demographic stochasticity via random birth and death events. We do not include any form of environmental stochasticity in the data-generating process. (3) We simulate realistic observations from each of the 10 replicates by sampling every 12–24 hours, each measurement includes observation error. (4) We subsequently fit the appropriate dynamical equations in R with Markov Chain Monte Carlo (MCMC) sampling using Stan (Stan Development Team, 2018). Here, we use three different statistical approaches by accounting for observation error only (OBS), process error only (PROC), or fitting state-space models accounting for both (SSM). For each ecological scenario as described below, we conducted 10,000 experiments, each, while varying the levels

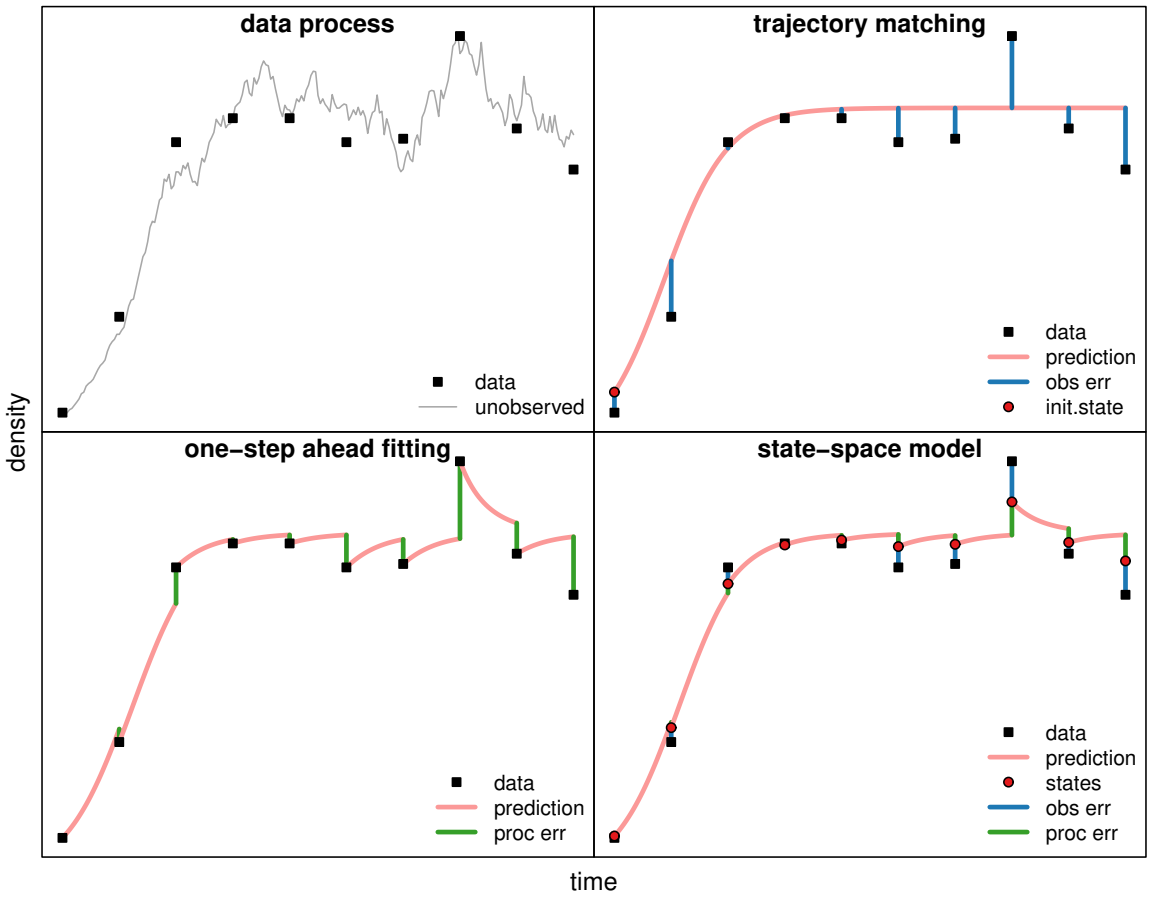

Figure 1: Possible statistical approaches for fitting a population model to a time series of population densities. The upper left panel depicts the underlying stochastic birth-death process (logistic growth) and the black squares show the sampled data with corresponding errors. The simplest statistical model assumes only observation errors and matches the calculated 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 (bottom right) takes both sources of error (observation and process) into account.

of introduced process and observation error. Finally, we complement our work with an analysis of real population dynamic data from microbial laboratory systems using published data (Fronhofer et al., 2020).

Overall, in our analysis, process error models fare least well and state-space models outperform simpler approaches for biological models as complex as predator-prey systems when data is irregular or incomplete. Additionally, we provide a comprehensive tutorial for fitting these models in R.

Materials and methods

Mathematical models

The following ecological models are used for generating the observed data. Example time series are depicted in Fig. S1, see Table S1 for an overview of the parameters.

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 of 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

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

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 conversion factor, which includes consumer-resource body-size ratio as well as assimilation efficiency. In our analyses we will focus on a Holling type II (Holling, 1959), that is, saturating functional response of the form $\frac{aR}{1+ahR}$ 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. The main difference to a classical Gillespie algorithm (exact stochastic simulation algorithm; Gillespie, 2007) consists in calculating maximum rate constants which speeds up the simulation as updating occurs less frequently as detailed in Allen and Dytham (2009). More generally, our approach assumes that birth and death events happen stochastically. Via increasing or decreasing birth and death rates but keeping the resulting intrinsic rate of increase constant ($d \in [0, 1]$, $b = r_0 + d$, $r_0 = 0.1$ [h^{-1}]), we can simulate biologically relevant increases or decreases in demographic stochasticity and therefore process error σ_{proc}. This approach has the advantage that we do not need to assume any scaling of demographic stochasticity with population size a priori. The correct scaling emerges from our model. For predator-prey dynamics, we treat the resource's intrinsic growth rate as above. The predator's rate of change (i.e., conversion factor

e and mortality rate d) is simultaneously varied. Stochasticity here increases the system's frequency, but keeps equilibria ($\frac{dN}{dt} = \frac{dR}{dt} = 0$) intact.

We initially verified computationally that a stochastic time series' variation σ^2 around its deterministic counterpart's equilibrium K scales with $\sigma \sim \sqrt{Kd}$ (Fig. S15). Hence we identify σ_{proc} with $\sqrt{d}$ for fixed K. For example, with the logistic growth model we observed on average $\sigma_{\text{proc}} = 3.072\sqrt{K}$ for $d = 1.0$.

Finally, we sample from the generated time series at varying time intervals to include observation error (Fig. 1). Concretely, here the first three days, sampling occurs every 12 hours, then every 24 hours until 14 days. We count abundances N_{count} in a fixed fraction p of the space monitored or the faction of volume sampled. Assuming a random distribution of individuals in space, this is described by a binomial sampling process $N_{\text{count}} \sim \text{Binomial}(N, p)$. This implies that the estimated abundance $\frac{N_{\text{count}}}{p}$ in the total volume has a mean of $\frac{pN}{p} = N$ and a standard deviation of $\frac{1}{p}\sqrt{p(1-p)N} = \sqrt{\frac{1-p}{p}N}$. Hence we identify σ_{obs} with $\sqrt{\frac{1-p}{p}}$ for fixed N. For example, any population at its carrying capacity $N = K$ features an expected observation error $\sigma_{\text{obs}} = 9.950\sqrt{K}$ for $p = 0.01$, which is approximately three times higher than the process error as described above.

For each scenario, that is, each population or community model, we simulated 10,000 microcosm experiments. Each experiment was assigned a fixed process error and a fixed observation error (death rate $d \in [0.0, 1.0]$ and sampling fraction $p \in [0.01, 1.0]$, random draws from a uniform distribution, for p on logscale). In each experiment we simulated 10 time series replicates with identical parameterisation. In the two-species scenarios, 5 additional replicates for the two single-species dynamics were added. This additional single-species data, in conjunction with the 10 two-species time series, informs intra-specific model parameters like $r_{0,i}$ and $\alpha_{i,i}$ ($i = 1, 2$).

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 a microbial laboratory system.

This dataset represents a empirical example that matches our single species model examples introduced above (Eqs. 1 and 3). Furthermore, it fulfils the prerequisites of low environmental stochasticity discussed in the introduction. Nevertheless, it provides room for testing whether and how the different statistical models can pick up non-linearities in density-regulation (Fronhofer et al., 2020).

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 $20\,\mathrm{mL}$. Data was collected using a computer-vision and video-analysis pipeline on sample volumes of $31\,\mu\mathrm{L}$. For details see Fronhofer et al. (2020).

Statistical approach

Here, we describe the statistical approaches for analyzing our simulated and empirical data. In summary, we combine a deterministic prediction model $U(t)$ with a statistical model to fit it to an observed time series Y_i at times t_i ($i = 1, \ldots, n$) and to estimate model parameters θ. Our three proposed statistical models vary in their treatment of observation and process error, that is, their variance structure. Mathematically, they differ in the way $U_i = U(t_i)$ is predicted from information at the previous time point t_{i-1}: either from the previous prediction U_{i-1} (OBS), the previous observation Y_{i-1} (PROC), or a previous latent state variable Z_{i-1} (SSM). Each model is presented in a separate section below.

More specifically, let Y_i denote untransformed observations (counted abundances in sampling fraction p) of a single time series. Estimated abundances in the total volume are given by $\frac{Y_i}{p}$. Each observation is univariate for single species systems ($Y_i \in \mathbb{N}$), or bivariate for two-species systems ($Y_i \in \mathbb{N}^2$). Generally, we use $m = 10$ non-aggregated time series replicates denoted by Y_{ij} ($i = 1, \ldots, n$, $j = 1, \ldots, m$). In the following, we describe the model fitting of single time series 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 in the total volume $\frac{Y}{p}$. 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 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.

Observation error model (OBS)

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, yielding predictions $U_1, \ldots, U_n$. 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 predictions $p \cdot U_i$ for the sampled fraction p are confronted with the data Y_i by evaluating the likelihood $\mathcal{L}(\theta, U_1)$. For the likelihood, we chose a negative binomial distribution

$$Y_i \sim \mathrm{NB}(p \cdot U_i, \tau), \quad i = 1, \ldots, n \tag{9}$$

which has mean $\mu = p \cdot U_i$ and variance $\sigma^2 = \mu + \frac{\mu^2}{\tau}$. Parameter $\tau > 0$ describes the amount of overdispersion which makes the negative binomial a flexible choice, both for error variances scaling with μ (large τ) or with μ^2 (small τ). Also, it is suited for the integer- and potentially zero-valued observations Y_i (O'Hara and Kotze, 2010). For time series that did not include any 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 single set of model parameters (e.g. $\theta = (r_0, K)$, logistic growth scenario), but allowing individual initial values U_{1j} ($j = 1, \ldots, m$) in one statistical model, i.e. using a joint likelihood function $\mathcal{L}(\theta, \tau, U_{11}, \ldots, U_{1m})$.

Process error model (PROC)

When observation error is ignored, we assume that the observations for the total volume $\frac{Y_i}{p}$ are sufficiently close to the true abundances. Predictions are generated one-step-ahead

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

i.e. predicting U_i from the previous observation Y_{i-1} only (Fig. 1c). The observations' deviation from this piecewise deterministic process is modeled again with a negative binomial distribution

$$Y_i \sim \text{NB}(p \cdot U_i, \tau), \quad i = 2, \ldots, n \tag{11}$$

for reasons stated above, but here it accounts for process error. 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 single set of model parameters θ, to evaluate the joint likelihood $\mathcal{L}(\theta, \tau)$ in one statistical model.

State-space model (SSM)

This approach assumes both observation error and process error are present. It requires explicitly modelling the time series of estimates of true abundances Z_i ($i = 1, \ldots, n$) as latent states. These 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 these predictions is modelled with a lognormal distribution

$$Z_i \sim \text{lognormal}(\log(U_i), \sigma), \quad i = 2, \ldots, n \tag{13}$$

with a scale parameter σ. Hence, $\log(Z_i)$ follows a normal distribution with mean $\log(U_i)$ and standard deviation σ. A continuous distribution is required here for $U_i, Z_i \in \mathbb{R}$. Additionally, the state parameters Z_i are confronted with the integer-valued observations Y_i with a negative binomial distribution as in the previous models

$$Y_i \sim \text{NB}(p \cdot Z_i, \tau), \quad i = 1, \ldots, n \tag{14}$$

accounting for observation error. Eqs. 13,14 define the likelihood function $\mathcal{L}(\theta, \sigma, \tau, Z_1, \ldots, Z_n)$. 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 $\mathcal{L}(\theta, \sigma, \tau, Z_{11}, \ldots, Z_{nm})$ in one statistical model.

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 not converge (more than 100 divergent iterations).

Evaluation

For each of the 10,000 model fits per scenario, we computed bias and root mean squared error (RMSE) for every parameter θ_i 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 (e.g. $\theta_1 = r_0$, $\theta_2 = K$ for logistic growth). Relative bias

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

is a measure of the point estimate (posterior mean $\text{E}(\widehat{\theta}_i)$), while the relative RMSE

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

also accounts for the uncertainty of the estimation (by including the posterior variance $\text{Var}(\widehat{\theta}_i)$). Additionally, we investigated the effects of process and observation error on model accuracy to test, e.g., if the statistical approaches OBS, PROC and SSM are affected differently by the amount of both error sources in the data. First, for each fit, we computed overall measures of accuracy as geometric means $\overline{\text{Bias}} = \prod_{i=1}^{k} |\text{Bias}(\theta_i)|^{1/k}$ and $\overline{\text{RMSE}} = \prod_{i=1}^{k} \text{RMSE}(\theta_i)^{1/k}$ over absolute Bias and RMSE of all model parameters θ_i, respectively. Then, we fitted generalized additive models (GAMs) to these measures using observation error and process error as predictors and visualized the response surfaces, see Figs. S2–S13. We used the "mgcv" package (Wood, 2011) with model formulae "$\overline{\text{Bias}} \sim s(\sigma_{\text{proc}}, \sigma_{\text{obs}})$" and "$\overline{\text{RMSE}} \sim s(\sigma_{\text{proc}}, \sigma_{\text{obs}})$".

Results

Fitting simulated data

When analyzing time series with OBS and PROC, almost all fits converged properly and only less than 1% of all experiments had problems with divergent iterations. For SSM, around 5% of all fits did not converge, with a maximum of 9% for our most complex model (predator-prey system). There are generic strategies when encountering convergence issues, like employing more informative priors, using a longer warmup phase for the MCMC sampler, or increasing the sampler's accuracy at the cost of computation speed (Stan Development Team, 2022), which we did not further investigate case-by-case given the amount of involved experiments.

In most scenarios, OBS and SSM 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 mostly produced the least accurate parameter estimates.

For the **logistic growth** model (Eq. 1), all models produced unbiased estimates. PROC estimates of the growth rate r were slightly less precise, 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 ($\sigma_{\text{proc}} \sim \sqrt{N \cdot d}$). 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 $d < 0.4$ (approximately). Therefore, the exact values of large d 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 mostly biased. OBS and SSM still produced more accurate estimates than PROC. Its estimates were

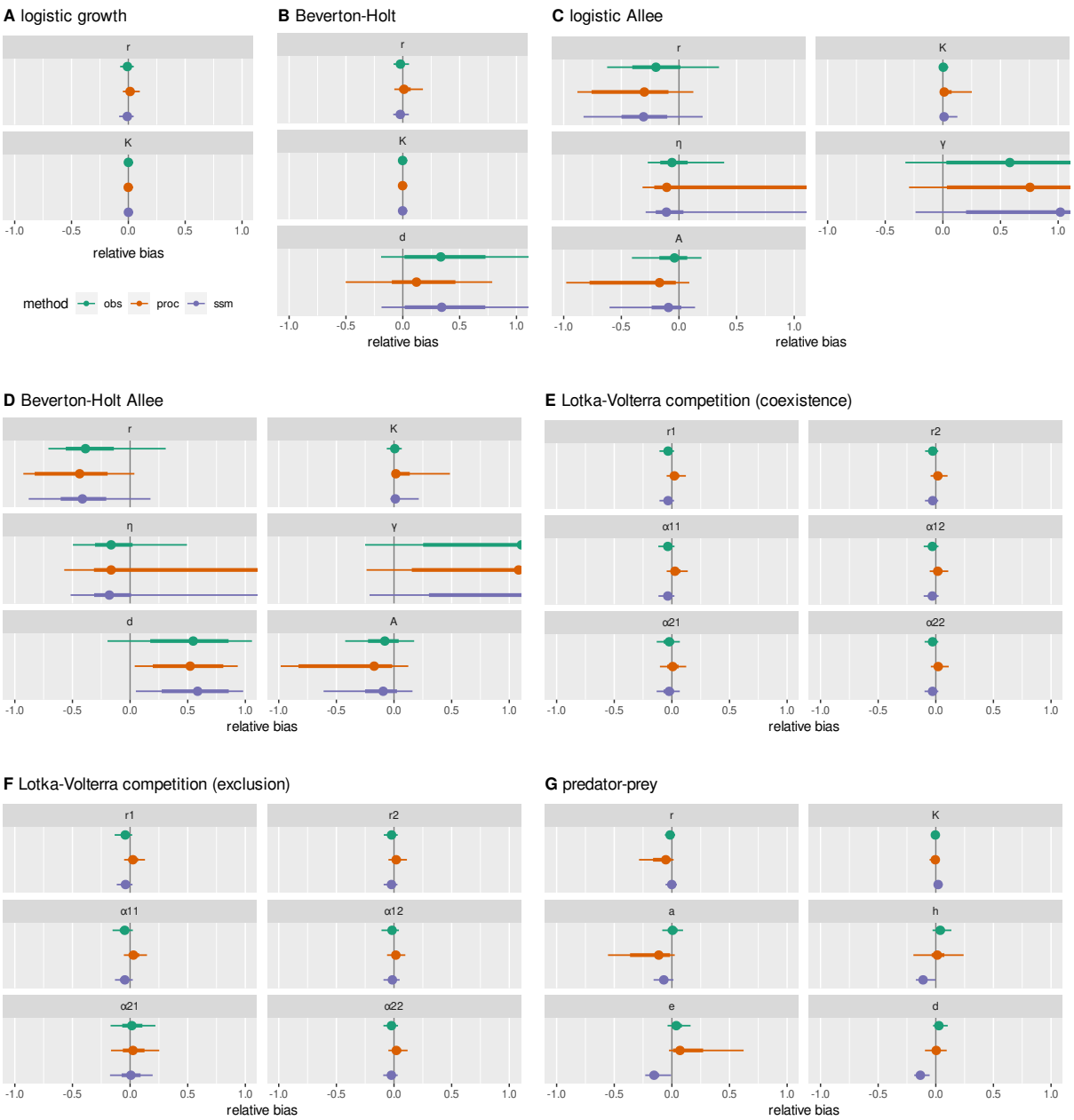

Figure 2: Distribution of relative bias, model parameters vs. statistical models for all population dynamics models (10,000 experiments per scenario). Dots are mean, thin lines are 95%, and bold lines are 66% quantiles. For Allee models, A is computed from other model parameters and not a free parameter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

heavily biased if not both process error and observation error were low (Fig. S4). The inaccuracy was most pronounced 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 emphasize that this is not a problem with MCMC convergence, but rather a problem with practical identifiability (Raue et al., 2009). Information on this nonlinear, four-parameter population growth rate is lost due to process and observational noise, and cannot be recovered

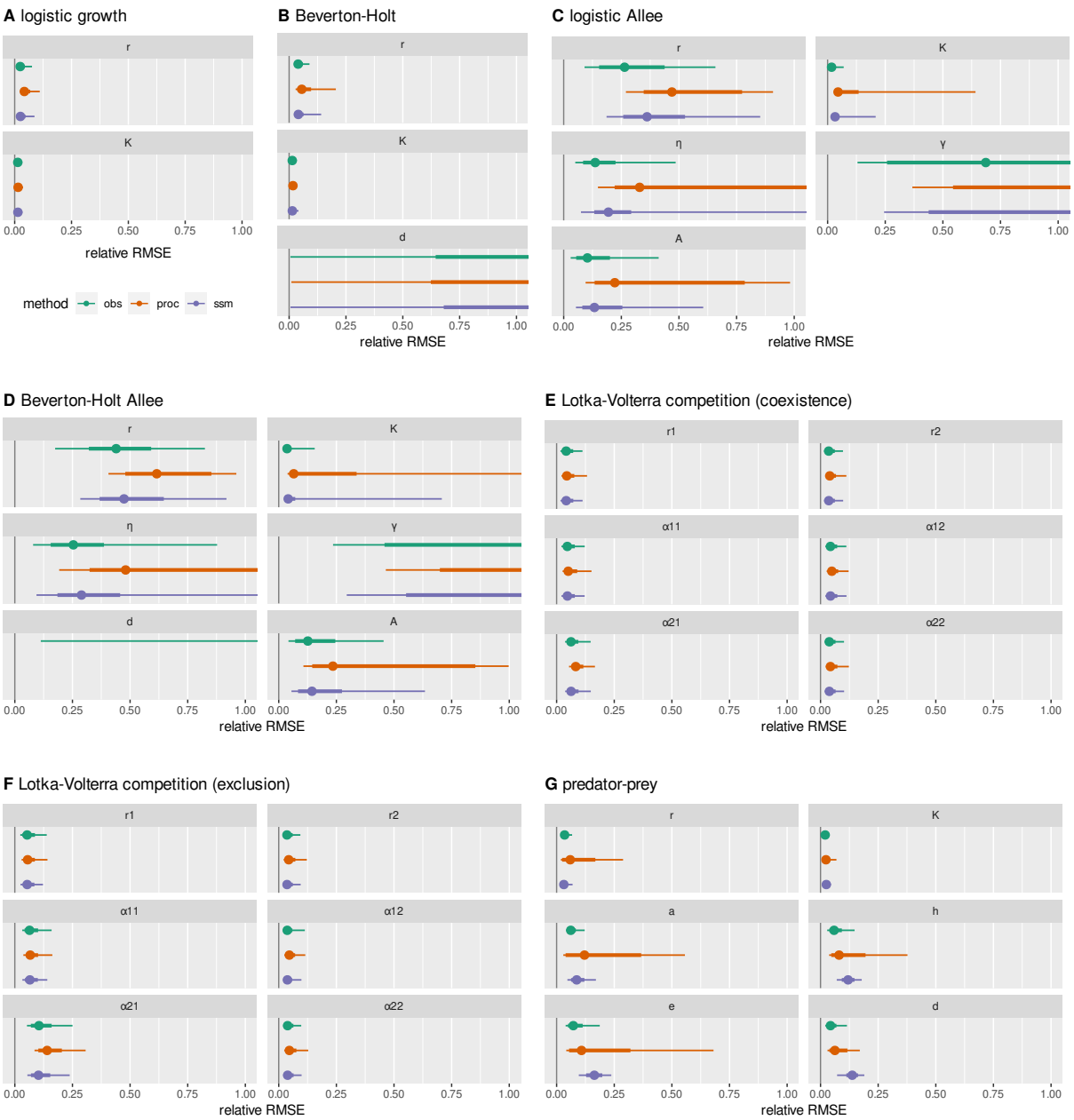

Figure 3: Distribution of relative RMSE, model parameters vs. statistical models for all population dynamics models (10,000 experiments per scenario). Dots are mean, thin lines are 95%, and bold lines are 66% quantiles. For Allee models, A is computed from other model parameters and not a free parameter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

accurately by statistical inference. However, the critical population density $A(r, K, \eta, \gamma)$ (see Supplementary Information), which separates positive and negative population growth, was estimated comparably accurately by OBS and 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. Interestingly, model performance suffered for low d, which is associated with low pro-

cess error and strong Beverton-Holt density-regulation (Fig. S5). We assume that the two non-linear density-effects cannot be distinguished accurately by any of the statistical models, and that this five parameter model is practically not identifiable.

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 (Fig. S6). Without control data, results were similar albeit slightly less accurate. Especially PROC performed worse under the presence of observation error (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 generally were high and just slightly lower than in the coexistence scenario. OBS and SSM produced marginally better estimates than the PROC approach. Estimation errors grew with observation error (Fig. S7), especially for interspecific competition coefficient α_{21} of the out-competed species 1. Without control data however, estimation accuracy decreased in total (Fig. S10), partially 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 suffered, and most strikingly for PROC.

We tested a **predator-prey** model (Eq. 6a, 6b) with cyclic dynamics. The parameters were chosen such that the system experienced approximately two to three full cycles in the observed time (depending on the level of process error). 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 underestimation of predator's model parameters for SSM. Here, the posterior distributions of the individual fittings showed correlations in these parameters. PROC generally featured biased estimates and was highly sensitive to observation error (Fig. S8). Second, we fitted the model without any control data (Fig. S11). While accuracy decreased for parameters estimated with OBS and PROC, SSM results still were comparable to the estimates with control data when observation error was not too high.

Additional scenarios

The OBS and the SSM fitting approach performed quite 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 replicates 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 joint average growth rate) correctly. We found that this generally holds even with fewer replicates, but results were slightly less accurate regarding bias (Fig. S12). However, fewer observations also meant higher posterior uncertainty and therefore higher RMSE, especially for α_{21}, PROC performed worst here. 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 five to seven full cycles. 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 regularity of the process (constant cyclic frequency) was not seriously disturbed, and that it would only decrease with longer time series, leading to less accurate results especially for the OBS model. Fitting longer time series (Fig. S13), results for OBS were indeed slightly less precise than for the shorter time series, while SSM results did not change. We conclude that, while OBS is still quite robust against the level of process variation and the resulting irregularity tested here, SSM should be used for longer time series with irregular cycles.

Finally, we simulated experiments for **logistic growth with a lower carrying capacity** ($K = 10^3$ instead of $K = 10^4$). Generally, process and observation error scale with square root of population abundance as verified above ($\sigma \sim \sqrt{N}$), therefore relative errors decrease with abundance ($\frac{\sigma}{N} \sim \frac{1}{\sqrt{N}}$). As expected, the accuracy of estimated parameters descreased in experiments with lower abundances (Fig. S14), but PROC suffered most.

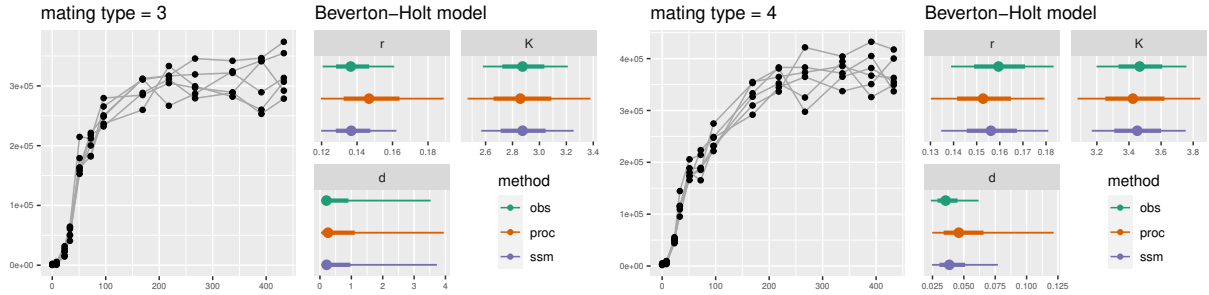

Figure 4: Two out of seven empirical datasets and posterior distributions of fitted Beverton-Holt models, including model parameters r, K, d. For parameter K, x-axis labels show $K/10^5$. Left: d unidentifiable indicating logistic growth, right: small d identifiable.

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, see Fig. S16 for full dataset and Figs. S17–S23 for posterior predictive checks). Since all observed densities were positive and non-integer, we used lognormally distributed residuals for process and observation errors. Note that densities are non-integer here because the data was collected using video-recording and analysis and densities are calculated as averages over the observation time of one video. 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.

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 in the context of deterministic prediction models. It is important to keep in mind that our work is motivated by data from experimental laboratory systems, that is, systems where environmental stochasticity is typically low.

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 if d is small and data is not logistic, Lotka-Volterra competition model) or when the biological models lead to regular cycles, such as in some predator-prey dynamics. With additional single-species time series as control data in the predator-prey systems, OBS produced even better estimates than SSM, while the opposite was found when no control data was used. For data featuring irregular cycles (e.g., in longer time series), we also endorse the use of SSMs. The good performance of OBS for predator-prey models is in line with recent studies (Rosenbaum et al., 2019; DeLong and Lyon, 2020). Also, Barraquand and Gimenez (2021) demonstrated with a discrete-time stochastic model, that parameters can be more easily estimated from (noisy) limit cycles than from time series converging to a steady state in general.

Therefore, OBS may be a viable alternative to SSM, especially if SSMs suffer, for example, from extensive model complexity or convergence problems (Bolker, 2008). Auger-Méthé et al. (2016) discuss identifiability issues with linear SSMs, especially if measurement error is larger than process error, or that σ_{obs} and σ_{proc} often cannot be estimated accurately (e.g. Knape, 2008). Here, we consider these as nuisance parameters and focus on model parameters. While we observe a decrease of model performance with measurement error for all three models (Figs. S2–S13), this does not generally reach a level we would consider as unidentifiable for SSMs. Contrary to the studies cited above, which investigate stationary time series, our datasets feature transient or cycling dynamics, containing more information on the nonlinear processes parameters.

We modelled process variation in PROC and SSM by assuming a piecewise deterministic process between sampling times t_i and t_{i+1} (ODE integration), where error is introduced in t_{i+1} (Auger-Méthé et al., 2021). A more elaborate treatment of process error, that potentially could increase estimation accuracy, requires either an analytical expression of error propagation in the stochastic population process to compute the likelihood, or stochastic simulations during model fitting for a likelihood approximation. While the former is often not available, the latter requires computationally even more demanding methods like sequential Monte Carlo (SMC) or Approximate Bayesian Computing (ABC) (Hartig et al., 2011). While we acknowledge that SSMs classically use piecewise (log-)linear predictions models instead of ODEs, we here aimed at comparing three ways of statistically treating error structure and used identical continuous prediction models for each. How linearized ODEs (Euler's method) and temporal resolution (Clark and Bjørnstad, 2004) affect model performance was not tested here.

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 for predator-prey systems. Biological settings that lead to the exclusion of one species, such as some of the Lotka-Volterra competition models we have used here, make estimates imprecise without such control data, unless 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) can 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 statistical 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 error models and state-space models outperform process error models for data that one may expect to be collected from experimental laboratory populations. Importantly, our continuous-time models allow us 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 under the conditions specified above, deterministic models seem to be sufficient to describe the stochastic dynamics emerging from process and observation errors.

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 au-

thors contributed to 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 91(4): e01470.

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.

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

Chesson, P. 2000. Mechanisms of maintenance of species diversity. – Annual Review of Ecology, Evolution, and Systematics 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 in Ecology and Evolution 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.

Fronhofer, E. A. and Altermatt, F. 2015. Eco-evolutionary feedbacks during experimental range expansions. – Nature Communications 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 .

Gillespie, D. T. 2007. Stochastic Simulation of Chemical Kinetics. – Annu. Rev. Phys. Chem. 58: 35–55.

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.

Hartig, F., Calabrese, J. M., Reineking, B., Wiegand, T. and Huth, A. 2011. Statistical inference for stochastic simulation models–theory and application. – Ecology Letters 14(8): 816–827.

Hefley, T. J., Tyre, A. J. and Blankenship, E. E. 2013. Fitting population growth models in the presence of measurement and detection error. – Ecological Modelling 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. – Ecology Letters 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. – The Canadian Entomologist 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 33(9): 676–688.

Knape, J. 2008. Estimability of density dependence in models of time series data. – Ecology 89(11): 2994–3000.

Lunn, D., Goudie, R. J. B., Wei, C., Kaltz, O. and Restif, O. 2013. Modelling the Dynamics of an Experimental Host-Pathogen Microcosm within a Hierarchical Bayesian Framework. – PLOS ONE 8(8): 1–15.

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

Mühlbauer, L. K., Schulze, M., Harpole, W. S. and Clark, A. T. 2020. gauseR: Simple methods for fitting Lotka-Volterra models describing Gause's "Struggle for Existence". – Ecology and Evolution 10(23): 13275–13283.

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. – Ecology Letters 24(4): 739–750.

O'Hara, R. B. and Kotze, D. J. 2010. Do not log-transform count data. – Methods in Ecology and Evolution 1(2): 118–122.

Paine, C. E. T., Marthews, T. R., Vogt, D. R., Purves, D., Rees, M., Hector, A. and Turnbull, L. A. 2012. How to fit nonlinear plant growth models and calculate growth rates: an update for ecologists. – Methods in Ecology and Evolution 3(2): 245–256.

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. – Frontiers in Ecology and Evolution 6: 234.

Rosenbaum, B. and Rall, B. C. 2018. Fitting functional responses: Direct parameter estimation by simulating differential equations. – Methods in Ecology and Evolution 9(10): 2076–2090.

Rosenzweig, M. L. and MacArthur, R. H. 1963. Graphical Representation and Stability Conditions of Predator-Prey Interactions. – The American Naturalist 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.

— 2022. Runtime warnings and convergence problems. – URL. https://mc-stan.org/misc/warnings.html. Accessed: 2023-02-06.

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.

Wootton, K. L., Curtsdotter, A., Jonsson, T., Banks, H. T., Bommarco, R., Roslin, T. and Laubmeier, A. N. 2022. Beyond body size-new traits for new heights in trait-based modelling of predator-prey dynamics. – PloS one 17(7): e0251896.

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. – Fisheries Research 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.