

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20

Comparison Between Lotka-Volterra and Multivariate Autoregressive Models of Ecological Interaction Systems

Daniel V. Olivença*, Jacob D. Davis, Eberhard O. Voit

The Wallace H. Coulter Department of Biomedical Engineering, Georgia Institute of Technology and
Emory University, 950 Atlantic Drive, Atlanta, Georgia, 30332-2000, USA

* Corresponding Author; orcid.org/0000-0001-5474-2657

Contact Information:

Tel: +1 (404) 731-1736

Email: dvolivenca3@gatech.edu, jddavis@gatech.edu, eberhard.voit@bme.gatech.edu

Running Headline:

Comparison of Lotka-Volterra and MAR models

21 **Abstract:**

22 1. Lotka-Volterra (LV) and Multivariate Autoregressive (MAR) models are computational frameworks with
23 different mathematical structures that have both been proposed for the same purpose of extracting
24 governing features of dynamic interactions among coexisting populations of different species from
25 observed time series data.

26 2. We systematically compare the feasibility of the two modeling approaches, using four synthetically
27 generated datasets and seven ecological datasets from the literature.

28 3. The overarching result is that LV models outperform MAR models in most cases and are generally
29 superior for representing cases where the dependent variables deviate greatly from their steady states. A
30 large dynamic range is particularly prevalent when the populations are highly abundant, change
31 considerably over time, and exhibit a large signal-to-noise ratio. By contrast, MAR models are better suited
32 for analyses of populations with low abundances and for investigations where the quantification of noise
33 is important.

34 4. We conclude that the choice of either one or the other modeling framework should be guided by the
35 specific goals of the analysis and the dynamic features of the data.

36

37 **Key words:**

38 Algebraic Lotka-Volterra Inference (ALVI), Lotka-Volterra models, Multivariate Autoregressive (MAR)
39 models, community dynamics, parameter estimation, population dynamics, structure inference, systems
40 biology.

41

42 **Availability of algorithms used:**

43 <https://github.com/LBSA-VoitLab/Comparison-Between-LV-and-MAR-Models-of-Ecological-Interaction->

44 Systems

45 1. Introduction

46 The growth of populations has been a topic of human interest since prehistoric times: Babylonian clay
47 tablets documented exponential growth in cuneiform lettering as early as about 4,000 years ago (Sachs &
48 Goetze, 1945; Savageau, 1979). The quantitative representation and analysis of population dynamics is
49 also one of the early roots of mathematical modeling in biology. In the mid-1920s, Alfred Lotka (Lotka,
50 1925) studied periodic increases and decreases in the populations of lynx and hare in Canada, while Vito
51 Volterra (Volterra, 1926) independently analyzed fish catches and the competition among populations in
52 the Adriatic Sea. Since these early days, Lotka-Volterra (LV) models have become a mainstay—and typical
53 default—in computational ecology (May, 2001) .

54 With the discovery of complex microbiomes and their surprisingly strong effects on human health and the
55 environment, the interest in interactions among different species has received renewed attention (Gavin,
56 Pokrovskii, Prentice, & Sobolev, 2006; Stein et al., 2013; Shenhav et al., 2019). As an example, we recently
57 inferred the temporally changing interactions among bacterial communities in different lake
58 environments with over 12,000 Operational Taxonomic Units (OTUs) (Dam et al., 2016; Dam et al., 2020).
59 We chose as our computational framework an LV model, which we augmented with LV equations for
60 environmental variables that affected the OTUs (see also (Stein et al., 2013)). Our rationale for this choice
61 was a combination of (1) the successful history of LV models, (2) their mathematical simplicity and
62 tractability and (3) the important fact that parameter values (and thus signs and strengths of interactions)
63 can be obtained from time series data of OTU abundances with methods of linear regression (Voit & Chou,
64 2010).

65 Multivariate Autoregressive (MAR) models were proposed a few decades ago as a viable alternative to LV
66 models. Originally proposed for problems in economics (Sims, 1980), Ives suggested their use for
67 predicting responses of populations to environmental changes (Ives, 1995). His specific motivation was

68 to establish techniques for studying how population abundances change in response to long-term
69 environmental trends and for partitioning different factors driving changes in mean population densities
70 in response to these trends. Since this early work, MAR models have been chosen to represent the
71 interaction dynamics between biotic and abiotic drivers, infer the intra- and interspecific effects of species
72 abundances on population growth rates, identify environmental drivers of community dynamics, predict
73 the fate of communities submitted to environmental changes and extract measures of community
74 stability and resilience. The latter was initially applied to lake and marine systems and later in terrestrial
75 ecology (Certain, Barraquand, & Gårdmark, 2018).

76 Thus, two modeling frameworks with entirely different structures have been proposed for essentially the
77 same purpose of extracting key features of dynamic interactions among coexisting populations of
78 different species from observed time series data. Both methods have had successes, but a direct
79 comparison of the two approaches has not been reported. Such a comparison is the subject of this article.

80 LV are ODE models, whereas MAR are statistical models. The former were designed to elucidate the long-
81 term dynamics of interacting populations, whereas the latter were conceived to also describe the
82 stochastic structure of a dataset. Our focus for their comparison is the ability of each model structure to
83 produce an acceptable fit to the available data and to capture the process dynamics underlying the
84 observed trends in population abundances.

85 We use four versions of MAR: MAR without any data transformation, MAR with log transformation, MAR
86 upon data smoothing and MAR with log transformation upon data smoothing. Log transformation is
87 necessary for comparing the general mathematical interpretation of a MAR model with a typical ecological
88 interpretation, where they can be viewed as multispecies competition models with Gompertz density
89 dependence (Ives, 1995; Certain et al., 2018) (Section 1.3 of the *Supplements*). Data smoothing is explored
90 to assess if the advantages of LV models are in fact due to this preparatory step. It is clear that data

91 smoothing will impede the ability of MARs to describe stochastic structures in the data, but this aspect is
92 not the focus of this study.

93 We begin with a description and comparison of the main features of LV models and MAR models,
94 subsequently analyze small synthetic systems, which offer the advantage of simplicity and full knowledge
95 of all model features, and then assess several real-world systems. It is quite evident that it is impossible
96 to compare distinct mathematical approaches with absolute objectivity and without bias (Rykiel, 1996),
97 and it sometimes happens that inferior choices of models in specific cases outperform otherwise superior
98 choices. We will attempt to counteract these vagaries by selecting case studies we consider representative
99 and by stating positive and negative facts and features as objectively as possible.

100

101 **2. Materials and Methods**

102 **2.1. Lotka-Volterra models**

103 Lotka-Volterra (LV) models (Lotka, 1925; Volterra, 1926) are systems of first-order ordinary differential
104 equations (ODEs) of the format

$$105 \quad \frac{dX_{i,t}}{dt} = a_i X_{i,t} + \sum_{j=1}^n b_{ij} X_{i,t} X_{j,t}, \quad i = 1, 2, \dots, n. \quad \text{eqn 1}$$

106 The left side of equation 1 represents the change in species X_i with respect to time. The equation with
107 only the first term on the right side, $a_i X_{i,t}$, yields exponential growth, while the sum captures interactions
108 between pairs of populations. Most of these terms represent interactions between different species, such
109 as predation, competition for the same resources or cooperation, but one term in each equation, $b_{ii} X_{i,t} X_{i,t}$
110 , accounts for interactions among the members of the same species and is sometimes interpreted as
111 “crowding effect.” Background and further details regarding these models are presented in *Supplements*

112 Section 1.1. Because ODEs are natural representations of dynamic processes, the index t is usually
113 omitted.

114

115 **2.2. Estimation of LV Parameters Based on Slopes of Time Courses**

116 Any of the numerous generic parameter estimation approaches for systems of nonlinear ODEs may be
117 used to estimate the parameter values of LV systems; reviews include (Mendes & Kell, 1998; Wedelin &
118 Gennemark, 2007; Chou & Voit, 2009). Here, we use a combination of smoothing, slope estimation, and
119 parameter inference, for which we use the recently introduced Algebraic Lotka-Volterra Inference (ALVI)
120 method (Voit *et al.*, 2021). We begin by smoothing the raw time series data in order to reduce noise in
121 the data as well as in their slopes, where the effects of noise are known to be exacerbated (Knowles & Renka,
122 2014). Many options are available, but smoothing splines and local regression methods are particularly
123 useful (Cleveland, 1981); they are reviewed in *Supplements* Section 1.2.1. Splines have degrees of freedom
124 and we will refer to a spline with, say, 8 degrees of freedom as “8DF-spline”.

125 The estimation of slopes allows us to convert the inference problem from one involving ODEs into one
126 exclusively using algebraic functions. This conversion is accomplished by substituting the left side of
127 equation 1 with estimated slopes that correspond to values of the dependent variables on the right side,
128 which leaves the parameters as the only unknowns (Voit & Savageau, 1982; Varah, 1982; Voit & Almeida,
129 2004; see also *Supplements* Sections 1.2.2. and 1.2.3.).

130 After the differentials are replaced with estimated slopes, two options permit the inference of the
131 parameter values of LV-models. We can apply simple multivariate linear regression (ALVI-LR), where we
132 either use all datapoints or iterate the regression several times with subsets of points, which is a natural
133 approach of creating ensembles of solutions. As an alternative, if n is the number of variables, one may

134 use $n+1$ of the datapoints and slopes, which results in a system of linear equations that can be solved with
135 simple algebraic matrix inversion (ALVI-MI). For a thorough description of the ALVI method please see
136 (Voit et. al., 2021) and an example in *Supplements* Section 1.2.4.

137

138 **2.3. Multivariate Autoregressive (MAR) models**

139 In contrast to the ODEs of the LV format, Multivariate Autoregressive (MAR) models are discrete recursive
140 linear models. They have the general format

$$141 \quad X_{i,t+1} = \alpha_i + \sum_{j=1}^n \beta_{ij} X_{j,t} + \sum_{g=1}^m \gamma_{ig} u_{g,t} + w_{i,t}; i = 1, 2, \dots, n; w_{i,t} \sim N(0, \delta_i) \quad , \quad \text{eqn 2}$$

142 where $u_{g,t}$ are environmental variables and $w_{i,t}$ represents normally distributed noise. This set of
143 equations, for different i , is usual represented in the matrix form

$$144 \quad \mathbf{X}_{t+1} = \boldsymbol{\alpha} + \boldsymbol{\beta} \mathbf{X}_t + \boldsymbol{\gamma} \mathbf{u}_t + \mathbf{w}_t, \quad \mathbf{w}_t \sim MVN(0, \boldsymbol{\delta}). \quad \text{eqn 3}$$

145 Explained in words, the “state” of the system at time $t+1$, expressed by the vector \mathbf{X}_{t+1} , depends
146 exclusively on the state of system one time unit earlier, \mathbf{X}_t , as well as on external inputs and stochastic
147 environmental effects. Furthermore, $\boldsymbol{\alpha}$ is the vector of intercepts and $\boldsymbol{\beta}$ is one row of the population
148 interaction matrix. The term $\boldsymbol{\gamma} \mathbf{u}_t$ describes how cofactors affect the dependent variables. Specifically, \mathbf{u}_t
149 is a vector of external variables and $\boldsymbol{\gamma}$ is the vector of weights associated with these external variables.
150 Finally, the term \mathbf{w}_t is a vector representing stochastic noise affecting the dependent variables.
151 Background and further details regarding these models are presented in *Supplements* Section 1.3.

152

153 **2.4. Parameter estimation for MAR**

154 Software packages for the estimation of MAR model parameters greatly facilitate the use of these models.
155 An example is the package MARSS, which uses an expectation maximization algorithm (Holmes, Ward, &
156 Wills, 2012; Holmes, Ward, & Scheuerell, 2020). Some details of MARSS are discussed in in *Supplements*
157 Section 1.4. and in the next section.

158

159 **2.5. Structural Similarities between Modeling Formats**

160 Although LV and MAR models have both been proposed for characterizing the interactions among
161 populations within a mixed community, they are distinctly different in structure and appearance.
162 Nonetheless, they also exhibit fundamental mathematical similarities, which are sketched below; a
163 detailed analysis is presented in *Supplements* Section 1.5.

164 To assess these similarities, we focus on models without environmental factors, and thus on

$$165 \quad X_{i,t+1} = \alpha_i + \sum_{j=1}^n \beta_{ij} X_{j,t} + w_{i,t}; \quad i = 1, 2, \dots, n; \quad w_{i,n} \sim N(0, \delta_i). \quad \text{eqn 4}$$

166 We also suppose that the MAR variables represent the logarithms of abundances, as proposed in
167 (Dennis & Taper, 1994; Ives, 1995; Certain *et al.*, 2018). Borrowing the principles of solving ODEs with
168 Euler's method, we discretize the LV model, which yields

$$169 \quad X_{i,t+h} = X_{i,t} + h \cdot \frac{dX_i}{dt} \Big|_{X_i=X_{i,t}} = X_{i,t} + h \cdot X_{i,t} (a_i + \sum_{j=1}^n b_{ij} X_{j,t}), \quad i = 1, 2, \dots, n. \quad \text{eqn 5}$$

170 If the dynamics remains close to the steady state, then $X_{i,t+1} - X_{i,t} \approx 0$ for any given t . Substituting this
171 approximation into equations 4 and 5 yields

$$172 \quad \alpha_i + \sum_{j=1}^n \beta_{ij} X_{j,t} + w_{i,t} - X_{i,t} \approx 0 \quad \text{eqn 6}$$

173 and

174
$$a_i + \sum_{j=1}^n \tilde{b}_{ij} X_{j,t} - X_{i,t} \approx 0, \quad \text{eqn 7}$$

175 respectively. Ignoring the noise term in the MAR model, the two sets of near-steady-state equations 6 and 7 are
176 the same if $\alpha_i = a_i$, $\beta_{ij} = b_{ij}$ for all $i \neq j$ and $\beta_{ij} = b_{ij} + 1$ for $i = j$. Thus, the MAR and LV models are mathematically
177 equivalent at the steady state and similar close to it.

178

179 **3. Results**

180 The comparison between LV and MAR models may be executed in two ways. A purely mathematical
181 approach was sketched in Section 2.5. An alternative approach focuses on practical considerations and
182 actual results of inferences from data. It is described in the following.

183 For simplicity, we omit environmental inputs ($c_i X_i U_t$ and $\gamma_i u_t$, respectively) and begin by testing several
184 synthetic datasets with different dynamics. We suppose that these data are moderately sparse and noisy,
185 to mimic reality. In particular, we test whether the LV inference from synthetic LV data returns the correct
186 interaction parameters and whether the MAR inference from synthetic MAR data does the same.
187 Subsequently, we test to what degree LV inferences from MAR data yield reasonable results and *vice*
188 *versa*. Finally, we apply the inference methods to real data. As the main metric, we compare the sums of
189 squared errors and use a Wilcoxon rank test to assess the significance of the difference.

190

191 **3.1. Case study 1: Synthetic LV data**

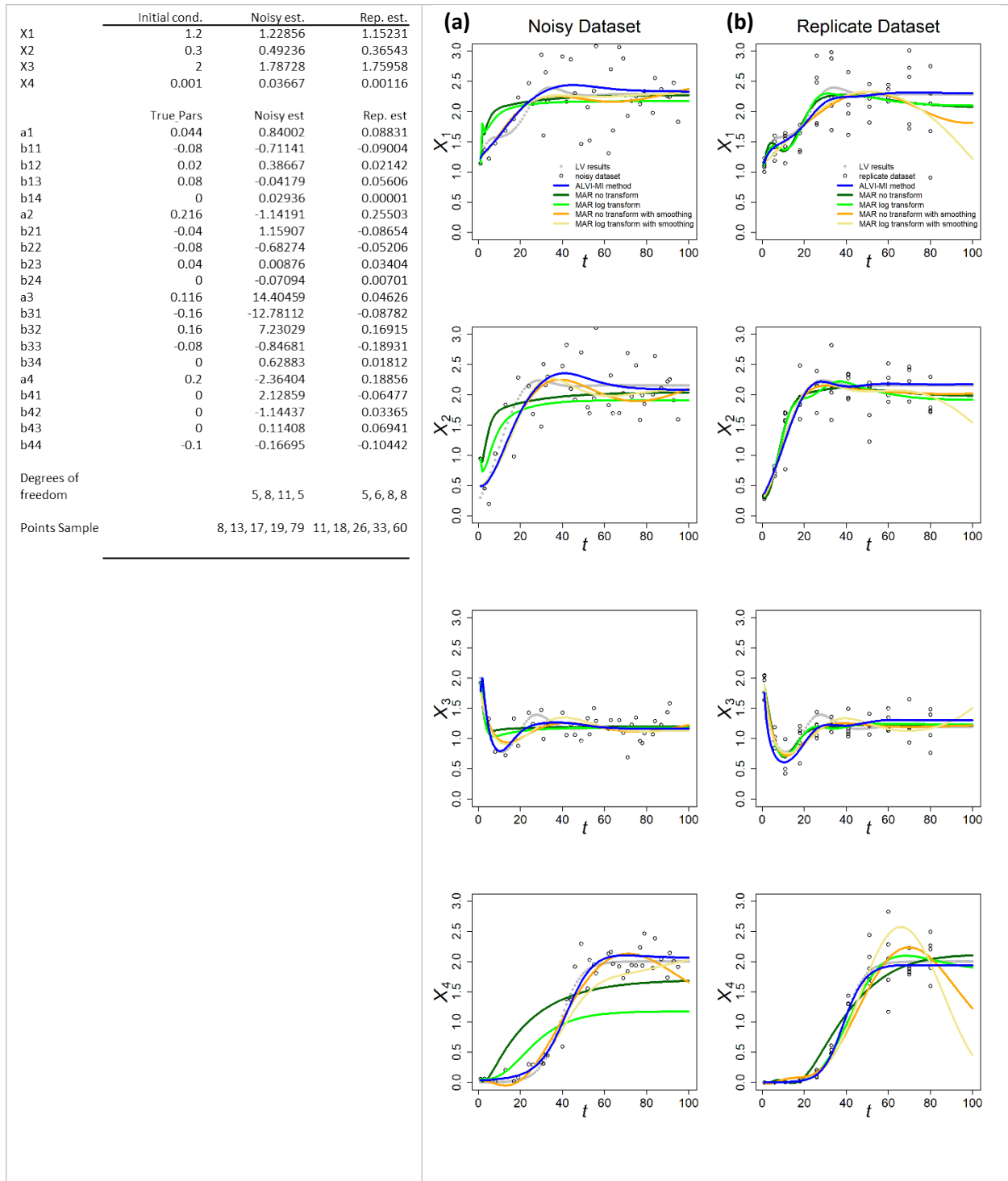
192 The first case consists of data that were generated with a four-variable LV model and superimposed with
193 synthetic, normally distributed noise (for details, see *Supplements* Section 2). We also generate a smaller

194 noisy dataset, which however comes with replicates. The specific question we address is whether the LV
195 and MAR inference methods return the true dynamics and parameter values.

196 The fits for the noisy and replicate LV datasets are presented in Figures 1 a, b, along with parameter
197 estimates. These generally possess the correct sign and could, if deemed beneficial, serve as the starting
198 point for an additional, refining optimization, for instance with a steepest-descent method. The inferred
199 and true values are quite similar for the replicate dataset. By contrast, the parameter values inferred for
200 the noisy dataset do not exactly recoup the true values. In fact, these inferences yield slightly better fits
201 to the noisy data (SSE=17.981) than the “true” values (SSE=18.197), due to the noise. Because we usually
202 obtain better results through ALVI-MI, we display those results here and ALVI-LR results in the
203 *Supplements*.

204 Figure 1 also displays the MARSS estimates with and without log-transformation of the data and with or
205 without data smoothing. With respect to X_1 , X_2 and X_3 , these estimates are of adequate quality. They
206 present good fits, although not as good as the LV inference, which is probably not surprising as the data
207 were generated with an LV system. MARSS did not perform well for the “detached” variable X_4 , especially
208 for the noisy dataset.

209



210 **Figure 1: ALVI-MI and MARSS methods applied to noisy (a) and replicate (b) LV datasets.** Original synthetic data
 211 are shown as gray dots and data with added noise as black circles. ALVI results are presented in blue. True
 212 parameters and ALVI-MI estimates are presented in the Table. MAR estimates are presented in green, orange
 213 and yellow. Data and parameter estimates for MAR can be seen in Table S1. SSEs for all fits are presented in Table
 214 1 toward the end of the article.

215

216 Because MARSS yields parameter values for a discrete recursive system, they are not directly comparable
217 to the true parameters of a LV system; nonetheless, their numerical values are recorded for completeness
218 in Tables S1.4 and S1.5.

219 For MARSS inferences from the replicate dataset we had to average points with the same time value.
220 MARSS did not converge for all parameters but it still presented a relatively good fit. Additional details
221 are presented in *Supplements* Section 2.

222 The ALVI-MI method also works well for more complicated dynamics, as demonstrated in *Supplements*
223 Section 2 and Figure S5.

224

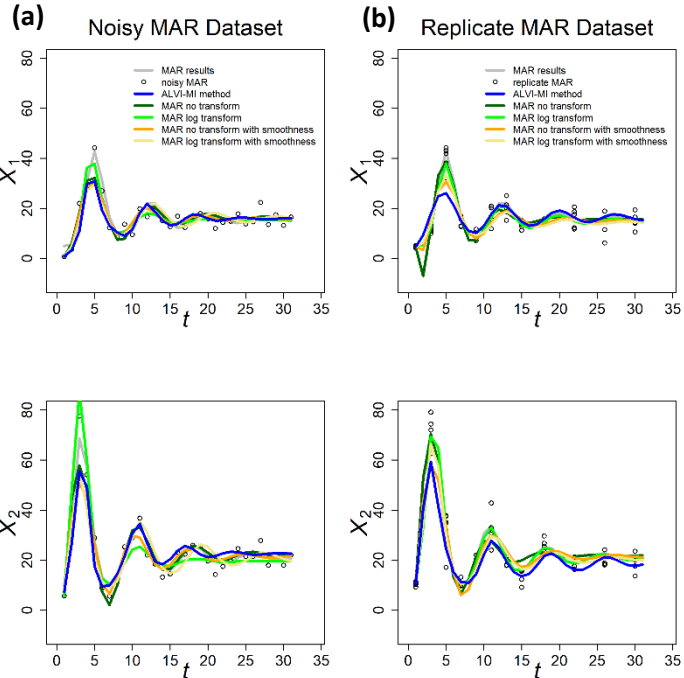
225 **3.2. Case study 2: Synthetic MAR data**

226 Here we reverse the set-up of Case Study 1 by creating synthetic data with an MAR model and test
227 whether either method can infer results corresponding to the original system.

228 As a representative example, we use a four-variable MAR system to create 31 synthetic datapoints. We
229 create a *noisy MAR* dataset by adding noise in the form of a random normal variable of mean zero and a
230 standard deviation of 20% of each variable's mean to the original 31 points sample. We also create a
231 *replicate MAR* dataset by choosing 12 points and generating 5 replicates by multiplying their values by a
232 random normal of mean 1 and standard deviation of 0.2. The initial conditions, parameters, dynamics,
233 ALVI and MAR fits are presented in Figure 2. All fits to the synthetic MAR data are satisfactory. The LV
234 parameters are not directly comparable to the MAR parameters.

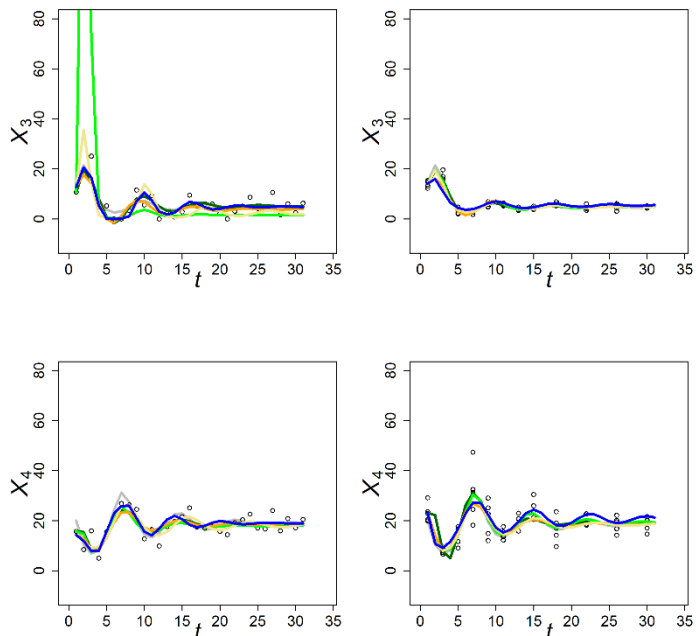
Noisy MAR

Pars	True Values	MAR	MAR log transform	MAR with data smoothing	MAR log transform with smoothing
β_{11}	0.45	0.1326	0.31343	0.39642	0.3582
β_{21}	-0.5	-0.7579	-0.71839	-0.56989	-0.72816
β_{31}	-0.2	-0.3014	-2.56121	-0.04234	-1.47731
β_{41}	0.3	0.0655	0.41855	0.32471	0.3857
β_{12}	0.5	0.4134	0.21377	0.39737	0.29638
β_{22}	0.544635	0.0613	0.97497	0.3656	1.35461
β_{32}	-0.3	-0.2501	2.01907	-0.2907	1.40072
β_{42}	-0.1	-0.0694	-0.75818	-0.14531	-0.59619
β_{13}	-0.1	-0.1515	0.00937	0.02061	0.01489
β_{23}	0.3	1.623	-0.02319	1.21771	-0.04951
β_{33}	0.9	0.76	-0.1124	0.96983	0.44232
β_{43}	-0.3	-0.4955	0.0407	-0.03142	0.04829
β_{14}	-0.3	-0.4951	-0.58434	-0.00845	-0.3693
β_{24}	0.1	-0.9279	0.28317	-0.09393	0.8837
β_{34}	0.3	-0.4852	1.51221	-0.10254	3.90492
β_{44}	0.3	0.2066	-0.3689	0.61166	0.15671
α_1	2	16.7985	2.91985	1.15325	0.09526
α_2	1.02106	42.1023	1.22855	19.75135	0.03591
α_3	0.740671	20.7671	-2.84687	8.94758	-0.0315
α_4	2	17.8742	5.04906	5.29832	0.00908
δ_1	3.320835	9.5406	0.03335	1.17525	0.00537
δ_2	4.771945	20.564	0.03175	7.84299	0.00869
δ_3	1.214544	8.3144	4.2181	3.09948	0.6299
δ_4	3.668288	9.9207	0.05057	1.71729	0.00224



Replicate MAR

Pars	True Values	MAR	MAR log transform	MAR with data smoothing	MAR log transform with smoothing
β_{11}	0.45	0.1313	0.7550	0.4045	0.8751
β_{21}	-0.5	-0.1997	-1.2900	-0.4620	-0.7421
β_{31}	-0.2	0.0452	-0.9240	-0.0046	-0.3758
β_{41}	0.3	0.5716	0.7020	0.3775	0.4040
β_{12}	0.5	0.6439	1.4000	0.3926	0.8919
β_{22}	0.544635	-0.0389	-0.2100	0.0853	0.4392
β_{32}	-0.3	-0.2161	-1.4000	-0.1873	-0.4188
β_{42}	-0.1	-0.2339	0.3570	0.0521	0.1148
β_{13}	-0.1	-1.4212	0.4890	-0.2147	0.5447
β_{23}	0.3	3.2295	-0.7320	2.0833	0.0302
β_{33}	0.9	1.3301	-0.1770	0.9647	0.6010
β_{43}	-0.3	0.5027	0.2900	-0.2409	-0.0582
β_{14}	-0.3	-0.1848	1.6900	0.1292	1.2788
β_{24}	0.1	-0.2463	-2.2500	-1.0586	-0.6291
β_{34}	0.3	-0.0189	-2.4100	-0.1712	-0.3419
β_{44}	0.3	0.5529	1.6100	0.8786	0.9728
α_1	1.02106	-0.2209	0.0437	0.2503	0.0305
α_2	2	1.8693	0.0597	0.3678	0.0240
α_3	0.740671	-0.1895	-0.0261	-0.2918	-0.0320
α_4	2	0.0909	-0.0257	-0.1392	-0.0073
δ_1	3.320835	0.0029	0.0000	3.2610	0.0060
δ_2	4.771945	3.7408	0.0010	10.8861	0.0108
δ_3	1.214544	0.0214	0.0018	0.6440	0.0123
δ_4	3.668288	0.8913	0.0000	1.9498	0.0038



235 **Figure 2: MARSS and ALVI-MI methods applied to noisy (a) and replicate (b) MAR datasets.** Original synthetic
 236 data are shown as gray dots, data with added noise as black circles. MAR estimates are presented in green,
 237 orange and yellow. ALVI results are in blue. The variables of the noisy dataset were smoothed with 15DF-splines
 238 and the ALVI-MI solution was calculated with spline points at times 2, 6, 15, 18 and 26. The variables of the replicate
 239 dataset were smoothed with 15DF-splines and the ALVI-MI solution was calculated with spline points corresponding

240 to times 1, 3, 11, 13 and 15. Data and parameter estimates for ALVI can be seen in *Supplements* Table S5. SSEs for
241 these fits are presented in Table 1 toward the end of the article.

242

243 In most cases, the different MAR models had difficulties retrieving the true parameters of the system, and
244 sometimes even the correct sign (Figure 2). This is probably due to the small number of datapoints: Certain
245 *et al.* (Certain et al., 2018) suggest that the length of the time series should be at least 5 times greater
246 than the number of *a priori* nonzero elements in the matrix B in order to recover interaction signs
247 correctly. Our sample has 31 observations and should have at least 80. For X_1 and X_4 , all models show a
248 similar fit, but not for X_2 and X_3 , where MAR models with log transformation show a considerable deviation
249 from the data.

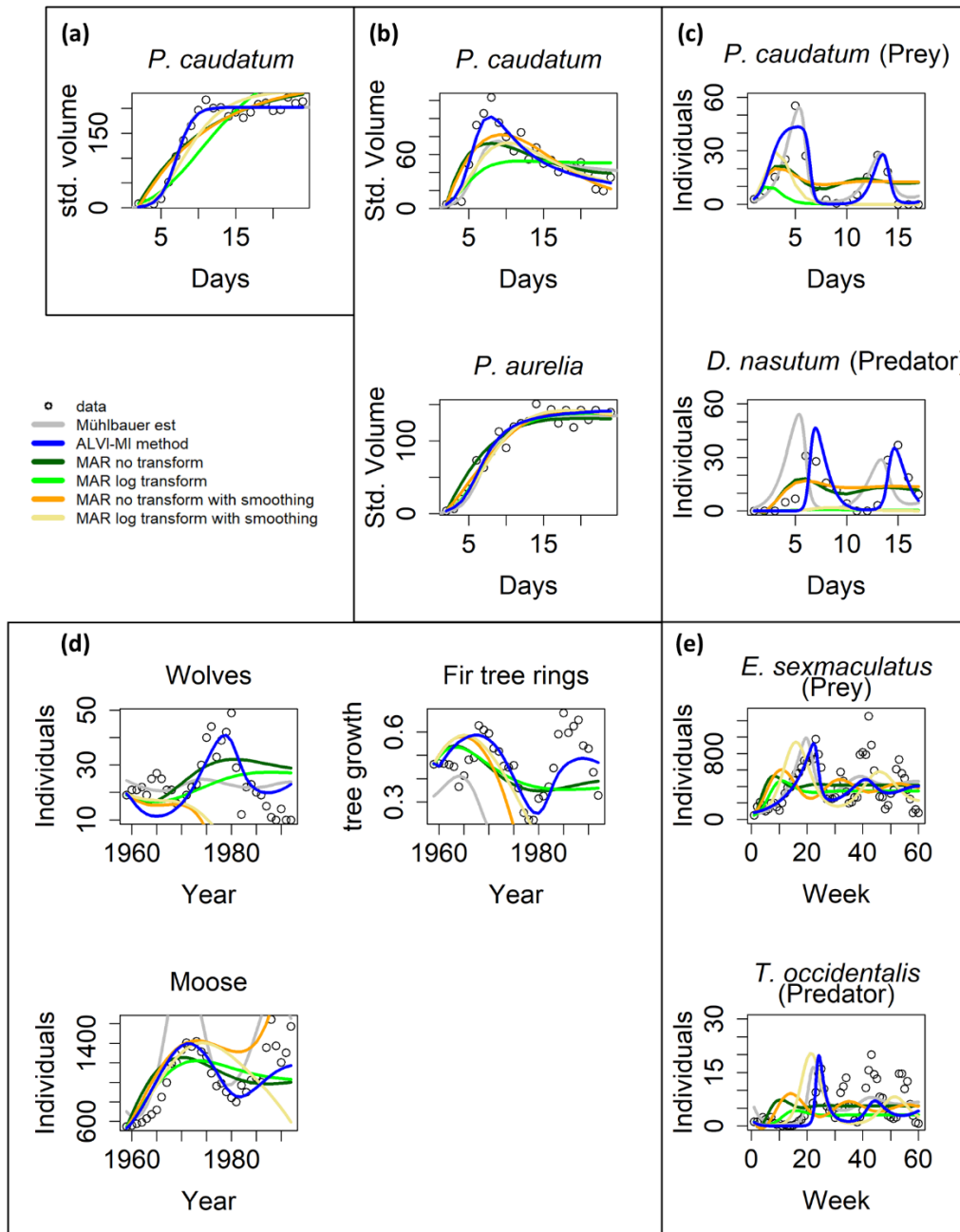
250

251 **3.3. Case study 3: Experimental data from the literature, previously used for inferences with LV and** 252 **MAR models**

253 **3.3.1. Published LV inferences**

254 Data from Georgy Gause's 1930s experiments and others were recently compiled in the R package *gauserR*
255 (Mühlbauer *et al.* 2020). In the accompanying paper, the authors present five examples to test their
256 method for estimating LV model parameters. We use the exact same examples to demonstrate to what
257 degree LV and MAR methods are compatible with these real-world data and compare our results to those
258 presented by Mühlbauer and colleagues. For more information regarding the original experimental data,
259 see (Gause, 1934), (Huffaker, Shea, & Herman, 1963), (McLaren & Peterson, 1994) and (Mühlbauer et al.,
260 2020). The results are presented in Figure 3, with data as symbols and various estimates as lines. SSEs of
261 the different estimates for these and other test examples are presented in Table 1.

262 The MAR method never outperforms the other methods considered. To be fair, these examples had been
 263 used to test actual data for compatibility with the LV structure, which may explain the superior
 264 performance of the LV model. Nonetheless, these are the types of data the MAR method is supposed to
 265 capture.



266

267 **Figure 3: Model inferences associated with Gause's data (Gause, 1934).** Circles show observations, gray lines are
268 the estimates from Mühlbauer *et al.* (Mühlbauer *et al.*, 2020). ALVI-MI method estimates are presented as blue lines
269 and MAR estimates as green, orange and yellow lines. See text and *Supplements* Tables S4.1 and S4.3 for further
270 details.

271

272 For the case in Figure 3a, ALVI-MI yields the same results as found in (Mühlbauer *et al.*, 2020). In contrast,
273 the MAR estimates are poor, with a very high estimate for the noise (Table S4.3), especially if one does
274 not use log-abundances; this problem occurs in all cases presented in Figure 3. The data in Figure 3a are
275 close to a logistic function, similar to X_4 in the previous noisy dataset, where MAR also did not perform
276 well.

277 Figure 3b shows data from a competition experiment between *Paramecium caudatum* and *Paramecium*
278 *aurelia* that were co-cultured. Estimates for *P. aurelia* are similar for all methods but ALVI-MI exhibits
279 clear superiority for *P. caudatum*.

280 The data in Figure 3c are complicated. Mühlbauer and colleagues noted that additional quantities of
281 bacteria were introduced to avoid species extinction. Furthermore, many datapoints in this dataset are
282 zero, which causes problems for the parameter estimators. As a remedy, we changed the zeros to 10^{-5} ,
283 but our initial estimates still produced poor fits. However, if we use the estimated trajectories from
284 Mühlbauer *et al.* as “data,” quasi as a diagnostic measure, ALVI-MI captures the parameters that
285 reproduce the fit of Mühlbauer *et al.*. This finding suggests that the initial poor fit is not a problem of LV
286 adequacy. Instead, we hypothesize that the problem was caused by insufficient datapoints or almost-
287 linear dependence, which affects the matrix inversion. To test this hypothesis, we used the first splines as
288 data to create a second set of splines that has more datapoints to create the subsample to be used on
289 ALVI-MI. We were able to achieve the presented fit, which is still somewhat inferior to the one by
290 Mühlbauer *et al.*, but a considerable improvement over our initial fits.

291 When calculating splines for this dataset, it is difficult to choose degrees of freedom that capture both
292 maxima. High degrees of freedom capture the global maxima but overshoot the local maxima. Low
293 degrees of freedom capture the local but undershoot the global maxima. We suspect this to be the cause
294 of ALVI's initial poor performance. Still, ALVI yields better fits than MAR.

295 The data in Figure 3d are also complicated, in this case due to two aspects. First, they show a stark
296 difference in absolute numbers, with the abundance values for moose being several magnitudes higher
297 than the numbers of tree rings. As a potential remedy, we normalized the fitting error for each dependent
298 variable by dividing it by its mean to balance the SSE. The result is shown in Figure 3d. The LV models
299 perform better than MAR, and MAR with log-abundances produces a better noise estimate than with the
300 untransformed data.

301 The second issue is the fact that, around 1980, the wolves were exposed to a disease introduced by dogs
302 that caused a precipitous drop in the population between 1981 and 1982 (Park Service, 2021). Typical
303 mathematical models are not equipped to simulate such a black swan event, and the totality of results
304 from the various methods suggests that neither LV nor MAR may be good models for this system, because
305 none of the fits, by Mühlbauer *et al.*, ALVI, and MAR, are entirely satisfactory. Nonetheless, our ALVI
306 results present a decent fit for moose and fir tree rings. To improve the fit to the wolf data, we divided
307 the data into two groups, from 1959 to 1980 and from 1983 to the end of the series and estimated
308 parameters for the two intervals. The results are presented in Figure S7 in orange lines. The fit is greatly
309 improved, although still not perfect.

310 Figure 3e describes yet another complicated example. According to the inference, the ALVI-MI estimates
311 fit the first peak well but the oscillations die down, in contrast to the data. Estimates from Mühlbauer *et*
312 *al.* produce even poorer estimates, suggesting that the data may not be compliant with the LV structure.
313 As in the previous example, MAR models do not capture the dynamics, although MAR with log-

314 abundances produces good noise estimates. Surprisingly, MAR with smoothing presented very poor fits
315 to these data.

316 We repeated the analysis using ALVI-LR instead of ALVI-MI. The results were by and large similar and
317 slightly inferior; they are shown in *Supplements* Figure S6 and Table S4.2.

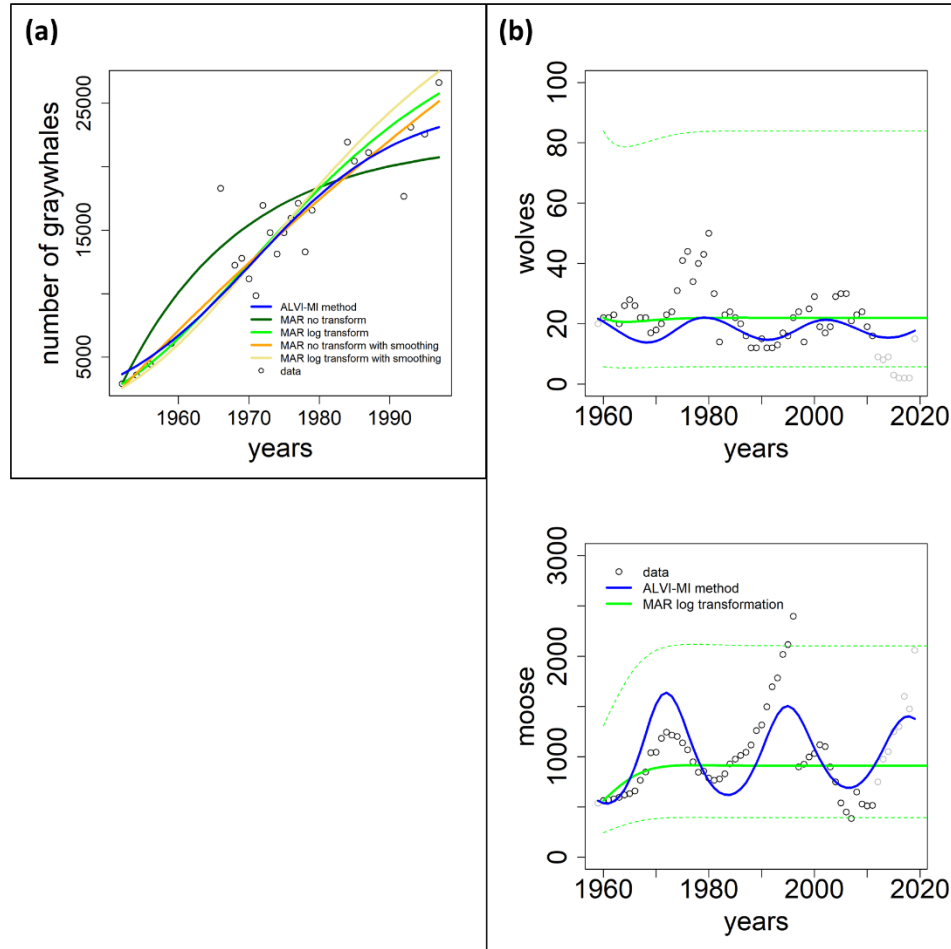
318 One should note that Mühlbauer *et al.* used a steepest-descent method, while our method did not.
319 Therefore, our results could possibly be further improved by adding a refinement cycle of steepest-
320 descent optimization. The main problem of these algorithms, getting stuck in local minima, would
321 presumably not be an issue, since the ALVI results are already close to the optimum.

322

323 **3.3.2. Published MAR inferences**

324 Here we use two datasets presented in the MAR inference package MARSS. The first dataset, “gray
325 whales,” consists of 24 annual abundance estimates of eastern North Pacific gray whales during recovery
326 from intensive commercial whaling prior to 1900 (Gerber, Demaster, & Kareiva, 1999). It is thus to be
327 expected that the whales are initially far from the carrying capacity of the system. The second case
328 consists of data for wolf and moose populations on Isle Royale in Lake Superior between 1960 to 2011;
329 this dataset was used by Holmes and colleagues (Holmes et al., 2020) to demonstrate usage of the MARSS
330 R package.

331



332

333 **Figure 4: Two datasets of wildlife observations.** Column a: Abundance data of gray whales (Gerber et al., 1999). The
334 plot shows results from ALVI-MI in blue; MAR estimates are displayed with green, orange and yellow lines. Column
335 b: Wolves and moose on Isle Royale (Vucetich, 2021). The original data used for parameter estimation are displayed
336 with black circles, data not used by the estimation processes are shown in gray, ALVI-MI results are in blue, MAR
337 estimates using log-abundances are displayed with green lines. The dashed lines indicate confidence intervals for
338 the MAR estimates. Values of the estimates can be seen in Table S6.

339

340 Figure 4a show fits to the gray whale data (Gerber *et al.*, 1999). ALVI-MI noticeably outperforms MAR,
341 even though the data came from a MAR demonstration. In particular, the MAR results suggest that the
342 whales are close to regaining their carrying capacity, which seems to contradict the trend in the data. The
343 SSEs can be seen in Table 1. It is unclear why the MAR method without transformation does not perform
344 better. As it stands, the estimates are inadequate (with the highest SSE) and have a very high variance for

345 the error. An LV model with one variable is a logistic function, and the LV fit represents initial quasi-
346 exponential growth that starts to slow down after a while. This behavior nicely reflects the fact that the
347 whales were recovering from very small numbers due to overfishing but are apparently still much below
348 the carrying capacity.

349 Figure 4b returns to the Isle Royale dataset from Vucetich (Vucetich, 2021), which we already used in the
350 context of examples from the collection of Mühlbauer and colleagues (Mühlbauer et al., 2020); *cf.* Figure
351 3. Holmes *et al.* (Holmes et al., 2020) used only the data of wolves and moose for a MAR analysis but
352 extended them over a longer time horizon. Specifically, eight datapoints were added since the former
353 usage of this dataset by Holmes and colleagues, from 2012 to 2020 (gray symbols in Figure 4b).

354 The results of the MAR model are identical with those published in (Holmes et al., 2020), with the same
355 log transformation and z-scoring of the data, and the same parameter values were inferred. The result
356 consists of acceptable estimates, although we found a slightly better fit without the z-scoring. Still, for a
357 direct comparison, we opted to present the example exactly as Holmes *et al.* did. Interestingly, these fits
358 miss all oscillatory behavior seen in the data. The ALVI results do show oscillations but clearly suffer from
359 the disruption in the wolf population in 1981 and 1982, as discussed before.

360 Because we used in this example only MAR with log transformation, we display the confidence intervals
361 for the MAR model as dashed green lines. Very few datapoints are outside the confidence intervals.

362

363 **3.4. Comparative summary of the performance of LV and MAR in the presented examples**

364 Inspection of Table 1 renders is evident that ALVI clearly performs better than MAR. In a few cases, the
365 ALVI-LR solution gives a better SSE than ALVI-MI, but the difference between the two is not substantial.
366 ALVI-LR appears to be superior when the data are noise-free.

367

368 **Table 1 – Sum of squared errors (SSE) of data fits for all experiments with ALVI-LR, ALVI-MI and four**
 369 **variants of the MAR methods.** We also include SSEs for the estimates obtained by Mühlbauer *et al.* (2020)
 370 for LV data presented in Figure 3. Bold values identify the lowest SSE score for each example. Examples
 371 used in the Wilcoxon rank test are marked with asterisks.

	Shown in	ALVI-LR	ALVI-MI	MAR	MAR logTrans	MAR with smoothing	MAR with log and smoothing	Mühlbauer <i>et al.</i>	Test
Noisy LV dataset	Fig. 1a	9.009	4.362	37.31	53.31	7.364	8.260		*
Replicate LV dataset	Fig. 1b	11.41	2.953	6.247	5.439	13.87	42.83		*
Noisy MAR dataset	Fig. 2a	2,727	1,142	908.3	37,447	1,191	1,604		*
Replicate MAR dataset	Fig. 2b	2,518	1,462	933.1	177.967	1,773	1,482		*
SynthData1	Fig. S5a	3.29E-07	9.04E-07	10.49	4.626				
SynthData2	Fig. S5b	8.68E-06	8.57E-06	3.029	5.876				
SynthData3	Fig. S5c	0.002	0.018	5.111	4.378				
SynthData4	Fig. S5d	5.65E-04	2.10E-02	2.47E+04	3.99E+35				
SynthData5	Fig. S5e	7.048	18.94	16.59	15.70				
SynthData6	Fig. S5f	71.27	28.36	7,379,825	174.2				
Mühlbauer et al. 1	Fig. 3a	2,588	2,516	18,160	51,738	17,877	16,399	3271	*
Mühlbauer et al. 2	Fig. 3b	6,204	4,781	13,016	22,162	9,293	12,447	40604	*
Mühlbauer et al. 3	Fig. 3c	24,183	2,191	4,619	10,146	4,780	8,795	578.1	*
Mühlbauer et al. 4	Fig. 3d	17,700,169	1,118,350	2,199,174	17,77,559	9,904,281	3,973,876	7934136	*
Mühlbauer et al. 5	Fig. 3e	12,114,811	2,251,804	4,245,588	4,450,871	5,596,187	4,554,089	2893764	*
Holmes et al. 1	Fig. 4a	168,134,153	153,703,251	217,478,452	166,186,993	143,728,190	212,424,910		*
Holmes et al. 2	Fig. 4b	32,326,913	5,556,585	10,267,681	10,947,294	11,136,306	14,342,449		*

372

373

374 We used a one-sided Wilcoxon rank test to see if the differences in performance are significant. The
 375 results and alternative hypotheses for these tests are presented in Table 2.

376 **Table 2 - Test results of one-sided Wilcoxon rank.**

Null hypothesis H_0 :	Alternative hypothesis H_1 :	p -value
SSE values of ALVI-MI are equal or higher than corresponding MAR values	SSE values of ALVI-MI are less than corresponding MAR values	0.0093
SSE values of ALVI-MI are equal or higher than corresponding MAR values for log transformed variables	SSE values of ALVI-MI are less than corresponding MAR values for log transformed variables	0.0024
SSE values of ALVI-MI are equal or higher than corresponding MAR values for smoothed data	SSE values of ALVI-MI are less than corresponding MAR values for smoothed data	0.0269
SSE values of ALVI-MI are equal or higher than corresponding MAR values for log transformed variables and smoothed data	SSE values of ALVI-MI are less than corresponding MAR values for log transformed variables and smoothed data	0.0005
SSE values of the ALVI-MI are equal or higher than SSE values obtained with the ALVI-LR	SSE values of ALVI-MI are less than SSE values obtained with the ALVI-LR	0.0005
SSE values of MAR with data smoothing are equal or higher than MAR values without	SSE values of MAR with data smoothing are less than MAR values without	0.7676
SSE values of MAR with log transformation and data smoothing are equal or higher than those for MAR with log transformation values	SSE values of MAR with log transformation and data smoothing are less than those for MAR with log transformation values	0.7935

377

378

379 The data supports that the ALVI-MI method produces smaller SSE's than the other methods considered.

380 Also, in the last two tests, the data do not show evidence that data smoothing reduces the SSE's in MAR.

381 Comparing the results ALVI-LR and ALVI-MI with respect to the absolute value of the difference between

382 true and estimated parameters, we obtained mixed results (Table 3). Indeed, a one-sided Wilcoxon rank

383 test with the alternative hypothesis that the absolute errors in parameter values associated with ALVI-LR

384 were smaller than those associated with ALVI-MI did not yield a significant p -value (0.2783).

385

386 **Table 3 – Absolute differences between true and estimated parameters for ALVI-LR and ALVI-MI.** Bold
387 font indicates the lower difference in each case. A statistical test did not reject the null hypothesis of no
388 difference between the results of the two methods.

	ALVI-LR	ALVI-MI
Noisy Data 5%	1.97538	2.68019
Replicate Data 5%	0.24117	0.32537
Noisy Data 20%	1.35692	46.60082
Replicate Data 20%	0.8835	0.66838
Synthetic Data 1	0.00021	0.00038
Synthetic Data 2	0.20832	0.15534
Synthetic Data 3	1.06975	0.0982
Synthetic Data 4	0.19183	0.60795
Synthetic Data 5	0.82403	0.36212
Synthetic Data 6	0.18794	0.97794

389

390

391 Comparing the values in Tables 1 and 3 for noisy and replicate datasets, the latter presented smaller

392 values. However, using the one-sided Wilcoxon rank test for the values in Table 1 with the alternative

393 hypothesis—that the SSE values for the replicate dataset were smaller than those for the noisy datasets—

394 did not yield a significant p -value (0.2734). For the values in Table 3, the test produced a p -value (0.0625)

395 smaller than 0.1, suggesting that time series data are better than clustered data for parameter estimation
396 in LV using ALVI.

397

398 **4. Discussion**

399 We have compared LV and MAR models using, as parameter estimation strategies, a recently introduced
400 ALVI method for LV models (Voit et al., 2021) and the established MARSS method for MAR models
401 (Holmes et al., 2012). Summary Table 1 renders it evident that ALVI usually outperforms MARSS, with
402 often substantially lower SSE values, and that LV models in the vast majority of cases provide better
403 representations than MAR models. Furthermore, Certain *et al.* (Certain et al., 2018) prescribed the
404 length of the data series as at least 5 times greater than the number of nonzero interaction elements in
405 order to recover interaction signs correctly. A caveat is that we presented the MAR model equations
406 without noise, thereby ignoring an important part of these models that characterizes the stochastic
407 structure of the data.

408 ODEs and MAR models are derived from different philosophies and comparing them fairly is not
409 straightforward. Here we compared the two approaches from a point of view of someone who is more
410 interested in capturing the dynamics of a phenomenon, which poses an immediate disadvantage for MAR
411 models, because they were created for phenomena with random noise and for characterizing the
412 structure of this noise. Consequently, we found much of the dynamics quantified as noise in the estimates.
413 Thus, an overarching conclusion is that the two approaches are different tools that are adequate in
414 different situations. MAR models are well suited for simulations with variables presenting low abundances
415 and investigations where the quantification of noise is relevant. By contrast, LV models, and ODE models
416 in general, are better suited to capture the dynamics of a phenomenon and for cases where the
417 dependent variables have high abundances and a high signal-to-noise ratio. Nonetheless, if the

418 characterization of noise is of interest, one might subtract the LV fit from the raw data and then assessing
419 the remainder. Another advantage of LV models is that ALVI permits *a priori* tests for the adequacy of the
420 LV structure for a given dataset (Voit et al., 2021).

421 Because the LV structure is continuous, we can evaluate it at any point or choose any interval between
422 the points in the numerical solution. By contrast, MAR does not truly reveal a time resolution higher than
423 its intrinsic interval between solution points. However, Holmes *et al.* (Holmes et al., 2012) demonstrated
424 with the MARSS R function that it is feasible to interpolate any number of missing values between the
425 known datapoints, and that this method can be used to decrease the time unit for stepping forward. While
426 this step does not make the MAR model as densely time-resolved as an ODE model, it mitigates the
427 apparent granularity disadvantage considerably.

428 ALVI allows a choice between linear regression and matrix inversion. The former is simpler, because it
429 uses all points available, and faster due to the fact that no data sample needs to be chosen. In most cases
430 tested, it also produces good fits and estimates. However, the latter usually produces slightly better
431 results and works well even in cases where the ALVI-LR solution fails (Table 1). It also offers a natural
432 approach to inferring whole ensembles of well-fitting model parameterizations.

433 Results obtained with MARSS or with ALVI-LR are rather robust if the data are noisy, whereas solutions
434 with ALVI-MI may be sensitive to small alterations in the data. As an example, consider the synthetic MAR
435 data presented in Figure 2, where we added noise as a normally distributed variable with mean zero and
436 standard deviation of 0.2 of the dependent variable mean. For comparison, consider now an alternative
437 sample (Figure S8) obtained by applying the same procedure, but with a different seed for noise creation.
438 The alternative dataset is almost indistinguishable from the original dataset in Figure 2, but using the same
439 point sample determined for the original dataset, ALVI-MI produces different parameter values. This
440 means that if we calculate a new set of splines, we should also search for a new point sample. The

441 conclusion is that, although the inferred fits are similarly good, the parameter values associated with the
442 best fit for a noisy point sample may not be optimal for another noisy sample, which may not be surprising.
443 In fact, we showed in a different example with noise that the inferred parameter values yielded a better
444 SSE than even the true values (Section 3.1). The argument may also be turned around into a positive
445 feature: Different noisy datasets or subsamples of these datasets can easily be used to create natural
446 ensembles of models that characterize the underlying data in a robust manner and yield additional
447 insights into the variability of the model parameters.

448 The MARSS software makes modeling with MAR models easy, although not entirely automatic, as many
449 options must be tested to find the one that returns the best fit in each case. For example, one must decide
450 whether to use estimated initial conditions or the initial datapoints and which variables should have the
451 same noise level. By contrast, the ALVI method for LV models is novel (Voit et al., 2021), and while all
452 steps are straightforward and code is available on GitHub [[https://github.com/LBSA-VoitLab/Comparison-
453 Between-LV-and-MAR-Models-of-Ecological-Interaction-Systems](https://github.com/LBSA-VoitLab/Comparison-Between-LV-and-MAR-Models-of-Ecological-Interaction-Systems)], no formally published software
454 currently exists that encompasses all these steps in a streamlined manner. As a new tool, ALVI offers
455 several avenues for further refinement. One important component is optimal data smoothing with
456 splines, which requires the determination of a suitable number of degrees of freedom and may also
457 employ weights for different variables within a dataset.

458 A comparison of MARSS results with or without log transformation of the dependent variable abundances
459 did not yield clear results. If the MAR models are to be viewed as a multispecies competition models with
460 Gompertz density dependence (Certain et al., 2018), the log transformation is required (*Supplements*
461 Section 1.3). While inferences for LV models usually benefit from smoothing, the same is not true for MAR
462 model, where smoothing in some cases, but certainly not always, led to improved data fits (Table 2).

463 MARSS uses a steepest decent optimization step, which ALVI presently does not. Although ALVI already
464 performs better than MARSS (Table 1), it might be possible to improve its results even further by adding
465 a refinement step based on steepest-descent optimization. Steepest-descent methods tend to get
466 trapped in local minima if the initial guesses are poor, but ALVI would not likely encounter this problem,
467 as the solutions are already very good and could be used directly as initial guesses for the refinement step.
468 Estimation and inference methods typically do not scale well. ALVI bucks this trend, at least to some
469 degree, as both the smoothing and estimation of slopes occur one equation at a time. Thus, instead of
470 scaling quadratically, the inference problem scales linearly. The ultimate matrix inversion or linear
471 regression is essentially the same for all realistically sized models. Thus, the only time-consuming step
472 within ALVI-MI is the choice of datapoints. An exhaustive test for all combinations grows quickly in
473 complexity, but it is always possible to opt for a random search. The resulting solution is not necessarily
474 the best possible, but can still provide an excellent fit or at least a valuable starting point for a steepest-
475 decent refinement optimization. Importantly, many random solutions can also be collated to establish an
476 ensemble of well-fitting models, which in most cases yields more insight than a single optimized solution.

477

478 **5. Acknowledgements**

479 This work was supported in part by the following grant: NIH-2P30ES019776-05 (PI: Carmen Marsit).

480 The funding agency is not responsible for the content of this article.

481

482 **6. Author Contributions**

483 D.V.O., J.D.D., and E.O.V. conceived this project, performed the literature review, created the synthetic
484 data examples and wrote of the manuscript. D.V.O. produced the code required for the study. All
485 authors reviewed and edited the manuscript.

486

487 **7. Data Accessibility**

488 The R code for running the experiments is available on GitHub:

489 [https://github.com/LBSA-VoitLab/Comparison-Between-LV-and-MAR-Models-of-Ecological-Interaction-](https://github.com/LBSA-VoitLab/Comparison-Between-LV-and-MAR-Models-of-Ecological-Interaction-Systems)
490 [Systems](https://github.com/LBSA-VoitLab/Comparison-Between-LV-and-MAR-Models-of-Ecological-Interaction-Systems)

491

492 **8. Additional Information**

493 **Competing financial interests**

494 The authors declare no competing interests.

495

496 **9. References**

497 Certain, G., Barraquand, F., & Gårdmark, A. (2018). How do MAR(1) models cope with hidden
498 nonlinearities in ecological dynamics? *Methods in Ecology and Evolution*, 9(9), 1975–1995.

499 doi:10.1111/2041-210X.13021

500 Chou, I. C., & Voit, E. O. (2009). Recent developments in parameter estimation and structure

501 identification of biochemical and genomic systems. *Mathematical Biosciences*, 219(2), 57–83.

- 502 doi:10.1016/j.mbs.2009.03.002
- 503 Cleveland, W. S. (1981). LOWESS: A program for smoothing scatterplots by robust locally weighted
504 regression. *The American Statistician*, *35*(1), 54. doi:10.2307/2683591
- 505 Dam, P., Fonseca, L. L., Konstantinidis, K. T., & Voit, E. O. (2016). Dynamic models of the complex
506 microbial metapopulation of lake mendota. *Npj Systems Biology and Applications*, *2*(1), 16007.
507 doi:10.1038/npjbsa.2016.7
- 508 Dam, P., Rodriguez-R, L. M., Luo, C., Hatt, J., Tsementzi, D., Konstantinidis, K. T., & Voit, E. O. (2020).
509 Model-based comparisons of the abundance dynamics of dacterial communities in two lakes.
510 *Scientific Reports*, *10*(1), 1–12. doi:10.1038/s41598-020-58769-y
- 511 Dennis, B., & Taper, M. L. (1994). Density Dependence in Time Series Observations of Natural
512 Populations: Estimation and Testing. *Ecological Monographs*, *64*(2), 205–224.
513 doi:10.2307/2937041
- 514 Gause, G. F. (1934). Experiemental analysis of Vito Volterra’s mathematical theory of the struggle for
515 existence. *Science*, *79*(2036), 16–17. doi:10.1126/science.79.2036.16-a
- 516 Gavin, C., Pokrovskii, A., Prentice, M., & Sobolev, V. (2006). Dynamics of a Lotka-Volterra type model
517 with applications to marine phage population dynamics. *Journal of Physics: Conference Series*,
518 *55*(1), 80–93. doi:10.1088/1742-6596/55/1/008
- 519 Gerber, L. R., Demaster, D. P., & Kareiva, P. M. (1999). Gray whales and the value of monitoring data in
520 implementing the U.S. endangered species act. *Conservation Biology*, *13*(5), 1215–1219.
521 doi:10.1046/j.1523-1739.1999.98466.x
- 522 Holmes, E. E., Ward, E. J., & Scheuerell, M. D. (2020). Analysis of multivariate timeseries using the

- 523 MARSS package. Seattle: Northwest Fisheries Science Center, NOAA. Retrieved from [https://cran.r-](https://cran.r-project.org/package=MARSS/vignettes/UserGuide.pdf)
524 [project.org/package=MARSS/vignettes/UserGuide.pdf](https://cran.r-project.org/package=MARSS/vignettes/UserGuide.pdf)
- 525 Holmes, E. E., Ward, E. J., & Wills, K. (2012). MARSS: multivariate autoregressive state-space models for
526 analyzing time-series data. *The R Journal*, 4(1), 11. doi:10.32614/RJ-2012-002
- 527 Huffaker, C. B., Shea, K. B., & Herman, S. G. (1963). Experimental studies on predation: dispersion
528 factors and predator-prey oscillations. *Hilgardia*, 34, 305–330. Retrieved from
529 <http://hilgardia.ucanr.edu/fileaccess.cfm?article=152594&p=ZPTIMD>
- 530 Ives, A. R. (1995). Predicting the response of populations to environmental change. *Ecology*, 76(3), 926–
531 941. doi:10.2307/1939357
- 532 Knowles, I., & Renka, R. J. (2014). Methods for Numerical Differentiation of Noisy Data. *Electronic*
533 *Journal of Differential Equations*, 21, 235–246. Retrieved from [https://ejde.math.txstate.edu/conf-](https://ejde.math.txstate.edu/conf-proc/21/k3/knowles.pdf)
534 [proc/21/k3/knowles.pdf](https://ejde.math.txstate.edu/conf-proc/21/k3/knowles.pdf)
- 535 Lotka, A. J. (1925). *Elements of Physical Biology*. Baltimor: Williams & Wilkins Company. Retrieved from
536 <https://archive.org/details/elementsofphysic017171mbp>
- 537 May, R. M. (2001). *Stability and complexity in model ecosystems*. Princeton University Press.
538 doi:10.1515/9780691206912
- 539 McLaren, B. E., & Peterson, R. O. (1994). Wolves, moose, and tree Rings on Isle Royale. *Science*,
540 266(5190), 1555–1558. doi:10.1126/science.266.5190.1555
- 541 Mendes, P., & Kell, D. (1998). Non-linear optimization of biochemical pathways: applications to
542 metabolic engineering and parameter estimation. *Bioinformatics*, 14(10), 869–883.
543 doi:10.1093/bioinformatics/14.10.869

- 544 Mühlbauer, L. K., Schulze, M., Harpole, W. S., & Clark, A. T. (2020). gauseR: Simple methods for fitting
545 Lotka-Volterra models describing Gause's "Struggle for Existence". *Ecology and Evolution*, *10*(23),
546 13275–13283. doi:10.1002/ece3.6926
- 547 Park Service, N. (2021). Why relocate wolves to Isle Royale? Retrieved from
548 <https://home.nps.gov/isro/learn/why-relocate-wolves-to-isle-royale.htm>
- 549 Rykiel, E. J. (1996). Testing ecological models: the meaning of validation. *Ecological Modelling*, *90*(3),
550 229–244. doi:10.1016/0304-3800(95)00152-2
- 551 Sachs, A., & Goetze, A. (1945). *Mathematical cuneiform texts: 29 (American Oriental)*. (O. Neugebauer,
552 Ed.). New Haven: American Oriental Society. Retrieved from
553 <https://www.jstor.org/stable/1359232?seq=1>
- 554 Savageau, M. A. (1979). Allometric morphogenesis of complex systems: Derivation of the basic
555 equations from first principles. *Proceedings of the National Academy of Sciences*, *76*(12), 6023–
556 6025. doi:10.1073/pnas.76.12.6023
- 557 Shenhav, L., Furman, O., Briscoe, L., Thompson, M., Silverman, J. D., Mizrahi, I., & Halperin, E. (2019).
558 Modeling the temporal dynamics of the gut microbial community in adults and infants. *PLOS*
559 *Computational Biology*, *15*(6), e1006960. doi:10.1371/journal.pcbi.1006960
- 560 Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1. doi:10.2307/1912017
- 561 Stein, R. R., Bucci, V., Toussaint, N. C., Buffie, C. G., Räscher, G., Pamer, E. G., ... Xavier, J. B. (2013).
562 Ecological modeling from time-series inference: insight into dynamics and stability of intestinal
563 microbiota. *PLoS Computational Biology*, *9*(12), e1003388. doi:10.1371/journal.pcbi.1003388
- 564 Varah, J. M. (1982). A spline least squares method for numerical parameter estimation in differential

565 equations. *SIAM Journal on Scientific and Statistical Computing*, 3(1), 28–46. doi:10.1137/0903003

566 Voit, E. O., & Almeida, J. (2004). Decoupling dynamical systems for pathway identification from
567 metabolic profiles. *Bioinformatics*, 20(11), 1670–1681. doi:10.1093/bioinformatics/bth140

568 Voit, E. O., & Chou, I.-C. (2010). Parameter estimation in canonical biological systems models.
569 *International Journal of Systems and Synthetic Biology*, 1(June), 1–19.

570 Voit, E. O., Davis, J. D., & Olivença, D. V. (2021). Inference and validation of the structure of Lotka-
571 Volterra models. *BioRxiv*. doi:10.1101/2021.08.14.456346

572 Voit, E. O., & Savageau, M. A. (1982). Power-law approach to modeling biological systems; II. Application
573 to ethanol production. *J. Ferment. Technol.*, 60(3), 229–232.

574 Volterra, V. (1926). *Variazioni fluttuazioni del numero d'individui in specie conviventi. Men Acad Lincei.*

575 Vucetich, J. A. (2021). Wolves and moose of Isle Royale. Retrieved from <https://isleroyalewolf.org/>

576 Wedelin, D., & Gennemark, P. (2007). Efficient algorithms for ordinary differential equation model
577 identification of biological systems. *IET Systems Biology*, 1(2), 120–129. doi:10.1049/iet-
578 syb:20050098

579

580

581

582

583

584

585

586

587

588

589

590 **SUPPLEMENTS**

591 **Comparison Between Lotka-Volterra and Multivariate Autoregressive Models of** 592 **Ecological Interaction Systems**

593

594 **Daniel V. Olivença, Jacob D. Davis, Eberhard O. Voit**

595

596 **1. Materials and Methods**

597 **1.1. Lotka-Volterra models**

598 For a single variable, Lotka-Volterra (LV) models (Eq. [1] in the main text) reduce to the well-known logistic
599 growth law

$$600 \quad \frac{dX_{i,t}}{dt} = a_i X_{i,t} - b_{ii} X_{i,t}^2 \quad [S1]$$

601 where the ratio a_i/b_{ii} is called the “carrying capacity” of the system, which corresponds to the non-trivial
602 steady state. If time-dependent environmental inputs are to be considered, one may add terms $\gamma_i X_i U_t$,
603 where U_t is an element of a vector of these inputs and the coefficients γ_i are weights that quantify the
604 effects of the factors on species X_i (Stein et al., 2013; Dam et al., 2016; Dam et al., 2020). The index t is
605 usually omitted, and the left-hand side is often written as \dot{X}_i .

606 The LV system is a *canonical* model in the sense that its mathematical structure is immutable and scalable
607 to any dimension (Voit, 2000). Such a canonical model may serve as a template to construct models of
608 different systems that reasonably satisfy the following assumptions:

- 609 • encounters between and within species are representable by mass action kinetics;

- 610 • the environment does not change during the process, unless environmental variables are
- 611 explicitly formulated as described above;
- 612 • the parameter values do not change during an experiment;
- 613 • the species respond to one another instantaneously;
- 614 • for very small population sizes, interactions are negligible and the rate of change (growth) of each
- 615 population is initially proportional to its size, resulting in initial exponential growth;
- 616 • adaptations of species are absent or negligible.

617 Although the model structure and these assumptions might appear to be unduly rigid, LV models are
618 extremely rich in the repertoire of their possible responses. In fact, the LV structure was shown to be
619 capable of modeling any type of differentiable nonlinearities, including different kinds of oscillations and
620 chaos (Vano et al., 2006), if sufficiently many auxiliary variables are permitted, some of which have
621 mathematical, but no real biological meaning (Voit & Savageau, 1986; Peschel & Mende, 1986; Savageau
622 & Voit, 1987). At the same time, the LV structure has severe limitations. For example, it is not well suited
623 for metabolic pathway systems, because a simple conversion of a substrate X_1 into a product X_2 would
624 require X_2 to appear in its own synthesis term, although the generation of X_2 in truth depends only on X_1
625 and possibly some modulators (see (Voit, 2013) for this and other limitations).

626 LV models were initially used to describe the dynamics of predator and prey populations or of populations
627 that compete for the same resources, but the same equations have also been used in entirely different
628 contexts and fields, including physics (Nambu, 1986; Hacinliyan, Kusbeyzi, & Aybar, 2010), pollution
629 assessment (Haas, 1981), economy (Zhou & Chen, 2006; Gandolfo, 2008), manufacturing (Chiang, 2012),
630 and sales (Hung, Chiu, Huang, & Wu, 2017).

631 Beyond the fact that LV models can be formulated very easily, another significant advantage over other
632 systems of nonlinear ODEs is the fact that LV models can be parameterized with linear regression methods
633 if time series data are available (Voit & Chou, 2010). As an intriguing alternative, the linearity also permits
634 us to select variable and slope values at $n+1$ time points and to obtain parameter inferences by solving a
635 set of linear algebraic equations (see below). It is furthermore possible to estimate parameter values from
636 sufficiently many profiles of species that initially coexist and ultimately survive under comparable
637 conditions (Voit et al., 2021).

638

639 **1.2. Estimation of LV Parameters Based on Slopes of Time Courses**

640 This section explains in detail an approach to parameter estimation that uses the Algebraic Lotka-Volterra
641 Inference (ALVI) method. For a detailed explanation of the ALVI method itself please see (Voit et al., 2021).

642

643 **1.2.1. Smoothing**

644 Even though one might consider smoothing a conceptually separate issue from the actual parameter
645 inference, the two are so closely intertwined in our analysis that it appears useful to discuss smoothing
646 options. The goal is two-fold. First, it is beneficial to reduce or even remove noise from the raw data, and
647 second, this smoothing greatly aids the determination of slopes of the experimental time courses (see
648 later).

649 A smoothed representation of a dataset implicitly integrates information that is not in the data. This
650 implicit integration step is not entirely unbiased and requires prudent judgment, because it must answer
651 the following questions, often without true knowledge of the system: Are the deviations between the
652 data and the smoothing function due to (stochastic) noise or are they part of the true signal? For instance,

653 are they the trace of true oscillations? Also, if very few data points deviate much more than all others
654 from the smoothing function, are they true peaks or valleys or are they statistical outliers? It is difficult to
655 answer these questions objectively, but two features of the data are of great benefit: First, if the variation
656 in noise amplitude is much smaller than the range of signal values (high signal-to-noise-ratio), the
657 distinction between signal and noise is relatively straightforward. Second, if the data come in replicates,
658 they may support or refute the potential of true oscillations or peaks at certain time points in the data.
659 Even if only one dataset is available, the biologist familiar with the phenomenon at hand usually has
660 developed an expectation regarding signal and noise, and if there is no biological rationale for expecting
661 oscillations or strong deviations from some simple trend, the smoothing strategies are flexible enough to
662 allow the integration of the biologist's knowledge and expectations. The result of the smoothing process
663 therefore is a synthesis of all relevant information, constrained by external knowledge and reasonable
664 expectations. Of course, it is also feasible to create alternative models with different thresholds between
665 signal and noise and to analyze them side by side.

666 Independent of the options and intricacies of obtaining smoothed time courses of all variables, it is well
667 known that the process of estimating slopes from data is more strongly affected by noise than the data
668 themselves (Knowles & Renka, 2014). Expressed differently, if the noise is left unchecked, its effect on the
669 estimated values of the slopes tends to be higher than its effect on the values of the variables.

670 We explored a number of methods for smoothing the time course data and keeping the noise in check
671 (Eilers, 2003; Vilela et al., 2007; Batista Júnior & Pires, 2014), cognizant of the fact that empirical raw data
672 alone do not provide enough information of what is noise and what is relevant signal in the dynamics of
673 the phenomenon under study.

674 One of the simplest approaches is the *three-point method*, where the slope at time point t_k is taken as the
675 average of the slopes at time points t_{k-1} and t_{k+1} (Burden, Faires, & Burden, 1993; Voit & Almeida, 2003).

676 More sophisticated methods were reviewed in (Cleveland & Grosse, 1991; Eilers & Marx, 1996; Batista
677 Júnior & Pires, 2014). For long, dense time series, moving average and collocation methods with or
678 without roughness penalty (Ramsay et al., 2007) are often very effective. However, they tend to be
679 unsuited for biological time series data because the measurements are usually quite sparse and obtained
680 over a relatively short time horizon.

681 Smoothing splines and *local regression* methods like LOESS (locally estimated scatterplot smoothing) and
682 LOWESS (locally weighted scatterplot smoothing) turned out to be particularly useful. A detailed
683 description of these methods can be found in (Cleveland, 1981).

684 In a nutshell, splines are piecewise polynomial functions that: pass through all sample points, are
685 continuous and have first and second derivatives that are continuous at junction points between adjacent
686 intervals. In a smoothing spline, the first condition is substituted by a least-squares fit that is balanced
687 with an additional criterion that penalizes splines with high second derivative values, which indicate local
688 roughness (Cleveland, 1979; Garcia, 2010; Loader, 2012).

689 LOESS and LOWESS algorithms use locally-weighted polynomial regression. LOWESS is used for univariate
690 smoothing and consists of computing a series of local linear regressions, with each local regression
691 restricted to a window of x-values. Smoothness is achieved by using overlapping windows and by gradually
692 down-weighting points in each regression according to their distance from the anchor point of the window.
693 LOESS is for fitting a smooth surface to multivariate data and it is a generalization of LOWESS in that locally
694 weighted univariate regressions are simply replaced by locally weighted multiple regressions. While LOESS
695 is more versatile, LOWESS is faster and sometimes succeeds when LOESS fails (Cleveland, 1979; Cleveland
696 & Devlin, 1988; Smyth, 2020). Locally-weighted polynomial regression methods have ‘span’ and splines
697 have ‘degrees of freedom,’ which are parameters that control the degree of smoothing.

698 The main result of smoothing with splines is a reduction or even removal of what is believed to be noise
699 in the data. The slope at each point can be computed directly from the smoothing spline, which after all
700 is an explicit function. This step of slope determination offers two options: it allows us to estimate slopes
701 only for the measured data points or to sample the smoothing function for any number of other points,
702 which yields a larger set of numerical values for variables and slopes (Voit & Almeida, 2004).

703 If we select many points from the smoothing spline, we overcome the problem of data scarcity that is
704 inherent in many datasets. In fact, sampling from the smoothing spline allows the subsequent parameter
705 inference method to access a larger amount of information and thereby to mitigate noise amplification.

706

707 **1.2.2. Conversion of ODEs into systems of algebraic equations**

708 If data are available as time series, it is mathematically feasible and beneficial to estimate slopes (for
709 instance, from smoothing splines) and to convert the inference problem from one based on ODEs into one
710 exclusively using algebraic functions (Voit & Savageau, 1982a, 1982b; Varah, 1982; Torres & Voit, 2002;
711 Voit et al., 2005).

712 Suppose the growth and interaction parameters of an LV system are to be estimated from time series
713 data of the dependent variables X_i . The smoothing of these data facilitates the estimation of slopes $S_k(X_i)$
714 of all variables at a set of time points t_k , $k = 1, \dots, K$. These time points may or may not correspond to the
715 measured data. In fact, the smoothing permits the computation of slopes at arbitrarily many time points
716 within the observation interval. However many slopes are computed, they correspond to derivatives of
717 the spline of X_i at the given time points. Substituting numerical values of all variables and slopes from the
718 smoothing splines into Eq. (1) yields a system of $n \times K$ linear algebraic equations containing all system
719 parameters:

720
$$S_{i,t} = a_i X_{i,t} + \sum_{j=1}^n b_{ij} X_{i,t} X_{j,t}, \quad i = 1, \dots, n \quad [S2]$$

721 If environmental inputs $\gamma_i X_{i,t} U_t$ are to be considered as well, they are added to the equations and either
722 substituted with numerical values, if known, or estimated with the parameters a_i and b_{ij} .

723 A caveat of this conversion of ODEs into algebraic equation is a possible time warp (see end of chapter 5
724 of (Voit, 2017)). The reason is that time is explicitly eliminated from the procedure. Nonetheless, the
725 estimates usually provide good results, or at least good initial guesses for other optimization approaches
726 such as traditional gradient methods.

727 Suppose the dependent variables are not zero within the dataset obtained from smoothing. If so, we can
728 divide both sides of the K equations for X_i in expression [S2] by the value of the dependent variable at
729 the appropriate time point. This step is not mandatory but linearizes the equations. The case of variables
730 with values of zero is typically not very interesting or can be handled by eliminating the variable or parts
731 of the time series.

732

733 **1.2.3. Parameter inference**

734 Once all differentials are replaced with estimated slopes, the inference of parameter values from LV-
735 models offers two options: because the system of algebraic equations is linear, we may optimize its
736 parameter values through simple multivariate linear regression (ALVI-LR), where we may use data points
737 or iterate the regression with subsets of points, which naturally leads to an ensemble of well-fitting
738 models.

739 An interesting alternative is to use just $n+1$ of the data points and slopes, if n is the number of variables,
740 which results in a system of linear equations that can be solved with simple algebraic methods (ALVI-MI).

741 Choosing different data points naturally creates ensembles of solutions. These can be further analyzed,
742 for instance, with respect to model robustness and identifiability. They can also be used to determine to
743 what degree the LV format is adequate for the available data (Voit et al., 2021).

744

745 **1.2.4. Example of parameter estimation with ALVI**

746 To explain the parameter estimation procedure with ALVI, we use the sparse noisy dataset presented in
747 *Supplements* Section 2 and also in Table S1.2. First, we smooth the data with a spline or LOESS. For this
748 example, we use 5, 8, 11 and 5DF-splines for X_1 , X_2 , X_3 and X_4 respectively and compute the first derivative
749 of the splines to estimate the slopes. At this point we discard the data and only use the spline values. We
750 may use the original data, especially if we think they characterize the studied phenomenon well, but using
751 the spline usually produces better results in the case of noisy data.

752 As an example, consider the first differential equation and the first datapoint, at $t = 0$:

$$753 \quad \frac{dX_1}{dt}(0) = a_1X_1(0) + b_{11}X_1(0)X_1(0) + b_{12}X_1(0)X_2(0) + b_{13}X_1(0)X_3(0) + b_{14}X_1(0)X_4(0)$$

754 We substitute numerical values for the slope and for all variables on the system equations,

time	Slope_ X_1	X_1	X_2	X_3	X_4
0	0.0353	1.229	0.492	1.787	0.0367

755

756 which yields

$$757 \quad 0.0353 = a_1 \times (1.229) + b_{11} \times (1.229)^2 + b_{12} \times (1.229) \times (0.492)$$
$$758 \quad \quad \quad + b_{13} \times (1.229) \times (1.787) + b_{14} \times (1.229) \times (0.0367).$$

759

760 We may divide the equation by the numerical value of the dependent variable, but this step is not
761 mandatory.

762

763
$$\frac{0.0353}{1.229} = a_1 + b_{11} \times (1.229) + b_{12} \times (0.492) + b_{13} \times (1.787) + b_{14} \times (0.0367)$$

764

765 The same steps are performed for every equation and every chosen time point. The result is a system of
766 linear equations with as many equations as chosen time points; each equation has $n+1$ unknown
767 parameters, where n corresponds to the number of dependent variables.

768 Now we have two options: We may use linear regression (ALVI-LR) or matrix inversion (ALVI-MI). ALVI-LR
769 uses every equation and every chosen time point and performs linear regression to produce estimates for
770 the parameters. For the alternative, ALVI-MI, we choose a sample of data points that, when combined
771 with the equations, generates a number of equations equal to the number of parameters to be estimated.
772 If these equations are linearly independent, the system is solvable and the solution is unique, allowing us
773 to obtain estimates for the parameters by simple matrix inversion.

774

775 **1.3. Multivariate Autoregressive (MAR) models**

776 Multivariate Autoregressive (MAR) models are discrete recursive models. Their format is shown in
777 eqn. 2 and 3 of the main text and conveys that the state of the system at time $t+1$ depends on the state
778 at t and possibly on environmental or stochastic input. As an alternative to this modeling structure with
779 “memory 1,” it is possible to extend MAR models to depend also on states farther in the past, such as X_t
780 ₁, X_{t-2} , and X_{t-3} , in addition to X_t . However, the commonly used models depend only on the immediately

781 prior state and are sometimes called MAR(1). Here, we only consider MAR(1) models and refer to them
782 simply as MAR models.

783 MAR models can be interpreted in two distinct ways. In generic mathematical terms, MAR models are
784 stochastic, linear approximations of nonlinear dynamic systems that evolve over time in the vicinity of a
785 fixpoint (steady state). According to this interpretation (Holmes et al., 2013), X_t is a vector of the
786 realization of random variables at time t . Noise captures natural variations in environmental conditions
787 and is modeled by a multivariate normal distribution with mean zero and variance-covariance matrix δ .
788 If stochasticity is omitted, MAR models are quite similar to LV models close to the steady state (see
789 *Supplements* Section 1.5).

790 One may also interpret MAR models in an ecological context, where they can be viewed as multispecies
791 competition models with Gompertz density dependence and an instantaneous growth rate that decreases
792 linearly over time as the population sizes increase (Ives, 1995; Certain, Barraquand, & Gårdmark, 2018).
793 In this context, X_t is a vector of the log-abundances of dependent variables at time t .

794 MAR models may be augmented with state variables that simulate the observation process and these
795 models are called Multivariate Autoregressive(1) State-Space (MARSS) (Holmes et al., 2012; Certain et al.,
796 2018); we will not analyze these for the sake of simplicity. For the comparisons in this study, we are not
797 considering the influence of environmental variables, so the γu_t term will be omitted henceforth.

798 It is considered an advantage in ecology if models explicitly take the influence of environmental factors
799 into account (for details, see (Certain et al., 2018) (Hytti et al., 2006)), which is the case for MAR. The
800 availability of estimation software like MARSS (Holmes et al., 2012, 2020) has greatly increased the appeal
801 of MAR models.

802

803 **1.4. MARSS**

804 MARSS is a software package for analyzing MAR models with or without log transformation of the
805 dependent variables. Its use requires several steps.

806 1 – Specify key MARSS settings:

B = "unconstrained"	Matrix with all elements potentially different
U = "unequal"	Vector with all elements potentially different
Q = "diagonal and unequal"	Matrix where all elements are zero except the main diagonal, where the elements have real values
Z = "identity"	Identity matrix
A = "zero"	All elements zero
R = "zero"	All elements zero
x0 =	Initial values of the time series

807

808 **A** and **R** correspond to the “observation variables,” which simulate the observation process of the system
809 variables. For our comparisons, these are set to zero because observation variables are not considered.

810 2 – In the MARSS function, the data must be formatted with variables in rows and observations in columns.
811 If the data points are not equally distributed in time, they must be augmented by “NA” to force the interval
812 between any two consecutive data entries to be of the same length. This is necessary to ensure the correct
813 time structure of the data for the estimator.

814 3 – With this regularization, MARSS finds estimates for **B**, **U** and **Z**, that correspond to α , β and δ in eqn 3
815 of the main text.

816 If MARSS does not converge, it is advisable to increase the max number of iterations. This step solves the
817 problem but is different from the suggestion offered by Holmes and colleagues, namely, that the model
818 assumptions should be checked (see p. 57 in (Holmes et al., 2020)).

819 Our setup is exactly equal to that used by Holmes *et al.* (Holmes et al., 2012, 2020) for the Isle Royale
820 dataset, which the authors used to exemplify the inference of species interaction parameters with and
821 without covariates. Some of the illustration examples were modeled differently in the literature but for
822 purpose of comparisons with LV models, this model structure was used. For example, the ‘gray whales’
823 dataset was modeled by Holmes and colleagues with β , the species interaction matrix, set to zero,
824 whereas \mathbf{R} , the matrix that captures the noise from the observational process, was estimated from the
825 data. Because we are interested in the interactions between species, we do not focus on observational
826 noise, and Holmes’ original setup was replaced with the one discussed above.

827

828 **1.5. Structural Similarities between Modeling Formats**

829 The two modeling formats appear to be very different mathematically. Nonetheless, they can be compared in terms
830 of their mathematical representations and also with respect to practical considerations. These comparisons
831 demonstrate that the two models can actually behave quite similarly if the community of populations operates
832 relatively close to a stable steady state. By contrast, if this assumption is violated, the two models often show
833 strongly diverging results, as the linearity of the MAR model can deviate considerably from the nonlinearities of the
834 LV model.

835 Purely considered on mathematical grounds, MAR is defined recursively in eqn 2 and 3 of the main text. By omitting
836 environmental variables, we directly obtain

$$837 \quad X_{i,t+1} = \alpha_i + \sum_{j=1}^n \beta_{ij} X_{j,t} + w_{i,t}; \quad i = 1, 2, \dots, n; \quad w_{i,n} \sim N(0, \delta_i), \quad [S3]$$

838 which can be interpreted as a multispecies competition model with Gompertz density dependence, if the
839 dependent variables represent logarithmic abundancies (Certain et al., 2018).

840 Suppose that the MAR model indeed uses log-abundances. To explain similarities between the MAR and LV
 841 formats, we rewrite this multispecies log-abundance model equivalently in Cartesian form, which yields the
 842 following:

843 For $i = 1, 2, \dots, n$; $w_{i,n} \sim N(0, d_i)$:

$$\begin{aligned} \ln(X_{i,t+1}) &= \alpha_i + \sum_{j=1}^n \beta_{ij} \ln(X_{j,t}) + w_{i,n} \\ \Leftrightarrow X_{i,t+1} &= e^{(\alpha_i + \sum_{j=1}^n \beta_{ij} \ln(X_{j,t}) + w_{i,n})} \\ \Leftrightarrow X_{i,t+1} &= e^{\alpha_i} e^{\left(\sum_{j=1}^n \ln((X_{j,t})^{\beta_{ij}})\right)} e^{w_{i,n}} \\ \Leftrightarrow X_{i,t+1} &= e^{\alpha_i} \prod_{j=1}^n e^{\left(\ln((X_{j,t})^{\beta_{ij}})\right)} e^{w_{i,n}} \\ \Leftrightarrow X_{i,t+1} &= e^{\alpha_i} \prod_{j=1}^n (X_{j,t})^{\beta_{ij}} e^{w_{i,n}} \quad [S4] \end{aligned}$$

844

845 In this form, MAR is similar to a discrete multivariate power-law function.

846 The similarity of this result to the LV model can be seen if we use Euler's method for determining the numerical
 847 solution. Euler's method is an approximation of more sophisticated methods and its simplicity makes it preferable
 848 for the comparisons between time series and ODEs.

849 Formulating the typical Euler step for the LV model transforms the ODE into a series of discrete steps of the type

$$850 \quad X_{i,t+h} = X_{i,t} + h * \left. \frac{dX_i}{dt} \right|_{X_i=X_{i,t}} = X_{i,t} + h * X_{i,t} (a_i + \sum_{j=1}^n b_{ij} X_{j,t}), \quad i = 1, 2, \dots, n, \quad [S5]$$

851 where h is the step size of Euler's method and dX_i/dt is the left-hand side of the differential equations in eqn 1,
 852 evaluated at time t .

853 A comparison of eqn [S4] and [S5] suggests that the MAR and LV models seem to be very different. Whereas the
 854 LV model captures nonlinear dynamic behaviors without variable transformations, the MAR model uses linearity in

855 log space. Nonetheless, there are similarities between the two formats. To see these, we compare eqn [S3] and [
 856 S5] instead of [S4] and [S5].

857 Furthermore, we suppose that the dynamics is near the steady state of the differential equations, so that $X_{i,t+1} - X_{i,t}$
 858 ≈ 0 for any given t . Using this approximate equality in eqn [S3], we obtain

859

$$860 \quad X_{i,t+1} - X_{i,t} \approx 0$$

$$861 \quad \Leftrightarrow \alpha_i + \sum_{j=1}^n \beta_{ij} X_{j,t} + w_{i,t} - X_{i,t} \approx 0; i = 1, 2, \dots, n; w_{i,n} \sim N(0, \delta_i). \quad [S6]$$

862

863 Using this approximate equality in format [S5] for LV yields, for $i = 1, 2, \dots, n$:

$$X_{i,t+h} - X_{i,t} \approx 0$$

$$\Leftrightarrow X_{i,t} + h * X_{i,t} \left(\alpha_i + \sum_{j=1}^n b_{ij} X_{j,t} \right) - X_{i,t} \approx 0$$

$$\Leftrightarrow \alpha_i + \sum_{j=1}^n b_{ij} X_{j,t} \approx 0$$

$$\Leftrightarrow \alpha_i + \sum_{j=1}^n \tilde{b}_{ij} X_{j,t} - X_{i,t} \approx 0 \quad [S7]$$

864 Here \tilde{b}_{ij} equals b_{ij} for all $i \neq j$ and $b_{ij} + 1$ for $i = j$.

865 If we disregard the Gaussian noise, w_{ij} , in the MAR model, the two sets of near-steady-state eqn [S6] and
 866 [S7] are the same. They are both linear, although the dynamic LV model itself is non-linear. Thus, if $\alpha_i =$
 867 a_i , $\beta_{ij} = b_{ij}$ for all $i \neq j$ and $\beta_{ij} = b_{ij} + 1$ for $i = j$, the MAR and LV models are mathematically equivalent
 868 at the steady state and similar close to it. As long as the nonlinearity are close to linear or close to power-
 869 law functions, MAR without or with log-transformation, respectively, may be expected to lead to
 870 acceptable fits.

871

872 **2. Case study 1: Synthetic LV data**

873 As a representative example, we use the four-variable LV system

$$874 \quad \frac{dX_{i,t}}{dt} = a_i X_{i,t} + b_{i1} X_{i,t} X_{1,t} + b_{i2} X_{i,t} X_{2,t} + b_{i3} X_{i,t} X_{3,t} + b_{i4} X_{i,t} X_{4,t}, \quad i = 1, \dots, 4 \quad [S8]$$

875 The parameters are presented in Figure S1. For a first analysis, we use this system to create one set of
876 synthetic time courses, consisting of 100 time points, which is presented in Table S1. The dynamics is
877 shown in Figures 1 and S1 as circles.

878 If we use these noise-free data, the inferences are close to perfect with respect to the trajectories and
879 parameter values (Figure S9).

880 To mimic a more realistic scenario, we created a noisy dataset, visualized in Figure S1a, which was
881 constructed by randomly choosing forty of the one hundred original datapoints obtained from the
882 synthetic system and adding to the chosen points a normal random variable with mean 0 and a standard
883 deviation of 20% of the mean of each variable. This *noisy dataset* is shown in Table S1.2.

884 A second realistic dataset (Figure S1b) was constructed by first choosing eleven points from the data that
885 characterize the dynamic (including extremes values). Next, each of the chosen points was multiplied by
886 a random normal variable with mean 1 and standard deviation of 0.2. This process was iterated to create
887 five replicates per chosen point. This *replicate dataset* is shown in Table S1.3.

888 Variable X_4 was designed as a (decoupled) logistic function. It is unaffected by the other variables and
889 does not affect them either. It was included to explore to what degree the methods to be tested can
890 detect this detachment.

891 The smoothing and slope estimation steps followed directly the procedures described in *Supplements*
892 Section 1.2. The first derivative of the smoothing function was used for estimates for the slopes.

893 To infer numerical values for the parameters of a given equation, we have the choice between linear
894 regression (ALVI-LR) and algebraic matrix inversion (ALVI-MI). For ALVI-MI, we choose points from the
895 sample and use the corresponding slope estimates to create a system of equations with the same number
896 of equations and unknowns. As we have 4 variables and 20 parameters, we need 20 independent
897 equations and thus 5 time points. For each time point we obtain the value for each of the 4 dependent
898 variables and use these to populate the equations.

899 As an illustration for the noisy dataset, we choose 5, 8, 11 and 5DF-splines and time points $t = 8, 13, 17,$
900 19 and 79 for the noisy dataset. For the replicate dataset, we use 5, 6, 8 and 8DF-splines, and the ALVI-MI
901 solution was calculated with spline points at times 11, 18, 26, 33 and 60. The time point selection for the
902 ALVI-MI solution can be automated using a random or exhaustive search of the possibilities.

903 The noisy dataset (Figures 1a, S1a, S3a and S4a) is representative of a study where each time point sample
904 corresponds to a single observation taken when it is possible or convenient. By contrast, the replicate
905 dataset (Figures 1b, S1b, S3b and S4b) simulates a series of experimental replicates where the
906 observations were conducted multiple times, but at fewer time points, which the researchers suspect
907 would contain valuable information.

908 As an illustration of how the smoothing techniques work, we used splines and LOESS with different
909 degrees of smoothing. The results for variable X_1 are shown in Figure S2. Choosing the optimal degree of
910 smoothing is not a trivial matter. Too much smoothing ignores important details in the variable dynamics,
911 while too little incorporates noise. In the programming language R, the function “loess.as” allows the
912 calculation of the optimum value for the spam, which controls the smoothing. The user still must decide

913 the degree of the polynomials to be used and choose from two criteria for automatic smoothing
914 parameter selection: a bias-corrected Akaike information criterion (AICC) and generalized cross-validation
915 (GCV). This choice is not always leading to the optimal solution, but it should be used to challenge our
916 assumptions.

917 Figure S3 presents the same treatment as presented in Figure 1 but for datasets with 5% noise. It is clear
918 and not surprising that the models produce better quality fits when the signal-to-noise ratio is higher.

919

920 **2.1.1. Application of ALVI to synthetic data**

921 An alternative to using the algebraic parameter inference method with matrix inversion (ALVI-MI) is the
922 linear regression method (ALVI-LR); fits for 20% noise are shown in Figure S4. As in ALVI-MI, the parameter
923 values are close to the true values and the fit is acceptable. However, the dynamics of the system is slightly
924 different. One reason is that the dynamic solutions are sometimes quite sensitive to the chosen initial
925 values. As a remedy, it is often beneficial to initiate the solution somewhere inside the overall time
926 interval, typically close to the midpoints of the variable ranges, and solve forward and backward.

927 ALVI also works for more complicated dynamics, as can be seen in Figure S5. Here we are interested in
928 finding out if the methods can recover the dynamics, which in some cases turns out to be challenging for
929 sparse data even without the introduction of noise. Thus, we used the synthetic data unaltered.
930 Specifically, data for early time points ($t = [1, 100]$) were fitted and then extrapolated for a total time
931 horizon of ($t = [1, 500]$). In these examples, ALVI-MI is used with 100DF-splines. It uses data samples with
932 points corresponding to timepoints $t = 5, 10, 20, 30$ and 50 for all cases except for the chaotic oscillations
933 where we used $t = 4, 6, 10, 15$ and 35 . For each case, we also present the MAR estimates. Of course, one
934 must recall that the original data were produced with LV models. While the MAR model extrapolations

935 are not always satisfactory, it is nevertheless comforting that the inference method returns good results
936 for the time interval used for data fitting.

937 ALVI-MI generally performed very well but did not adequately capture the deterministic chaos (chaos 1). For this
938 case only, we obtained a better fit using ALVI-LR, which may not be surprising given the extremely sensitive nature
939 of chaotic systems to noise. Apart from this situation, results with ALVI-LR are very similar to ALVI-MI results and will
940 not be displayed.

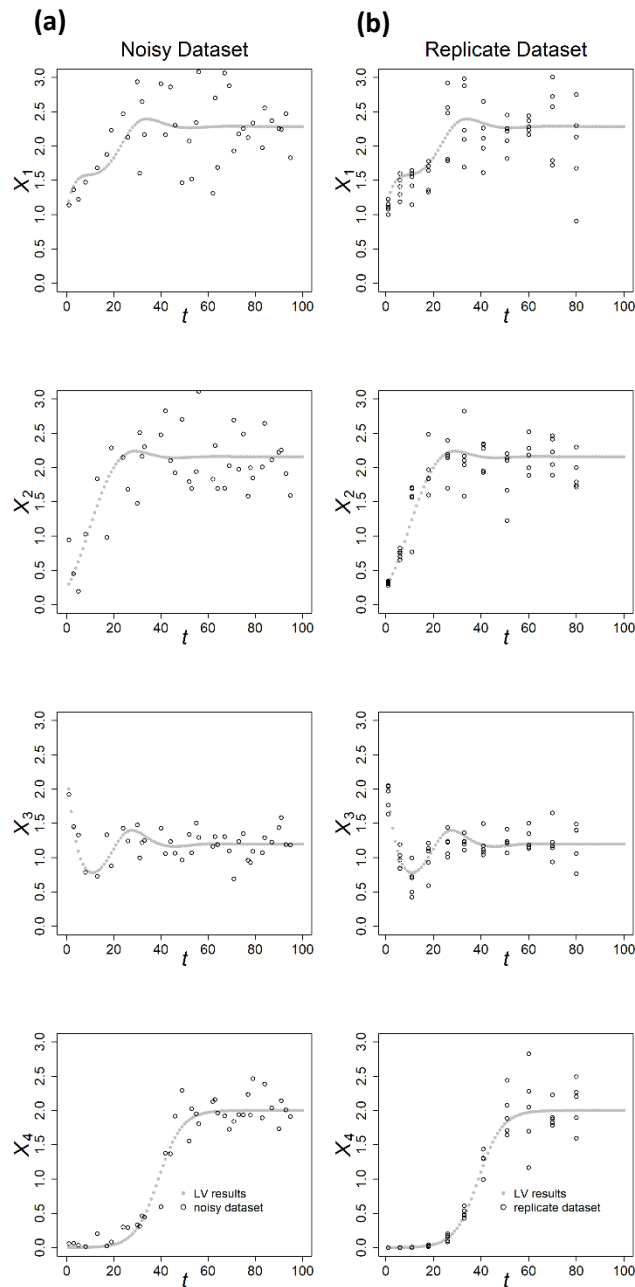
941 In Figure S5b, the MAR model performed well when log-abundances were used. In the remaining cases, it failed to
942 replicate the oscillations or these exploded by reaching amplitudes far bigger than in the dataset. One also notes
943 early discrepancies between the initial points used to create the estimates and the MAR estimates.

944

945 **3. Supplemental Figures**

946 **Figure S1**

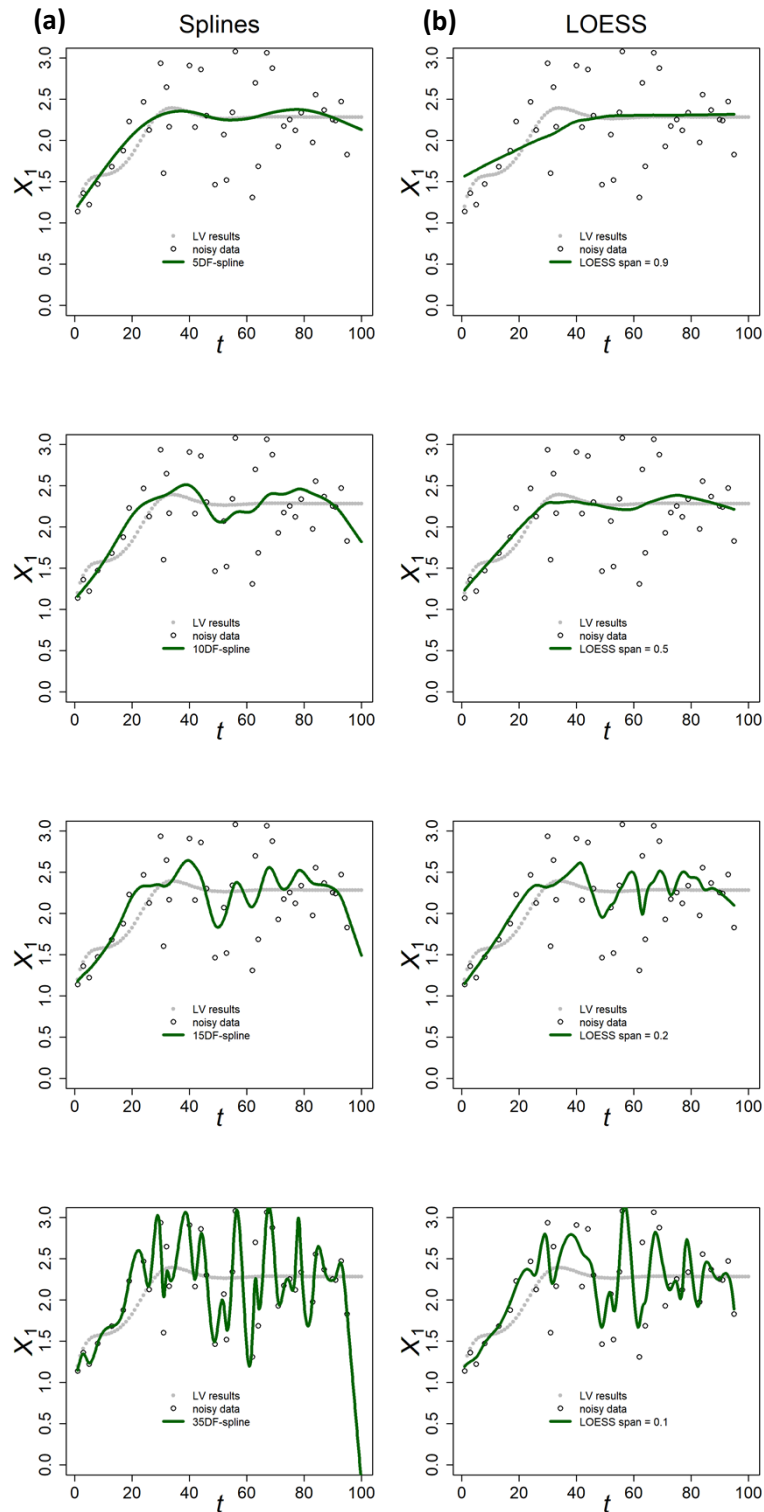
	Initial Conditions
X_1	1.2
X_2	0.3
X_3	2
X_4	0.001
	True Parameter Values
a_1	0.044
b_{11}	-0.08
b_{12}	0.02
b_{13}	0.08
b_{14}	0
a_2	0.216
b_{21}	-0.04
b_{22}	-0.08
b_{23}	0.04
b_{24}	0
a_3	0.116
b_{31}	-0.16
b_{32}	0.16
b_{33}	-0.08
b_{34}	0
a_4	0.2
b_{41}	0
b_{42}	0
b_{43}	0
b_{44}	-0.1



947 **Figure S1: Time courses with superimposed noise.** Initial conditions and parameter values for the
 948 synthetic LV example in equation S8 with four dependent variables. **Column a:** Noisy dataset – Based on
 949 40 points from the synthetic data with added random normal noise with mean 0 and standard deviation
 950 equal to 20% of each variable mean. **Column b:** Replicate dataset – 11 points were chosen from the
 951 synthetic data and at each point five “observations” were created by multiplying the value of the variable
 952 by a random normal value of mean 1 and standard deviation of 0.2.

953
 954

955 **Figure S2**

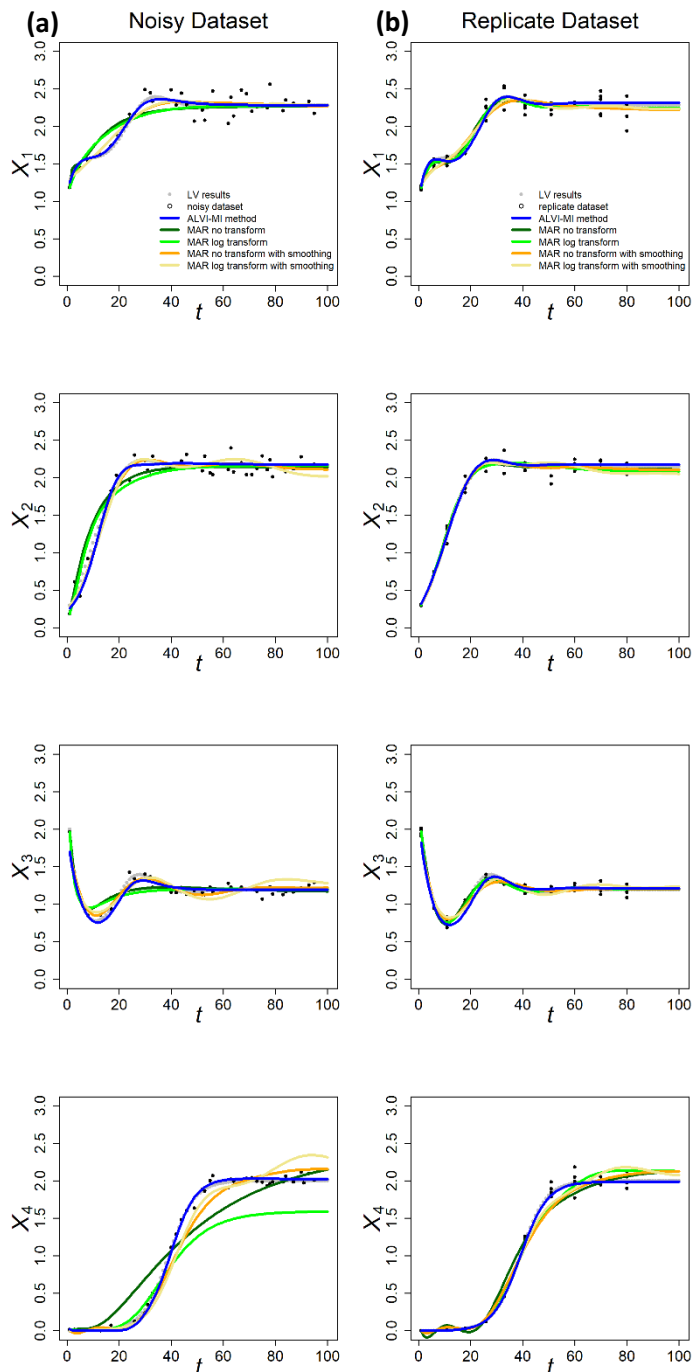


956
957 **Figure S2: Smoothing noisy variable X_1 .** Column a: Splines with different degrees of freedom. Column b:
958 LOESS with different span levels.

959 **Figure S3**

	Initial Conditions	Noisy Data	Replicate Data
X_1	1.2	1.25117	1.17824
X_2	0.3	0.25825	0.30397
X_3	2	1.69737	1.97235
X_4	0.001	0.02083	0.00104

	True Parameter Values	Noisy Data	Replicate Data
a_1	0.044	0.07373	0.03593
b_{11}	-0.08	-0.09759	-0.06540
b_{12}	0.02	0.02780	0.01935
b_{13}	0.08	0.07649	0.06434
b_{14}	0	-0.00149	-0.00401
a_1	0.216	0.24020	0.21072
b_{11}	-0.04	0.00429	-0.00647
b_{12}	-0.08	-0.08501	-0.09561
b_{13}	0.04	-0.04856	0.01263
b_{14}	0	-0.00323	-0.00423
a_2	0.116	0.10463	0.07192
b_{21}	-0.16	-0.13316	-0.16077
b_{22}	0.16	0.12651	0.16478
b_{23}	-0.08	-0.05638	-0.05221
b_{24}	0	-0.00441	0.00187
a_4	0.2	0.48906	0.19237
b_{41}	0	-0.79105	-0.00824
b_{42}	0	0.91565	0.00687
b_{43}	0	-0.30302	0.01005
b_{44}	-0.1	-0.15627	-0.09992

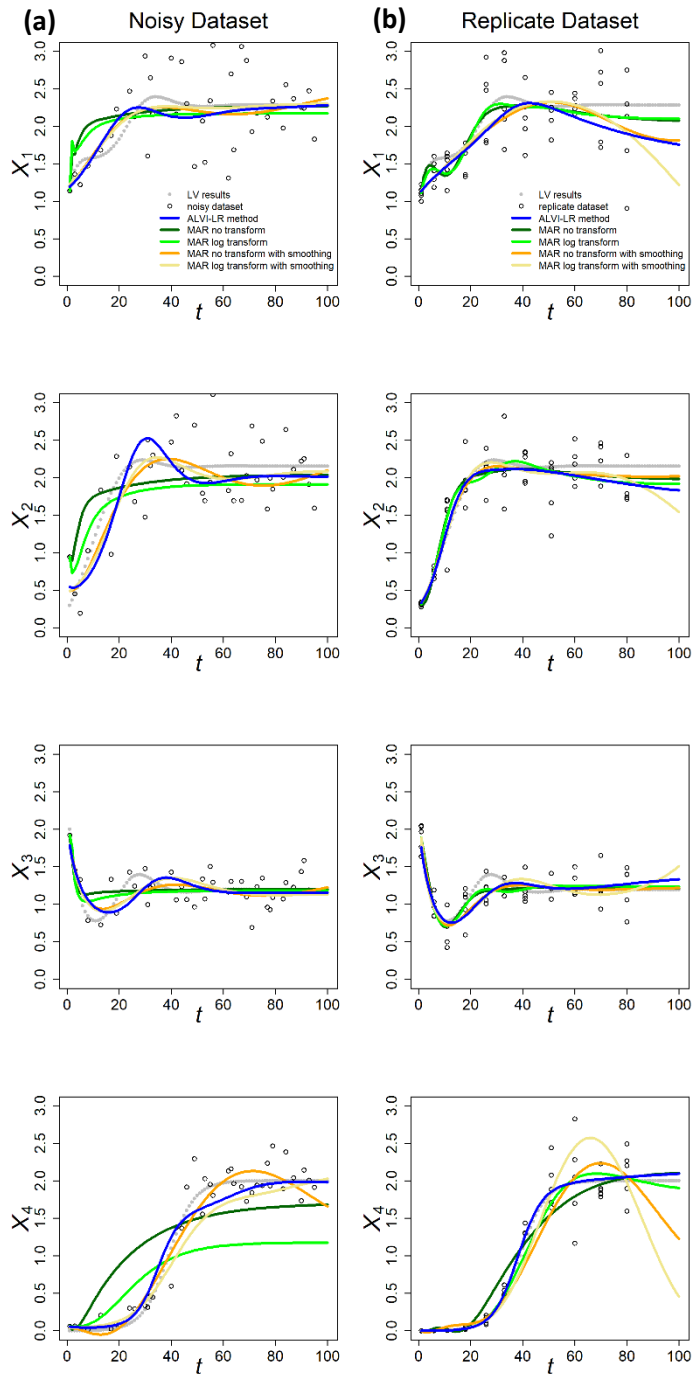


960 **Figure S3: Results of ALVI-MI and MAR applied to the noisy and replicate datasets with 5% of noise.**
 961 **Column a:** Noisy dataset. All variables were smoothed with 11DF-splines and the ALVI-MI solution was
 962 calculated with spline points at time points 8, 17, 30, 42 and 62. **Column b:** Replicate dataset. All variables
 963 were smoothed with an 8DF-spline and the ALVI-MI solution was calculated with spline points
 964 corresponding to time 6, 18, 26, 33 and 60. MAR estimates are presented in Tables S2.1 and S2.2.

965 **Figure S4**

	Initial Conditions	Noisy Data	Replicate Data
X_1	1.2	1.19177	1.13704
X_2	0.3	0.54552	0.33033
X_3	2	1.78728	1.75958
X_4	0.001	0.05701	0.00116

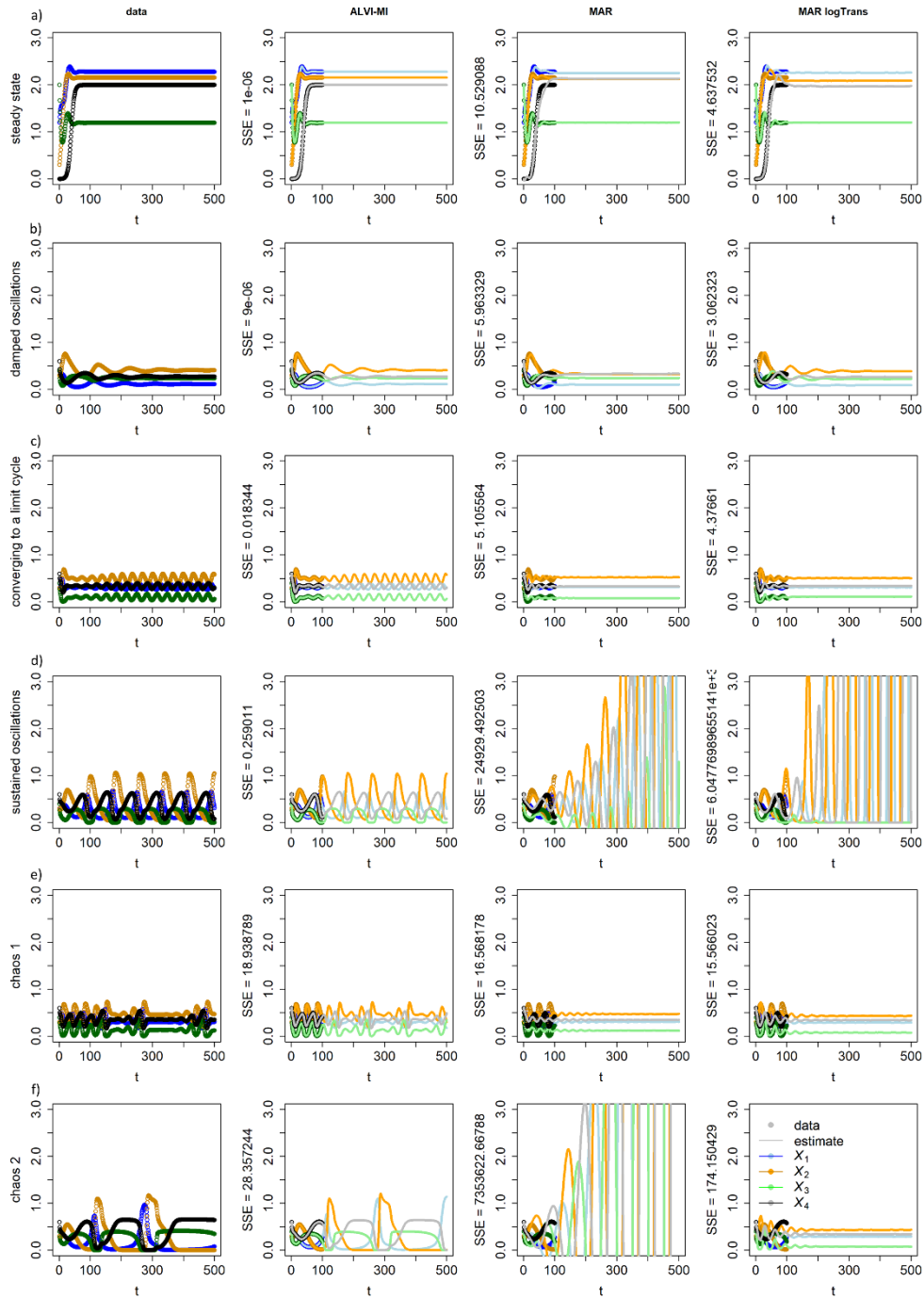
	True Parameter Values	Noisy Data	Replicate Data
a_1	0.044	0.05445	0.02433
b_{11}	-0.08	0.00671	-0.00738
b_{12}	0.02	-0.02130	-0.00173
b_{13}	0.08	-0.01411	0.01168
b_{14}	0	-0.00481	-0.01293
a_1	0.216	0.23293	0.21887
b_{11}	-0.04	0.00027	0.04438
b_{12}	-0.08	-0.01935	-0.11968
b_{13}	0.04	-0.14178	-0.04491
b_{14}	0	-0.01553	-0.00930
a_2	0.116	0.01081	-0.04261
b_{21}	-0.16	0.00380	-0.07696
b_{22}	0.16	0.04570	0.11963
b_{23}	-0.08	-0.08669	-0.02739
b_{24}	0	-0.00586	-0.00146
a_4	0.2	0.08891	0.16518
b_{41}	0	-0.02791	-0.02964
b_{42}	0	0.11969	0.02948
b_{43}	0	-0.11663	0.03522
b_{44}	-0.1	-0.06679	-0.10181



966

967 **Figure S4: Results of ALVI-LR and MAR applied to the noisy and replicate datasets. Column a:** Noisy
 968 dataset. Time courses of X_1 , X_2 , X_3 and X_4 were smoothed with 6, 11, 11 and 11DF-splines respectively.
 969 **Column b:** Replicate dataset. All variables were smoothed with 8DF-splines. MAR estimates are the same
 970 presented in Figure 1.

971 **Figure S5**



972

973

974

975 **Figure S5 – Data and results of inferences with ALVI-MI and MAR methods for LV systems with different**

976 **dynamics. Row a:** Data converging to a stable steady state; **Row b:** Damped oscillations; **Row c:** Initially

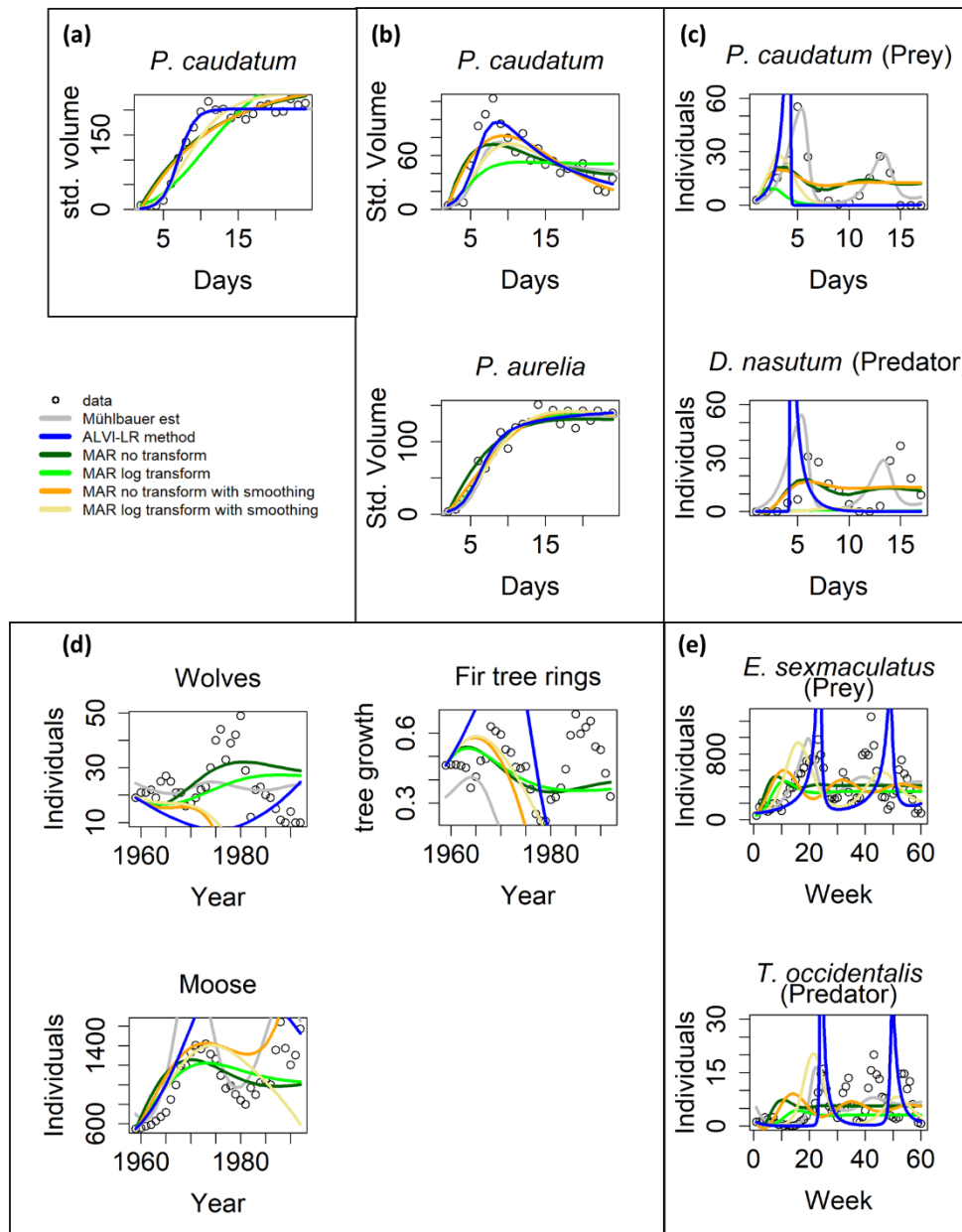
977 erratic oscillations converging to a limit cycle; **Row d:** Sustained oscillations; **Row e:** Deterministic chaos,

978 example 1; **Row f:** Deterministic chaos, example 2. Data, ALVI-MI and MARSS estimates are presented in

979 Table S3. The SSEs concerning the differences between the data and estimates for $t = [1, 500]$ are

980 presented as labels to the Y-axis. No smoothing preceded MAR because the data are noise free.

981 **Figure S6**

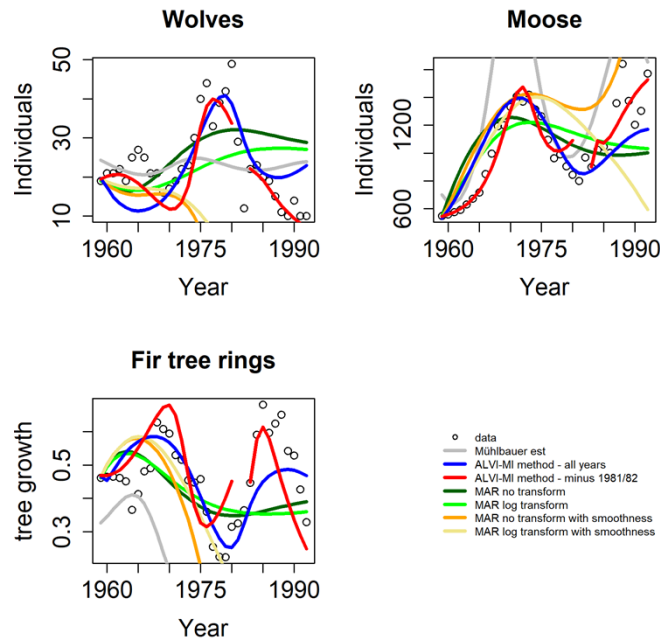


982 **Figure S6 – Examples of experimental data analyzed with ALVI-LR and MAR.** Black lines are estimates from
 983 Mühlbauer *et al.* (Mühlbauer *et al.*, 2020). ALVI-LR estimates are represented as blue lines; corresponding parameter
 984 values can be seen in Table S4.2. MAR estimates are presented in green, orange and yellow. Parameter estimates
 985 are presented in Table S4.3. **a:** Standardized volume of *Paramecium caudatum* grown in monoculture (Gause, 1934).
 986 **b:** Standardized volume of *Paramecium caudatum* and *Paramecium aurelia* grown in co-culture (Gause, 1934). **c:**
 987 Predator-prey interactions between *Didinium nasutum* and *Paramecium caudatum* grown in mixture (Gause, 1934).
 988 **d:** Multi-trophic dynamics for wolves, moose, and fir tree rings on Isle Royale from 1960 to 1994 (McLaren &
 989 Peterson, 1994). **e:** Predator-prey interactions between *Eotetranychus sexmaculatus* and *Typhlodromus occidentalis*
 990 in a spatially structured experiment (Huffaker, Shea, & Herman, 1963).
 991

992

993

994 **Figure S7**

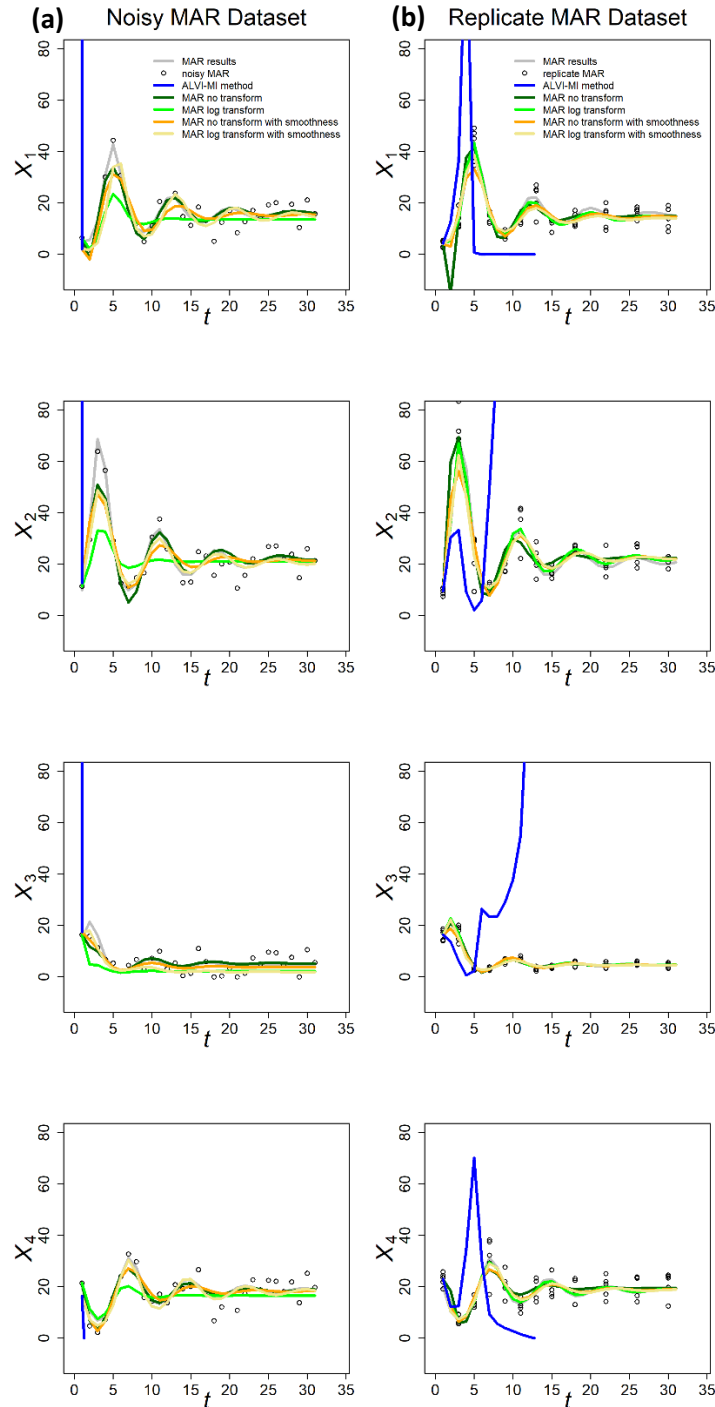


995

996 **Figure S7 - Multi-trophic dynamics for wolves, moose, and fir trees on Isle Royale from 1960 to 1994,**
997 **from McLaren & Peterson (1994) (McLaren & Peterson, 1994).** This panel is very similar to Figure 2 d) but
998 contains additional information. ALVI-MI estimates using all data are represented as blue lines. MAR
999 estimates are presented in green, orange and yellow. Red lines correspond to the estimates using ALVI-
1000 MI for two intervals, from 1959 to 1980 and from 1983 until the end of the series. This split was tested
1001 because around 1980 the wolves were exposed to a disease that drastically reduced their numbers, an
1002 event that dynamic models do not capture outside piecewise operation. MAR estimates are the same as
1003 in Figure 3.

1004

1005 **Figure S8**

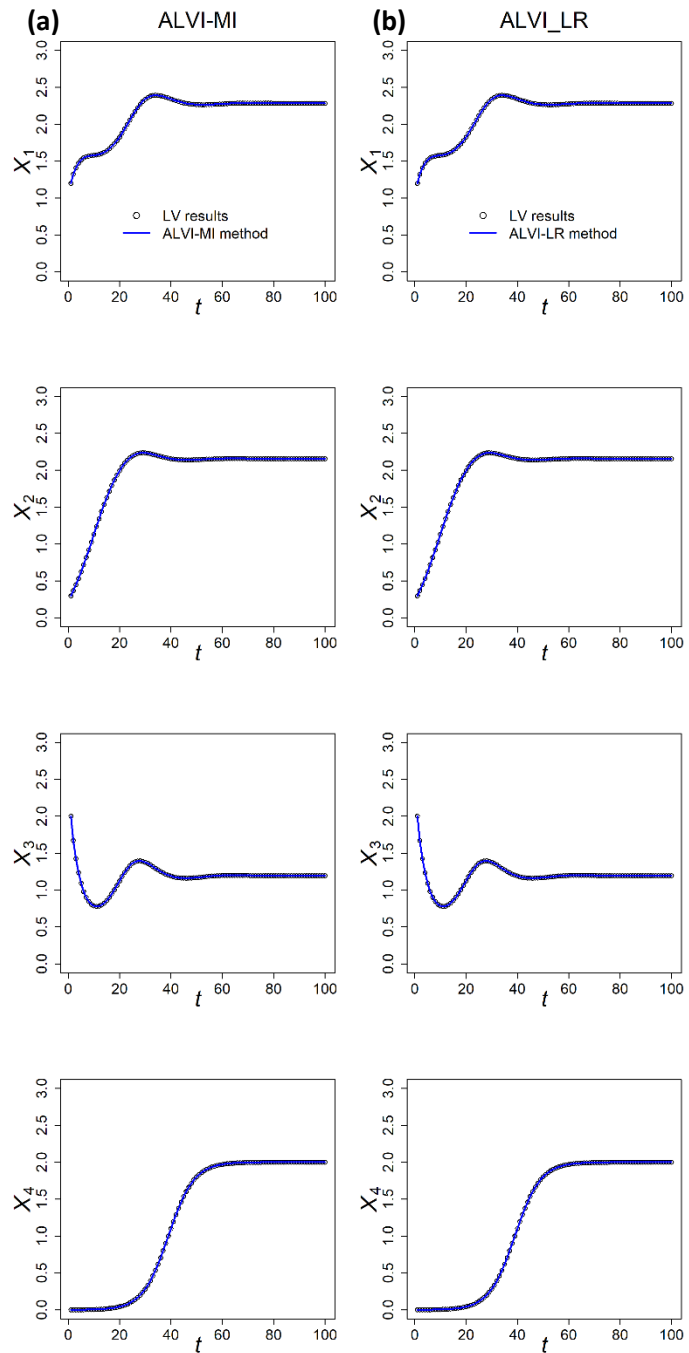


1006

1007 **Figure S8: ALVI-MI and MAR applied to an alternative sample of the same data presented in Figure 4, but with**
1008 **slightly changed noise.** Although the differences in noise are visually almost undetectable, very different results for
1009 the ALVI-MI fit are obtained if the same sample of spline points is used. However, if a new sample of spline points is
1010 determined, the fits are almost indistinguishable (not shown). See Text for further explanations.

1011 **Figure S9**

	Initial Conditions	ALVI-MI Estimates	ALVI-LR Estimates
X_1	1.2	1.2	1.2
X_2	0.3	0.3	0.3
X_3	2	2	2
X_4	0.001	0.001	0.001
	True Parameter values		
a_1	0.044	0.0440	0.0440
b_{11}	-0.08	-0.0800	-0.0800
b_{12}	0.02	0.0200	0.0200
b_{13}	0.08	0.0800	0.0800
b_{14}	0	0.0000	0.0000
a_1	0.216	0.2160	0.2160
b_{11}	-0.04	-0.0400	-0.0400
b_{12}	-0.08	-0.0800	-0.0800
b_{13}	0.04	0.0400	0.0400
b_{14}	0	0.0000	0.0000
a_2	0.116	0.1161	0.1161
b_{21}	-0.16	-0.1602	-0.1602
b_{22}	0.16	0.1601	0.1601
b_{23}	-0.08	-0.0799	-0.0799
b_{24}	0	0.0000	0.0000
a_4	0.2	0.2000	0.2000
b_{41}	0	-0.0003	-0.0003
b_{42}	0	0.0001	0.0001
b_{43}	0	0.0003	0.0003
b_{44}	-0.1	-0.0999	-0.0999



1012 **Figure S9: Estimates with alternative ALVI methods. Column a:** ALVI-MI and **Column b:** ALVI-LR with original
 1013 synthetic LV data. The fits are of high quality (ALVI-MI SSE = 1.162229e-05 and ALVI-LR SSE = 3.289283e-07) and the
 1014 parameter estimates are very close to the true parameters.

1015

1016 **4. Supplemental Tables**

1017

1018 **Table S1.1 – Synthetic LV data.** The data were generated with an LV system with four dependent
1019 variables with parameter values presented in Figure S1.

t	X_1	X_2	X_3	X_4
0	1.20000	0.30000	2.00000	0.00100
1	1.32083	0.37078	1.67253	0.00122
2	1.41112	0.44848	1.42395	0.00149
3	1.47488	0.53265	1.23448	0.00182
4	1.51746	0.62275	1.09051	0.00222
5	1.54431	0.71810	0.98219	0.00272
6	1.56033	0.81785	0.90227	0.00332
7	1.56959	0.92096	0.84535	0.00405
8	1.57534	1.02621	0.80746	0.00494
9	1.58006	1.13229	0.78565	0.00603
10	1.58566	1.23781	0.77778	0.00737
11	1.59356	1.34144	0.78227	0.00899
12	1.60486	1.44195	0.79788	0.01097
13	1.62038	1.53828	0.82360	0.01338
14	1.64076	1.62958	0.85846	0.01632
15	1.66648	1.71521	0.90140	0.01990
16	1.69783	1.79474	0.95118	0.02425
17	1.73496	1.86789	1.00625	0.02954
18	1.77779	1.93449	1.06475	0.03596
19	1.82599	1.99445	1.12450	0.04375
20	1.87894	2.04769	1.18311	0.05317
21	1.93567	2.09416	1.23810	0.06457
22	1.99490	2.13382	1.28715	0.07830
23	2.05502	2.16666	1.32828	0.09482
24	2.11427	2.19279	1.36011	0.11461
25	2.17079	2.21241	1.38190	0.13823
26	2.22286	2.22591	1.39369	0.16628
27	2.26899	2.23381	1.39615	0.19943
28	2.30811	2.23680	1.39047	0.23832
29	2.33961	2.23566	1.37821	0.28360
30	2.36336	2.23123	1.36102	0.33585
31	2.37966	2.22437	1.34057	0.39550
32	2.38913	2.21586	1.31834	0.46281
33	2.39266	2.20642	1.29562	0.53773
34	2.39123	2.19665	1.27343	0.61988
35	2.38587	2.18704	1.25253	0.70850
36	2.37758	2.17796	1.23345	0.80243
37	2.36727	2.16968	1.21655	0.90013
38	2.35575	2.16237	1.20201	0.99980
39	2.34368	2.15612	1.18987	1.09947
40	2.33162	2.15098	1.18009	1.19718
41	2.32001	2.14691	1.17257	1.29113
42	2.30917	2.14387	1.16713	1.37978
43	2.29935	2.14177	1.16358	1.46196
44	2.29070	2.14052	1.16170	1.53691
45	2.28329	2.14001	1.16126	1.60424
46	2.27717	2.14013	1.16203	1.66392
47	2.27231	2.14076	1.16377	1.71620
48	2.26866	2.14180	1.16626	1.76151

49	2.26613	2.14313	1.16930	1.80043
50	2.26462	2.14467	1.17270	1.83359
51	2.26400	2.14632	1.17627	1.86167
52	2.26415	2.14801	1.17987	1.88531
53	2.26494	2.14968	1.18337	1.90511
54	2.26623	2.15128	1.18668	1.92164
55	2.26790	2.15275	1.18970	1.93538
56	2.26982	2.15407	1.19237	1.94679
57	2.27189	2.15523	1.19467	1.95622
58	2.27400	2.15620	1.19658	1.96401
59	2.27609	2.15698	1.19809	1.97044
60	2.27808	2.15759	1.19921	1.97573
61	2.27991	2.15802	1.19998	1.98009
62	2.28156	2.15829	1.20042	1.98367
63	2.28299	2.15842	1.20058	1.98661
64	2.28418	2.15843	1.20049	1.98902
65	2.28515	2.15833	1.20021	1.99100
66	2.28588	2.15816	1.19977	1.99263
67	2.28640	2.15792	1.19922	1.99396
68	2.28673	2.15764	1.19859	1.99505
69	2.28688	2.15733	1.19793	1.99595
70	2.28688	2.15702	1.19726	1.99668
71	2.28675	2.15671	1.19660	1.99728
72	2.28653	2.15641	1.19598	1.99777
73	2.28623	2.15613	1.19541	1.99818
74	2.28588	2.15588	1.19491	1.99851
75	2.28551	2.15566	1.19448	1.99878
76	2.28512	2.15547	1.19411	1.99900
77	2.28473	2.15532	1.19383	1.99918
78	2.28436	2.15521	1.19361	1.99933
79	2.28402	2.15512	1.19346	1.99945
80	2.28371	2.15507	1.19337	1.99955
81	2.28344	2.15504	1.19333	1.99963
82	2.28322	2.15503	1.19334	1.99970
83	2.28303	2.15505	1.19338	1.99975
84	2.28289	2.15508	1.19346	1.99980
85	2.28279	2.15512	1.19355	1.99983
86	2.28272	2.15517	1.19367	1.99986
87	2.28269	2.15522	1.19378	1.99989
88	2.28269	2.15528	1.19391	1.99991
89	2.28270	2.15533	1.19403	1.99993
90	2.28274	2.15539	1.19414	1.99994
91	2.28279	2.15544	1.19425	1.99995
92	2.28286	2.15549	1.19434	1.99996
93	2.28292	2.15553	1.19442	1.99997
94	2.28299	2.15556	1.19449	1.99997
95	2.28306	2.15559	1.19455	1.99998
96	2.28313	2.15562	1.19459	1.99998
97	2.28320	2.15563	1.19462	1.99998
98	2.28325	2.15564	1.19464	1.99999
99	2.28330	2.15565	1.19465	1.99999
100	2.28335	2.15565	1.19465	1.99999
101	2.28338	2.15565	1.19464	1.99999

1020

1021

1022 **Table S1.2 – Noisy LV dataset.** From the synthetic data, generated with the LV system in Table S1.1, forty
1023 values were selected and random normal noise was added with mean 0 and standard deviation equal to
1024 20% of each variable mean.

t	X_1	X_2	X_3	X_4
1	1.14111	0.94716	1.92175	0.06256
3	1.36364	0.45349	1.45431	0.06645
5	1.22283	0.19582	1.33123	0.03641
8	1.47489	1.02716	0.78493	0.01857
13	1.68296	1.83783	0.72635	0.20645
17	1.87602	0.98039	1.33307	0.02590
19	2.23270	2.28656	0.88182	0.08578
24	2.46929	2.14703	1.42853	0.30301
26	2.12740	1.68427	1.24384	0.29370
30	2.93876	1.47752	1.47634	0.33315
31	1.60470	2.50979	0.99567	0.31345
32	2.64561	2.16238	1.21792	0.46671
33	2.16612	2.30080	1.25219	0.44936
40	2.90825	2.47699	1.42807	0.59721
42	2.16482	2.82494	1.06154	1.37928
44	2.86258	2.10103	1.23401	1.37031
46	2.30199	1.92131	1.06248	1.91622
49	1.46650	2.70250	0.96520	2.29495
52	2.07311	1.79577	1.33918	1.55672
53	1.52180	1.69359	1.07106	2.03041
55	2.34261	1.94310	1.50543	1.95215
56	3.07809	3.11285	1.29445	1.80656
62	1.30983	1.83120	1.16429	2.13265
63	2.70020	2.32062	1.30813	2.15788
64	1.68571	1.69570	1.19108	1.96368
67	3.06508	1.69727	1.30761	1.92002
69	2.87653	2.02694	1.10203	1.72796
71	1.92870	2.69214	0.69211	1.84299
73	2.17466	1.97223	1.23306	1.94046
75	2.25649	2.48677	1.35097	1.93681
77	2.12334	1.58311	0.96197	2.23364
78	3.38719	1.99818	0.93052	1.93376
79	2.33980	1.85041	1.09020	2.46553
83	1.97877	2.01006	1.07131	1.89393
84	2.55545	2.64187	1.29245	2.38731
87	2.36884	2.11299	1.22538	2.03970
90	2.25285	2.22309	1.43866	1.73292
91	2.24325	2.25534	1.58509	2.14186
93	2.47453	1.91424	1.19010	2.00689
95	1.82853	1.59450	1.18877	1.91223

1025

1026

1027 **Table S1.3 – Replicate LV dataset.** 11 points were selected from the synthetic data in Table S1.1. For each
 1028 time point, five observations were created by multiplying the original value by a normal random variable
 1029 with mean 1 and standard deviation 0.2.

t	X_1	X_2	X_3	X_4
1	1.079474	0.307892	1.968433	0.001177
1	1.228073	0.319118	1.767284	0.001143
1	1.001938	0.278408	2.035955	0.001019
1	1.151608	0.34439	2.049352	0.000994
1	1.106675	0.330651	1.634474	0.001462
6	1.409001	0.827832	1.033651	0.003136
6	1.292779	0.655128	0.840712	0.002841
6	1.186731	0.753583	0.964293	0.003671
6	1.501709	0.702128	0.846646	0.002595
6	1.600803	0.778034	1.191477	0.003243
11	1.553429	1.585188	0.501395	0.008283
11	1.420026	1.565142	0.721249	0.009309
11	1.599542	0.772725	0.708241	0.004804
11	1.642382	1.707548	0.424371	0.00881
11	1.142047	1.689577	0.992655	0.00613
18	1.644046	1.84217	0.929999	0.044789
18	1.780007	1.601518	1.134646	0.030728
18	1.710696	1.833336	1.096591	0.023249
18	1.331607	2.483739	0.591253	0.029612
18	1.357593	1.968956	1.2092	0.017282
26	2.560157	2.19029	1.010075	0.084837
26	2.478865	2.142542	1.441704	0.160821
26	2.920662	2.166496	1.227926	0.177711
26	1.783109	1.700208	1.23534	0.205836
26	1.809358	2.395388	1.056138	0.105769
33	2.230052	2.81995	1.194645	0.540826
33	1.692464	2.038453	1.113628	0.428625
33	2.981684	2.168246	1.363852	0.486374
33	2.095491	1.58247	1.231077	0.474691
33	2.875616	2.102594	1.238574	0.612312
41	1.970818	2.333379	1.04227	1.296557
41	1.611026	1.927812	1.114056	1.438404
41	2.112647	2.279129	1.081485	0.99357
41	2.652937	1.953001	1.498288	1.30327
41	2.261238	2.34695	1.170614	1.306394
51	2.079631	1.228253	1.209477	2.075646
51	1.819921	1.666988	1.070121	1.644326
51	2.454389	2.202211	1.41537	2.439964
51	2.256488	2.134285	1.231388	1.709964
51	2.213276	2.102253	1.234634	1.884561
60	2.165804	2.18245	1.155609	2.283614
60	2.279157	1.885294	1.137585	1.698353
60	2.36829	2.52211	1.349542	2.049815
60	2.234632	2.281866	1.184983	1.165772
60	2.439225	1.996245	1.501995	2.82506
70	1.719859	2.411846	1.227642	1.786889
70	2.570555	2.462909	1.1756	1.878108
70	1.790253	1.887747	1.142104	1.895824
70	2.723166	2.042576	1.65201	1.8243
70	3.007495	2.227207	0.93787	2.226255
80	2.296889	2.001374	1.396972	2.20474
80	2.749139	1.714839	1.059392	1.594539
80	0.9041	2.298373	1.489557	2.267917
80	1.676455	1.788501	0.766479	2.491924
80	2.131265	1.738182	1.403356	1.898047

1030

1031

1032 **Table S1.4 – MAR estimates for the noisy LV dataset in Figs. 1 and S4**

1033

	MAR without transformation	MAR with log transformation	MAR with smoothing	MAR with log transformation and smoothing
β_{11}	-0.09089	-0.03360	1.06000	1.02000
β_{21}	0.27209	0.43560	0.13800	0.07940
β_{31}	-0.21012	-0.43570	0.07490	-0.10500
β_{41}	0.12619	0.42840	-0.33900	-0.22100
β_{12}	0.43295	0.24860	-0.05360	-0.01990
β_{22}	0.47619	0.62610	0.91000	0.94700
β_{32}	0.18150	0.17100	0.01730	0.12000
β_{42}	0.00788	0.23190	0.22900	0.19400
β_{13}	-0.20530	-0.01010	-0.01980	-0.01510
β_{23}	-0.28009	-0.33410	-0.07090	-0.07300
β_{33}	0.50995	0.57920	0.86700	0.90900
β_{43}	-0.00869	-0.02950	-0.03190	0.04240
β_{14}	0.07772	0.00960	-0.00954	-0.00439
β_{24}	0.06929	0.02720	-0.02780	-0.01270
β_{34}	0.00957	0.01530	-0.00911	-0.00807
β_{44}	0.93364	0.88740	0.98800	0.95600
α_1	1.70925	0.63980	0.01030	0.00614
α_2	0.67025	-0.04340	0.01430	0.01390
α_3	0.67931	0.29480	-0.00571	-0.00386
α_4	-0.17950	-0.45920	0.02010	0.04190
δ_1	0.22184	0.05270	0.00005	0.00001
δ_2	0.19571	0.07520	0.00014	0.00003
δ_3	0.01924	0.01480	0.00049	0.00028
δ_4	0.04708	0.07570	0.00005	0.00010

1034

1035

1036

1037 **Table S1.5 – MAR estimates for the replicate LV dataset in Figs. 1 and S4**

1038

	MAR without transformation	MAR with log transformation	MAR with smoothing	MAR with log transformation and smoothing
β_{11}	0.77800	0.76500	1.01000	1.07000
β_{21}	0.07760	0.21000	0.00367	0.07720
β_{31}	-0.08920	-0.31200	-0.01630	-0.10700
β_{41}	0.09140	0.13400	0.23200	0.38400
β_{12}	0.13500	0.11600	-0.00781	-0.00292
β_{22}	0.87500	0.80500	0.93600	0.88300
β_{32}	0.14900	0.22500	0.10400	0.14800
β_{42}	-0.00518	0.17400	-0.08900	0.06510
β_{13}	0.24800	0.22200	-0.00982	-0.00328
β_{23}	-0.20400	-0.18300	-0.10300	-0.05940
β_{33}	0.94400	1.13000	0.86500	0.98300
β_{43}	0.01090	0.17500	-0.02880	0.04240
β_{14}	-0.02600	-0.01100	-0.02170	-0.00991
β_{24}	-0.00159	-0.00250	-0.00530	-0.00583
β_{34}	-0.00022	-0.00299	0.00004	-0.00317
β_{44}	0.97200	0.93000	0.96400	0.93000
α_1	0.02400	0.01470	0.00925	0.00610
α_2	0.02610	0.03690	0.01980	0.02190
α_3	-0.02240	-0.01470	-0.00880	-0.00573
α_4	0.02160	0.10700	0.02500	0.09300
δ_1	0.00557	0.00089	0.00008	0.00001
δ_2	0.00243	0.00073	0.00037	0.00009
δ_3	0.00001	0.00004	0.00024	0.00043
δ_4	0.00559	0.00267	0.00029	0.00022

1039

1040

1041 **Table S2.1 - MAR estimates for the noisy dataset in Fig. S3**
 1042

	MAR without transformation	MAR with log transformation	MAR with smoothing	MAR with log transformation and smoothing
β_{11}	0.23009	0.01769	0.90100	0.93300
β_{21}	0.11849	0.16822	-0.18000	-0.10300
β_{31}	-0.18449	-0.35186	-0.26400	-0.61400
β_{41}	0.14318	2.18894	0.40800	1.05000
β_{12}	0.34796	0.23560	0.03870	0.01540
β_{22}	0.84182	0.75617	1.05000	0.94800
β_{32}	0.19214	0.17605	0.18600	0.20600
β_{42}	-0.03130	0.01950	-0.15300	-0.06350
β_{13}	0.21488	0.18478	-0.00110	0.00908
β_{23}	-0.04121	0.02412	-0.10500	-0.04810
β_{33}	0.78353	0.91531	0.87900	1.02000
β_{43}	0.01040	1.00320	-0.02870	0.05100
β_{14}	0.03493	0.01754	-0.00471	-0.00229
β_{24}	-0.00776	0.00115	-0.00050	-0.00170
β_{34}	-0.00548	0.00382	0.01160	0.01240
β_{44}	0.97027	0.81890	0.95100	0.91800
α_1	-0.00061	0.00073	0.01020	0.00586
α_2	0.02045	0.03086	0.02000	0.02130
α_3	-0.01764	-0.01187	-0.00431	-0.00279
α_4	0.02229	0.04416	0.02100	0.06150
δ_1	0.01953	0.00318	0.00001	0.00000
δ_2	0.01746	0.01414	0.00002	0.00000
δ_3	0.00158	0.00121	0.00002	0.00004
δ_4	0.00363	0.39522	0.00038	0.00007

1043
 1044
 1045

1046 **Table S2.2 - MAR estimates for the replicate dataset in Fig. S3**
 1047

	MAR without transformation	MAR with log transformation	MAR with smoothing	MAR with log transformation and smoothing
β_{11}	0.83000	0.81000	0.88600	0.87400
β_{21}	-0.06680	0.02220	-0.05350	0.04420
β_{31}	-0.12600	-0.35400	-0.09890	-0.39000
β_{41}	0.23600	0.37800	0.23500	0.41700
β_{12}	0.07710	0.04860	0.05470	0.04280
β_{22}	0.97800	0.89700	0.97600	0.90000
β_{32}	0.16200	0.19000	0.13100	0.19100
β_{42}	-0.07360	0.10500	-0.08250	0.08650
β_{13}	0.12900	0.11100	0.05310	0.06250
β_{23}	-0.06320	-0.03200	-0.09610	-0.06310
β_{33}	0.87400	1.02000	0.89000	1.06000
β_{43}	-0.04980	0.06360	-0.02580	0.08010
β_{14}	-0.00609	-0.00121	-0.00871	-0.00304
β_{24}	-0.00592	-0.00445	-0.00621	-0.00488
β_{34}	-0.00154	0.00406	-0.00073	0.00508
β_{44}	0.96100	0.92600	0.96300	0.92500
α_1	0.02250	0.01380	0.01240	0.00749
α_2	0.03120	0.04080	0.02220	0.02350
α_3	-0.02690	-0.01750	-0.00777	-0.00512
α_4	0.01850	0.10700	0.02520	0.09450
δ_1	0.00052	0.00004	0.00007	0.00001
δ_2	0.00029	0.00007	0.00004	0.00001
δ_3	0.00002	0.00000	0.00002	0.00001
δ_4	0.00219	0.00114	0.00039	0.00014

1048

1049

1050

1051 **Table S3.1 – Initial conditions, parameter values and estimates for a four-variable LV system that**
 1052 **converges to a stable steady state, as presented in Fig. S5**

	Initial Condition	Estimate			
X_1	1.2	1.20993			
X_2	0.3	0.31991			
X_3	2	1.82146			
X_4	0.001	0.00104			
	True Parameter Value	ALVI-MI		MAR with log MAR	transformation
a_1	0.044	0.04398	β_{11}	0.89300	0.88800
b_{11}	-0.08	-0.07995	β_{21}	-0.07190	0.08510
b_{12}	0.02	0.01997	β_{31}	-0.17500	-0.43000
b_{13}	0.08	0.07999	β_{41}	0.19200	0.42800
b_{14}	0	-0.00001	β_{12}	0.04840	0.03110
a_2	0.216	0.21599	β_{22}	0.99000	0.91100
b_{21}	-0.04	-0.04000	β_{32}	0.16800	0.18000
b_{22}	-0.08	-0.08001	β_{42}	-0.05440	0.11000
b_{23}	0.04	0.04002	β_{13}	0.06600	0.06700
b_{24}	0	0.00000	β_{23}	-0.07540	-0.06800
a_3	0.116	0.11600	β_{33}	0.91100	1.03000
b_{31}	-0.16	-0.15993	β_{43}	-0.00639	0.03510
b_{32}	0.16	0.15998	β_{14}	-0.00719	-0.00229
b_{33}	-0.08	-0.08007	β_{24}	-0.00739	-0.00763
b_{34}	0	-0.00001	β_{34}	0.00126	0.00751
a_4	0.2	0.20002	β_{44}	0.96300	0.92400
b_{41}	0	-0.00001	α_1	0.07320	0.05740
b_{42}	0	0.00000	α_2	0.28900	0.01420
b_{43}	0	0.00000	α_3	0.14200	0.20700
b_{44}	-0.1	-0.10000	α_4	-0.22800	-0.38500
			δ_1	0.00014	0.00006
			δ_2	0.00004	0.00031
			δ_3	0.00064	0.00020
			δ_4	0.00032	0.00032

1053

1054

1055 **Table S3.2 – Initial conditions, parameter values and estimates for a four-variable LV system exhibiting**
 1056 **damped oscillations as presented in Fig. S5**

	Initial Condition	Estimate			
X_1	0.3	0.3			
X_2	0.3	0.3			
X_3	0.4	0.4			
X_4	0.6	0.6			
	True Parameter Value	ALVI-MI		MAR	MAR with log transformation
a_1	0.3	0.30245	β_{11}	1.05000	0.98800
b_{11}	-0.3	-0.30328	β_{21}	0.14700	0.01610
b_{12}	-0.27	-0.27098	β_{31}	-0.25300	-0.07880
b_{13}	-0.6	-0.60467	β_{41}	-0.23100	-0.05360
b_{14}	-0.045	-0.04708	β_{12}	-0.08350	-0.11600
a_2	0.4	0.40299	β_{22}	0.88800	0.92200
b_{21}	0.2	0.19592	β_{32}	0.04880	-0.03410
b_{22}	-0.4	-0.40118	β_{42}	-0.00228	-0.07260
b_{23}	-0.4	-0.40578	β_{13}	0.03660	-0.04410
b_{24}	-0.6	-0.60250	β_{23}	-0.07340	-0.06540
a_3	0.7	0.71519	β_{33}	0.71200	0.89700
b_{31}	-2.38	-2.39901	β_{43}	-0.20500	-0.02050
b_{32}	0.35	0.34344	β_{14}	-0.10400	-0.01220
b_{33}	-2.8	-2.82806	β_{24}	-0.12800	-0.03300
b_{34}	0.35	0.33650	β_{34}	0.03690	-0.14100
a_4	0.6	0.60791	β_{44}	0.96100	0.88100
b_{41}	-0.96	-0.97003	α_1	0.04710	-0.22200
b_{42}	-0.24	-0.24338	α_2	0.08150	-0.17700
b_{43}	-0.96	-0.97466	α_3	0.06750	-0.56600
b_{44}	-0.6	-0.60704	α_4	0.08660	-0.38800
			δ_1	0.00000	0.00004
			δ_2	0.00002	0.00016
			δ_3	0.00015	0.00271
			δ_4	0.00015	0.00065

1057

1058

1059

1060 **Table S3.3 – Initial conditions, parameter values and estimates for a four-variable LV system**
 1061 **displaying initially erratic oscillations, but converging to a limit cycle as presented in Fig. S5**

	Initial Condition	Estimate			
X_1	0.3	0.3			
X_2	0.3	0.3			
X_3	0.4	0.4			
X_4	0.6	0.6			
	True				
	Parameter				
	Value	ALVI-MI		MAR	MAR with log transformation
a_1	1	0.98743	β_{11}	0.87600	0.94673
b_{11}	-1	-0.99101	β_{21}	0.26700	0.19269
b_{12}	-1.09	-1.08083	β_{31}	-0.11100	-0.77772
b_{13}	-1.52	-1.51341	β_{41}	-0.17600	-0.18648
b_{14}	0	0.01346	β_{12}	-0.13600	-0.11408
a_2	0.72	0.72230	β_{22}	0.90900	0.88847
b_{21}	0	-0.00274	β_{32}	0.03980	0.59990
b_{22}	-0.72	-0.72223	β_{42}	0.06220	0.11789
b_{23}	-0.3168	-0.31966	β_{13}	-0.44100	0.00528
b_{24}	-0.9792	-0.97942	β_{23}	0.07280	0.01411
a_3	1.53	1.53466	β_{33}	0.86500	0.93625
b_{31}	-3.672	-3.67283	β_{43}	-0.03720	-0.00153
b_{32}	0	-0.00071	β_{14}	0.36500	0.00392
b_{33}	-1.53	-1.52615	β_{24}	-0.19100	-0.17211
b_{34}	-0.7191	-0.73207	β_{34}	-0.02410	0.33444
a_4	1.27	1.27361	β_{44}	0.93800	0.97758
b_{41}	-1.5367	-1.53884	α_1	0.02540	-0.12418
b_{42}	-0.6477	-0.64952	α_2	0.01980	-0.00507
b_{43}	-0.4445	-0.44519	α_3	0.03310	-0.28038
b_{44}	-1.27	-1.27581	α_4	0.04680	-0.16530
			δ_1	0.00001	0.00057
			δ_2	0.00001	0.00012
			δ_3	0.00013	0.00112
			δ_4	0.00005	0.00022

1062

1063

1064 **Table S3.4 – Initial conditions, parameter values and estimates for a four-variable LV system**
 1065 **displaying damped oscillations as presented in Fig. S5**

	Initial Condition	Estimate			
X_1	0.3	0.3			
X_2	0.3	0.3			
X_3	0.4	0.4			
X_4	0.6	0.6			
	True				
	Parameter				
	Value	ALVI-MI		MAR	MAR with log transformation
a_1	0.3	0.27831	β_{11}	0.96100	1.02036
b_{11}	-0.3	-0.30250	β_{21}	0.17900	0.10195
b_{12}	-0.27	-0.25033	β_{31}	-0.15100	-0.09094
b_{13}	-0.6	-0.59705	β_{41}	-0.20000	-0.04845
b_{14}	-0.045	-0.00606	β_{12}	-0.07430	-0.05021
a_2	0.4	0.36675	β_{22}	0.94200	0.98452
b_{21}	0.2	0.19193	β_{32}	-0.05210	-0.02708
b_{22}	-0.4	-0.36840	β_{42}	-0.01790	-0.03530
b_{23}	-0.4	-0.40045	β_{13}	-0.09270	0.00210
b_{24}	-0.6	-0.53672	β_{23}	0.00417	-0.00943
a_3	0.7	0.64222	β_{33}	0.82100	1.05860
b_{31}	-2.38	-2.37532	β_{43}	-0.15000	0.02181
b_{32}	0.35	0.39856	β_{14}	-0.00981	0.02438
b_{33}	-2.45	-2.42900	β_{24}	-0.10400	-0.00099
b_{34}	0.35	0.44414	β_{34}	-0.12500	-0.32180
a_4	0.6	0.55927	β_{44}	0.96300	0.89997
b_{41}	-0.96	-0.96318	α_1	0.05890	0.00507
b_{42}	-0.24	-0.20356	α_2	0.01900	0.11510
b_{43}	-0.96	-0.95279	α_3	0.13300	-0.41484
b_{44}	-0.3	-0.22817	α_4	0.09470	-0.17804
			δ_1	0.00003	0.00018
			δ_2	0.00010	0.00101
			δ_3	0.00014	0.00935
			δ_4	0.00008	0.00056

1066

1067

1068 **Table S3.5 – Initial conditions, parameter values and estimates for a four-variable LV system**
 1069 **displaying deterministic chaos (chaos 1) as presented in Fig. S5**

	Initial Condition	Estimate			
X_1	0.3	0.3			
X_2	0.3	0.3			
X_3	0.4	0.4			
X_4	0.6	0.6			
	True				
	Parameter				
	Value	ALVI-MI		MAR	MAR with log transformation
a_1	1	1.01561	β_{11}	0.85600	1.00059
b_{11}	-1	-1.01294	β_{21}	0.26400	0.20355
b_{12}	-1.09	-1.10110	β_{31}	-0.16900	-0.79454
b_{13}	-1.52	-1.52409	β_{41}	-0.21800	-0.23777
b_{14}	0	-0.01741	β_{12}	-0.14300	-0.10901
a_2	0.72	0.72904	β_{22}	0.88800	0.88053
b_{21}	0	-0.00777	β_{32}	0.01640	0.59018
b_{22}	-0.72	-0.72674	β_{42}	0.02020	0.08015
b_{23}	-0.3168	-0.32100	β_{13}	-0.47500	-0.03240
b_{24}	-0.9792	-0.98775	β_{23}	0.10500	0.03330
a_3	1.53	1.48293	β_{33}	1.07000	0.76513
b_{31}	-3.5649	-3.52631	β_{43}	0.06420	-0.03995
b_{32}	0	0.03573	β_{14}	0.38200	0.17765
b_{33}	-1.53	-1.50898	β_{24}	-0.24000	-0.23158
b_{34}	-0.7191	-0.67335	β_{34}	-0.21500	0.90744
a_4	1.27	1.25074	β_{44}	0.80900	1.08027
b_{41}	-1.5367	-1.52127	α_1	0.03510	0.01644
b_{42}	-0.6477	-0.63310	α_2	0.04320	-0.01131
b_{43}	-0.4445	-0.43667	α_3	0.11000	-0.11896
b_{44}	-1.27	-1.25061	α_4	0.11500	-0.24402
			δ_1	0.00002	0.00114
			δ_2	0.00002	0.00015
			δ_3	0.00022	0.00191
			δ_4	0.00006	0.00024

1070

1071

1072 **Table S3.6 – Initial conditions, parameter values and estimates for a four-variable LV system**
 1073 **displaying deterministic chaos (chaos 2) as presented in Fig. S5**

	Initial Condition	Estimate			
X_1	0.3	0.3			
X_2	0.3	0.3			
X_3	0.4	0.4			
X_4	0.6	0.6			
	True				
	Parameter				
	Value	ALVI-MI		MAR	MAR with log transformation
a_1	0.3	0.29277	β_{11}	1.05000	0.97700
b_{11}	-0.3	-0.25379	β_{21}	0.00560	0.02030
b_{12}	-0.27	-0.28051	β_{31}	-0.31600	-0.08880
b_{13}	-0.6	-0.55795	β_{41}	-0.23000	-0.06400
b_{14}	-0.045	-0.07045	β_{12}	-0.09830	-0.01780
a_2	0.4	0.39048	β_{22}	0.99000	1.00000
b_{21}	0.2	0.26661	β_{32}	0.08310	-0.02620
b_{22}	-0.4	-0.41601	β_{42}	-0.01940	-0.01610
b_{23}	-0.4	-0.33988	β_{13}	-0.00541	-0.06900
b_{24}	-0.6	-0.63818	β_{23}	-0.12100	-0.07290
a_3	0.8	0.77659	β_{33}	0.70200	0.87400
b_{31}	-2.38	-2.23922	β_{43}	-0.18000	-0.05050
b_{32}	0.35	0.31928	β_{14}	-0.08040	0.05960
b_{33}	-2.45	-2.32117	β_{24}	0.00941	-0.02020
b_{34}	0.35	0.27470	β_{34}	0.07400	-0.14500
a_4	0.6	0.58578	β_{44}	0.97300	0.94100
b_{41}	-0.96	-0.86962	α_1	0.05270	-0.11500
b_{42}	-0.24	-0.26048	α_2	0.02820	-0.09870
b_{43}	-0.96000	-0.87774	α_3	0.07760	-0.53700
b_{44}	-0.30000	-0.34969	α_4	0.10200	-0.28000
			δ_1	0.00000	0.00006
			δ_2	0.00002	0.00023
			δ_3	0.00009	0.00115
			δ_4	0.00007	0.00025

1074

1075

1076

1077 **Table S4.1 – ALVI-MI estimates for five experimental data examples from (Mühlbauer et al., 2020).** Data
 1078 came from experiments described in (Gause, 1934), (McLaren & Peterson, 1994) and (Huffaker et al.,
 1079 1963). See R package gauseR (Mühlbauer et al., 2020) for datasets “gause_1934_science_f02_03”,
 1080 “gause_1934_book_f32”, “mclaren_1994_f03” and “huffaker_1963” for details on observations.
 1081 Parameter estimates from Mühlbauer *et al.* can also be found in Table 2 in their paper.

1082

Example 1 - *Paramecium caudatum* in monoculture. The slopes were estimated from an 8DF-spline from data without log transformation and ALVI-MI using a subsample of spline points at the 3rd and 12th days.

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	1.259	0.92289	0.33611
b_{11}	-0.005	-0.00456	0.00044

Example 2 - *Paramecium caudatum* and *Paramecium aurelia* in a mixed population competition study. ALVI-MI was estimated from 10DF and 7DF-splines for *P. caudatum* and *P. Aurelia*, respectively. Spline points were taken at days 4, 8 and 11.

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	1.259	0.98677	0.27223
b_{11}	-0.005	-0.00409	0.00091
b_{12}	1.259	-0.00649	1.26549
a_2	-0.005	0.79868	0.80368
b_{21}	1.259	-0.00136	1.26036
b_{22}	-0.005	-0.00536	0.00036

Example 3 - Predator-prey interactions between *Didinium nasutum* and *Paramecium caudatum*. ALVI-MI estimates were calculated using 14DF and 10DF-splines, respectively, using a subsample of the 122nd, 140th, 168th points of the second spline.

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	1.099	1.70706	0.60806
b_{11}	-0.013	-0.03887	0.02587
b_{12}	-0.078	-0.11360	0.03560
a_2	-0.89	-1.27639	0.38639
b_{21}	0.084	0.14275	0.05875
b_{22}	-0.002	0.01565	0.01765

Example 4 - Multi-trophic dynamics for wolves, moose, and fir trees. ALVI-MI estimates were calculated using log-abundances and 8DF-splines. Spline points were chosen as a subsample corresponding to the years 1973, 1978, 1979 and 1982.

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	0.01	-1.901823	1.91182
b_{11}	-0.003	0.028812	0.03181
b_{12}	0.00004	0.000003	0.00004
b_{13}	0	2.754448	2.75445
a_2	2.021	0.331244	1.68976
b_{21}	-0.088	-0.006836	0.08116
b_{22}	0	-0.000107	0.00011
b_{23}	0.002	-0.090569	0.09257
a_3	0.238	2.779411	2.54141
b_{31}	0	-0.051545	0.05154
b_{32}	-0.0002	0.000494	0.00069
b_{33}	-0.139	-4.693609	4.55461

1083

1084

Example 5 - Predator-prey interactions between *E. sexmaculatus* and *T. occidentalis*. ALVI-MI estimates were calculated using 15DF and 20DF-splines, respectively. The splines were constructed using log-abundances of the dependent variables, using a subsample of spline points corresponding to the 17th, 48th and 55th datapoints.

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	0.187	0.11148	0.07552
b_{11}	0	0.00003	0.00003
b_{12}	-0.028	-0.02960	0.00160
a_2	-0.377	-0.80007	0.42307
b_{21}	0.0012	0.00251	0.00131
b_{22}	-0.024	-0.03144	0.00744

1085

1086

1087 **Table S4.2 – ALVI-LR estimates for five experimental data sets from (Mühlbauer et al., 2020).** Data came
 1088 from (Gause, 1934), (McLaren & Peterson, 1994) and (Huffaker et al., 1963) experiments. See R package
 1089 gauseR (Mühlbauer et al., 2020) datasets “gause_1934_science_f02_03”, “gause_1934_book_f32”,
 1090 ”mclaren_1994_f03” and “huffaker_1963” for details on observations. Parameter estimates from
 1091 Mühlbauer *et al.* can also be found in Table 2 of their paper.

Example 1 - *Paramecium caudatum* in monoculture, analyzed with 8DF-spline

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	1.259	0.93948	0.31952
b_{11}	-0.005	-0.00465	0.00035

Example 2 - *Paramecium caudatum* and *Paramecium aurelia* in mixed population, analyzed with 10DF and 7DF-splines

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	1.259	0.85524	0.40376
b_{11}	-0.005	-0.00289	0.00211
b_{12}	-0.008	-0.00580	0.00220
a_2	1.026	0.84423	0.18177
b_{21}	-0.002	-0.00187	0.00013
b_{22}	-0.007	-0.00553	0.00147

Example 3 - Predator-prey interactions between *Didinium nasutum* and *Paramecium caudatum*, analyzed with 14DF and 10DF-splines

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	1.099	0.45652	0.64248
b_{11}	-0.013	0.02117	0.03417
b_{12}	-0.078	-0.11495	0.03695
a_2	-0.89	-0.98922	0.09922
b_{21}	0.084	0.16549	0.08149
b_{22}	-0.002	-0.01146	0.00946

Example 4 - Multi-trophic dynamics for wolves, moose, and fir trees, analyzed with 28DF, 24DF and 28DF-splines

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	0.01	-0.06509	0.07509
b_{11}	-0.003	0.00164	0.00464
b_{12}	0.00004	0.00007	0.00003
b_{13}	0	-0.13411	0.13411
a_2	2.021	0.20754	1.81346
b_{21}	-0.088	-0.00483	0.08317
b_{22}	0	-0.00009	0.00009
b_{23}	0.002	0.06010	0.05810
a_3	0.238	-0.08580	0.32380
b_{31}	0	0.00343	0.00343
b_{32}	-0.0002	-0.00020	0.00000
b_{33}	-0.139	0.43352	0.57252

Example 5 - Predator-prey interactions between *E. sexmaculatus* and *T. occidentalis*, analyzed with 15DF and 20DF-splines

	Mühlbauer <i>et al.</i>	Estimate	Absolute Difference
a_1	0.344	0.03525	0.30875
b_{11}	0	0.00038	0.00038
b_{12}	-0.059	-0.03619	0.02281
a_2	-0.236	-0.44687	0.21087
b_{21}	0.0005	0.00159	0.00109
b_{22}	0	-0.03540	0.03540

1092

1093 **Table S4.3 – MAR estimates for five experimental data sets from (Mühlbauer et al., 2020).** Data came
 1094 from (Gause, 1934), (McLaren & Peterson, 1994) and (Huffaker et al., 1963) experiments. See R package
 1095 *gauseR* (Mühlbauer et al., 2020) for datasets “*gause_1934_science_f02_03*”, “*gause_1934_book_f32*”,
 1096 “*mclaren_1994_f03*” and “*huffaker_1963*” for details on observations. Parameter estimates from
 1097 Mühlbauer *et al.* can also be found in Table 2 on their paper.

1098 Example 1 - *Paramecium caudatum* in monoculture.

	MAR	MAR log transformation	MAR with data smoothing	MAR log transformation with smoothing
β_1	0.90000	0.85550	0.90400	0.74718
α_1	9.93000	0.15590	10.22600	0.21106
δ_1	287.57000	0.06920	126.15300	0.00525

1099

1100 Example 2 - *Paramecium caudatum* and *Paramecium aurelia* in coculture

	MAR	MAR log transformation	MAR with data smoothing	MAR log transformation with smoothing
β_{11}	0.81800	0.73590	0.91500	0.98309
β_{21}	0.07170	0.03240	0.11600	0.03625
β_{12}	-0.16030	-0.07020	-0.16900	-0.26372
β_{22}	0.82140	0.71450	0.87700	0.73153
α_1	27.54890	1.38260	0.91000	0.07810
α_2	20.30740	1.28150	5.82700	0.15430
δ_1	262.75870	0.17190	113.99000	0.03530
δ_2	253.59620	0.03470	11.12100	0.00202

1101

1102 Example 3 - Predator-prey interactions between *Didinium nasutum* and *Paramecium caudatum*

	MAR	MAR log transformation	MAR with data smoothing	MAR log transformation with smoothing
β_{11}	0.52400	0.80250	0.37800	0.86900
β_{21}	0.57700	0.10510	0.52100	0.18700
β_{12}	-0.57200	-0.21950	-0.45100	-0.25400
β_{22}	0.63300	0.59490	0.69100	0.76500
α_1	12.93300	-1.11100	0.40300	-0.75400
α_2	-2.76100	0.06450	0.23900	0.70000
δ_1	94.22700	10.97630	140.72600	7.15700
δ_2	27.95400	23.00840	83.62600	9.82300

1103

1104

1105

1106

1107 Example 4 - Multi-trophic dynamics for wolves, moose, and fir trees

	MAR	MAR log transformation	MAR with data smoothing	MAR log transformation with smoothing
β_{11}	0.62900	0.69390	1.13000	1.17360
β_{21}	-4.15000	-0.04540	-3.48000	-0.04426
β_{31}	-0.00083	-0.02210	0.00137	0.07236
β_{12}	0.00517	0.07570	0.00148	0.01665
β_{22}	0.84800	0.86270	0.90300	0.89080
β_{32}	-0.00007	-0.12840	-0.00009	-0.13234
β_{13}	-27.30000	-0.34650	10.70000	0.22128
β_{23}	96.60000	0.08020	245.00000	0.13054
β_{33}	0.89000	0.89160	1.11000	1.06497
α_1	15.90000	0.12600	-0.35700	-0.02613
α_2	241.00000	1.18230	27.80000	0.02986
α_3	0.14200	0.85670	-0.00485	-0.01065
δ_1	32.30000	0.06210	3.88000	0.00553
δ_2	15000.00000	0.01080	1290.00000	0.00127
δ_3	0.00382	0.02160	0.00098	0.00591

1108

1109 Example 5 - Predator-prey interactions between *E. sexmaculatus* and *T. occidentalis*

	MAR	MAR log transformation	MAR with data smoothing	MAR log transformation with smoothing
β_{11}	1.01500	0.88120	1.11000	1.05163
β_{21}	0.00900	0.53850	0.00718	0.47337
β_{12}	-17.79700	-0.09960	-16.00000	-0.10626
β_{22}	0.56200	0.75520	0.73200	0.86330
α_1	94.24600	0.80680	0.47700	-0.00088
α_2	-1.28500	-2.86420	-0.02580	-0.01771
δ_1	15440.27300	0.13470	1360.00000	0.02078
δ_2	5.55100	0.33790	1.93000	0.08104

1110

1111

1112

1113

1114 **Table S5.1 – Synthetic MAR data**

MAR Results

t	X_1	X_2	X_3	X_4
1	5	10	15	20
2	5.62418	35.2103	21.4125	10.3696
3	13.077	68.6603	16.2175	6.98986
4	30.8946	57.2944	7.75681	8.13234
5	42.7475	27.4855	3.71568	13.9927
6	31.3421	13.2593	2.63367	24.3619
7	16.5916	9.92497	3.02139	31.2622
8	9.868	12.4469	4.56419	27.4996
9	8.72248	20.3998	6.5998	19.5587
10	11.2768	30.657	7.33895	14.4998
11	16.7965	33.724	6.20514	13.314
12	21.9882	27.4413	4.66674	15.2345
13	22.1248	19.9474	3.79027	19.1227
14	18.0405	16.0646	3.69825	22.5386
15	14.0932	15.9557	4.22409	22.9275
16	12.3388	18.7495	5.03802	20.5772
17	12.7869	22.8183	5.59201	17.8653
18	14.8002	25.3774	5.51118	16.4484
19	17.1129	24.6804	4.99159	16.6605
20	18.1251	21.9731	4.48974	18.0462
21	17.3171	19.5701	4.27918	19.6396
22	15.6847	18.6862	4.39198	20.3941
23	14.456	19.3693	4.70288	19.9587
24	14.1822	20.9554	4.99681	18.8896
25	14.7787	22.3651	5.08888	17.9987
26	15.7517	22.7146	4.95911	17.7463
27	16.4482	21.9861	4.74141	18.1246
28	16.4701	20.8982	4.58774	18.7896
29	15.9468	20.1878	4.56995	19.2878
30	15.3325	20.1628	4.6678	19.3417
31	15.0101	20.6857	4.80099	19.0122

1115

1116

1117 **Table S5.2 –Synthetic MAR data with added noise (noisy MAR)**

Noisy MAR data

t	X_1	X_2	X_3	X_4
1	0.7537	5.7537	10.7537	15.7537
2	3.6789	33.265	19.4672	8.42431
3	22.0484	77.6318	25.189	15.9613
4	27.8408	54.24056	4.70298	5.0785
5	44.2644	29.0024	5.23259	15.5096
6	27.0183	8.93558	0.005	20.0381
7	12.285	5.61827	0.005	26.9555
8	8.64674	11.2257	3.34294	26.2784
9	13.7228	25.4001	11.6001	24.559
10	9.55584	28.9361	5.61798	12.7788
11	19.8884	36.8159	9.29705	16.406
12	16.6399	22.093	0.005	9.88615
13	20.6563	18.479	2.32184	17.6543
14	15.1923	13.2165	0.85012	19.6905
15	12.7385	14.601	2.86941	21.5728
16	16.8878	23.2986	9.58707	25.1263
17	12.3928	22.4243	5.19791	17.4712
18	15.4333	26.0105	6.14431	17.0815
19	18.0468	25.6143	5.92554	17.5945
20	15.8709	19.7188	2.23545	15.7919
21	12.0743	14.3274	0.005	14.3969
22	14.4706	17.472	3.17787	19.18
23	14.9269	19.8403	5.17384	20.4297
24	17.9169	24.6901	8.73157	22.6243
25	13.8412	21.4276	4.15137	17.0612
26	14.6419	21.6048	3.8493	16.6365
27	22.3732	27.9111	10.6664	24.0496
28	13.6319	18.06	1.74954	15.9514
29	17.5022	21.7432	6.12531	20.8432
30	13.1904	18.0207	2.52572	17.1996
31	16.5325	22.2082	6.32346	20.5347

1118

1119

1120

1121 **Table S5.3 – Synthetic MAR data with added noise replicates (replicate MAR)**

Replicate MAR data

t	X_1	X_2	X_3	X_4
1	4.49781	10.2631	14.7632	23.5471
1	5.11697	10.6373	13.2546	22.8581
1	4.17474	9.28028	15.2697	20.3851
1	4.79837	11.4797	15.3701	19.8827
1	4.61115	11.0217	12.2586	29.2412
3	11.9312	79.1524	17.0672	8.07106
3	10.947	62.6395	13.8815	7.31271
3	10.049	72.0532	15.922	9.44662
3	12.7162	67.1334	13.9795	6.6798
3	13.5553	74.391	19.6732	8.34617
5	41.8787	35.199	2.3953	15.7358
5	38.2823	34.7539	3.4456	17.6842
5	43.1218	17.1583	3.38346	9.12718
5	44.2767	37.916	2.02733	16.7366
5	30.7883	37.517	4.74218	11.6452
7	15.7223	9.78831	2.79243	47.4058
7	17.0225	8.50962	3.40691	32.5233
7	16.3596	9.74137	3.29265	24.6074
7	12.7344	13.1973	1.77531	31.342
7	12.9829	10.462	3.63077	18.2921
9	10.287	20.1958	4.824	12.0043
9	9.96034	19.7555	6.8854	22.756
9	11.7355	19.9764	5.86442	25.1459
9	7.16472	15.6769	5.89983	29.1256
9	7.27019	22.0869	5.04399	14.9662
11	15.6781	42.9179	5.62292	15.5585
11	11.8986	31.024	5.24159	12.3307
11	20.9623	32.9993	6.41934	13.992
11	14.7321	24.0842	5.79439	13.6559
11	20.2166	32.0002	5.82968	17.615
13	18.7011	21.639	3.34761	20.71
13	15.2871	17.8779	3.57818	22.9758
13	20.047	21.1359	3.47356	15.8704
13	25.1738	18.1115	4.81227	20.8173
13	21.4569	21.7648	3.75984	20.8672
15	12.942	9.13784	4.35658	25.9541
15	11.3257	12.4019	3.85461	20.5608
15	15.2742	16.3838	5.09822	30.5096
15	14.0426	15.8784	4.43551	21.3816
15	13.7737	15.6401	4.4472	23.5648
18	14.0831	25.677	5.31579	19.0626
18	14.8202	22.1809	5.23288	14.1771
18	15.3997	29.6731	6.20788	17.111
18	14.5306	26.8466	5.45091	9.73135
18	15.861	23.4862	6.90916	23.5824
22	11.7957	20.8907	4.50092	18.258
22	17.6303	21.3329	4.31011	19.19
22	12.2785	16.3511	4.18731	19.371
22	18.677	17.6922	6.05678	18.6402
22	20.6271	19.2914	3.43852	22.7473
26	15.8405	21.0941	5.80476	19.5684
26	18.9594	18.0741	4.40203	14.1525
26	6.23512	24.2245	6.18947	20.1291
26	11.5617	18.8505	3.18491	22.1173
26	14.6982	18.3201	5.83128	16.8463
30	10.6779	20.0665	4.69269	19.9748
30	16.5898	17.6329	5.2624	14.6922
30	19.4617	17.7483	5.08063	21.715
30	14.4153	13.7083	4.35089	17.093
30	14.1355	23.5724	5.30704	20.2612

1122

1123 **Table S5.4 – Initial conditions and ALVI-MI parameters estimates for the synthetic MAR data**

	Noise MAR	Replicate MAR
Spline	15DF-spline	15DF-spline
Point subsample	$t = 2, 6, 15, 18, 26$	$t = 1, 3, 11, 13, 15$
	Initial condition	Initial condition
X_1	0.814535	4.639806
X_2	7.333485	10.536398
X_3	12.89041	14.183246
X_4	14.301336	23.182858
Parameter	Estimate	Estimate
a_1	51.4374203	1.986927
b_{11}	-0.7981952	-0.029770
b_{12}	-0.1668495	-0.008283
b_{13}	-1.2191651	0.012518
b_{14}	-1.5223832	-0.068446
a_2	34.0728264	2.071364
b_{21}	-0.5711864	-0.054005
b_{22}	-0.1142499	-0.025029
b_{23}	-0.8100883	0.065384
b_{24}	-0.9677829	-0.051774
a_3	103.292766	0.327628
b_{31}	-1.7809655	-0.022297
b_{32}	-0.3029228	-0.008557
b_{33}	-2.6887155	-0.002260
b_{34}	-2.8840097	0.010410
a_4	30.8332588	-1.318063
b_{41}	-0.4422159	0.036978
b_{42}	-0.1141606	0.014019
b_{43}	-0.770332	-0.029955
b_{44}	-0.9159577	0.029689

1124

1125

1126

1127 **Table S6.1 – Parameter values and initial conditions estimated for the ‘grey whales’ dataset (Gerber,**
 1128 **Demaster, & Kareiva, 1999)**

Results generated with ALVI-MI with 3DF-spline and a data sample composed of spline values in 1959 and 1966.

Parameter	True value	Estimate
ALVI-MI		
a_1		0.0948
b_{11}		-3.79E-06
$X_1(0)$	2894	3663.9550
MAR		
α_1		1260
β_{11}		0.943
δ_1		7240000
$X_1(0)$	2894	
MAR with log transformation		
α_1		1.0368
β_{11}		0.9512
δ_1		0.0327
$X_1(0)$	2894	
MAR with smoothing		
α_1		597
β_{11}		0.9930
δ_1		199000
$X_1(0)$	2894	
MAR with log transformation and smoothing		
α_1		0.4902
β_{11}		0.9535
δ_1		0.0014
$X_1(0)$	2894	

1129

1130

1131 **Table S6.2 – Parameters and initial conditions estimated for the ‘Wolves and Moose’ dataset**
 1132 **(Vucetich, 2021)**

Results of ALVI-MI with 15DF-splines and a data sample composed of spline values in 1991, 1994 and 1997.

ALVI-MI

	Initial condition	Estimate
X_1	20	21.6545
X_2	538	560.5340
Parameter	Estimate	
a_1	-0.2732	
b_{11}	0.0073	
b_{12}	0.0001	
a_2	0.8431	
b_{21}	-0.0380	
b_{22}	-0.0001	

MAR with log transformation

	Initial condition	Estimate
X_1	20	22.0000
X_2	538	564.0000
Parameter	Estimate	
β_{11}	0.7670	
β_{21}	-0.1788	
β_{12}	0.0783	
β_{22}	0.8277	
δ_1	0.4485	
δ_2	0.1758	

1133

1134

1135

1136 5. References

- 1137 Batista Júnior, A. B., & Pires, P. S. M. (2014). An Approach to Outlier Detection and Smoothing Applied to
1138 a Trajectory Radar Data. *Journal of Aerospace Technology and Management*, 6(3), 237–248.
1139 doi:10.5028/jatm.v6i3.325
- 1140 Burden, R. L., Faires, J. D., & Burden, A. M. (1993). *Numerical Analysis* (Fifth edit). Boston, MA: PWS
1141 Publishing Co.
- 1142 Certain, G., Barraquand, F., & Gårdmark, A. (2018). How do MAR(1) models cope with hidden
1143 nonlinearities in ecological dynamics? *Methods in Ecology and Evolution*, 9(9), 1975–1995.
1144 doi:10.1111/2041-210X.13021
- 1145 Chiang, S.-Y. (2012). An application of Lotka–Volterra model to Taiwan’s transition from 200mm to
1146 300mm silicon wafers. *Technological Forecasting and Social Change*, 79(2), 383–392.
1147 doi:10.1016/j.techfore.2011.05.007
- 1148 Cleveland, W. S. (1979). Robust Locally Weighted Regression and Smoothing Scatterplots. *Journal of the*
1149 *American Statistical Association*, 74(368), 829–836. doi:10.1080/01621459.1979.10481038
- 1150 Cleveland, W. S. (1981). LOWESS: A Program for Smoothing Scatterplots by Robust Locally Weighted
1151 Regression. *The American Statistician*, 35(1), 54. doi:10.2307/2683591
- 1152 Cleveland, W. S., & Devlin, S. J. (1988). Locally Weighted Regression: An Approach to Regression Analysis
1153 by Local Fitting. *Journal of the American Statistical Association*, 83(403), 596–610.
1154 doi:10.1080/01621459.1988.10478639
- 1155 Cleveland, W. S., & Grosse, E. (1991). Computational methods for local regression. *Statistics and*
1156 *Computing*, 1(1), 47–62. doi:10.1007/BF01890836
- 1157 Dam, P., Fonseca, L. L., Konstantinidis, K. T., & Voit, E. O. (2016). Dynamic models of the complex
1158 microbial metapopulation of lake mendota. *Npj Systems Biology and Applications*, 2(1), 16007.
1159 doi:10.1038/npjbsa.2016.7
- 1160 Dam, P., Rodriguez-R, L. M., Luo, C., Hatt, J., Tsementzi, D., Konstantinidis, K. T., & Voit, E. O. (2020).
1161 Model-based comparisons of the abundance dynamics of dacterial communities in two lakes.
1162 *Scientific Reports*, 10(1), 1–12. doi:10.1038/s41598-020-58769-y
- 1163 Eilers, P. H. C. (2003). A Perfect Smoother. *Analytical Chemistry*, 75(14), 3631–3636.
1164 doi:10.1021/ac034173t
- 1165 Eilers, P. H. C., & Marx, B. D. (1996). Flexible smoothing with B -splines and penalties. *Statistical Science*,
1166 11(2), 89–121. doi:10.1214/ss/1038425655
- 1167 Gandolfo, G. (2008). Giuseppe Palomba and the Lotka-Volterra equations. *RENDICONTI LINCEI*, 19(4),
1168 347–357. doi:10.1007/s12210-008-0023-7
- 1169 Garcia, D. (2010). Robust smoothing of gridded data in one and higher dimensions with missing values.
1170 *Computational Statistics & Data Analysis*, 54(4), 1167–1178. doi:10.1016/j.csda.2009.09.020
- 1171 Gause, G. F. (1934). Experimental Analysis of Vito Volterra’s mathematical theory of the struggle for

- 1172 existence. *Science*, 79(2036), 16–17. doi:10.1126/science.79.2036.16-a
- 1173 Gerber, L. R., Demaster, D. P., & Kareiva, P. M. (1999). Gray whales and the value of monitoring data in
1174 implementing the U.S. endangered species act. *Conservation Biology*, 13(5), 1215–1219.
1175 doi:10.1046/j.1523-1739.1999.98466.x
- 1176 Haas, C. N. (1981). Application of predator-prey models to disinfection. *Journal of the Water Pollution*
1177 *Control Federation*, 53(3 I), 378–386. doi:10.2307/25041087
- 1178 Hacinliyan, A. S., Kusbeyzi, I., & Aybar, O. O. (2010). Approximate solutions of Maxwell Bloch equations
1179 and possible Lotka Volterra type behavior. *Nonlinear Dynamics*, 62(1–2), 17–26.
1180 doi:10.1007/s11071-010-9695-5
- 1181 Holmes, E. E., Ward, E. J., & Scheuerell, M. D. (2020). Analysis of multivariate timeseries using the
1182 MARSS package, version 3.11.3.
- 1183 Holmes, E. E., Ward, E. J., & Wills, K. (2012). MARSS: multivariate autoregressive state-space models for
1184 analyzing time-series data. *The R Journal*, 4(1), 11. doi:10.32614/RJ-2012-002
- 1185 Huffaker, C. B., Shea, K. B., & Herman, S. G. (1963). Experimental studies on predation: dispersion
1186 factors and predator-prey oscillations. *Hilgardia*, 34, 305–330. Retrieved from
1187 <http://hilgardia.ucanr.edu/fileaccess.cfm?article=152594&p=ZPTIMD>
- 1188 Hung, H.-C., Chiu, Y.-C., Huang, H.-C., & Wu, M.-C. (2017). An enhanced application of Lotka–Volterra
1189 model to forecast the sales of two competing retail formats. *Computers & Industrial Engineering*,
1190 109, 325–334. doi:10.1016/j.cie.2017.05.022
- 1191 Hytti, H., Takalo, R., & Ihalainen, H. (2006). Tutorial on Multivariate Autoregressive Modelling. *Journal of*
1192 *Clinical Monitoring and Computing*, 20(2), 101–108. doi:10.1007/s10877-006-9013-4
- 1193 Ives, A. R. (1995). Predicting the response of populations to environmental change. *Ecology*, 76(3), 926–
1194 941. doi:10.2307/1939357
- 1195 Knowles, I., & Renka, R. J. (2014). Methods for Numerical Differentiation of Noisy Data. *Electronic*
1196 *Journal of Differential Equations*, 21, 235–246. Retrieved from [https://ejde.math.txstate.edu/conf-](https://ejde.math.txstate.edu/conf-proc/21/k3/knowles.pdf)
1197 [proc/21/k3/knowles.pdf](https://ejde.math.txstate.edu/conf-proc/21/k3/knowles.pdf)
- 1198 Loader, C. (2012). Smoothing: Local Regression Techniques. In *Handbook of Computational Statistics* (pp.
1199 571–596). Berlin, Heidelberg: Springer Berlin Heidelberg. doi:10.1007/978-3-642-21551-3_20
- 1200 McLaren, B. E., & Peterson, R. O. (1994). Wolves, moose, and tree Rings on Isle Royale. *Science*,
1201 266(5190), 1555–1558. doi:10.1126/science.266.5190.1555
- 1202 Mühlbauer, L. K., Schulze, M., Harpole, W. S., & Clark, A. T. (2020). gauseR: Simple methods for fitting
1203 Lotka-Volterra models describing Gause’s “Struggle for Existence”. *Ecology and Evolution*, 10(23),
1204 13275–13283. doi:10.1002/ece3.6926
- 1205 Nambu, M. (1986). Plasma-maser effects in plasma astrophysics. *Space Science Reviews*, 44(3–4), 357–
1206 391. doi:10.1007/BF00200820
- 1207 Peschel, M., & Mende, W. (1986). *The Predator-Prey Model: Do we Live in a Volterra World?* Berlin:
1208 Akademie-Verlag.

- 1209 Ramsay, J. O., Hooker, G., Campbell, D., & Cao, J. (2007). Parameter estimation for differential
1210 equations: a generalized smoothing approach. *Journal of the Royal Statistical Society: Series B*
1211 *(Statistical Methodology)*, 69(5), 741–796. doi:10.1111/j.1467-9868.2007.00610.x
- 1212 Savageau, M. A., & Voit, E. O. (1987). Recasting nonlinear differential equations as S-systems: a
1213 canonical nonlinear form. *Mathematical Biosciences*, 87(1), 83–115. doi:10.1016/0025-
1214 5564(87)90035-6
- 1215 Smyth, G. (2020). Difference between LOESS and LOWESS. Retrieved 3 September 2021, from
1216 <https://stats.stackexchange.com/questions/161069/difference-between-loess-and-lowess>
- 1217 Stein, R. R., Bucci, V., Toussaint, N. C., Buffie, C. G., Räscht, G., Pamer, E. G., ... Xavier, J. B. (2013).
1218 Ecological modeling from time-series inference: insight into dynamics and stability of intestinal
1219 microbiota. *PLoS Computational Biology*, 9(12), e1003388. doi:10.1371/journal.pcbi.1003388
- 1220 Torres, N. V., & Voit, E. O. (2002). *Pathway Analysis and Optimization in Metabolic Engineering*.
1221 Cambridge, U.K.: Cambridge University Press. doi:10.1017/CBO9780511546334
- 1222 Vano, J. A., Wildenberg, J. C., Anderson, M. B., Noel, J. K., & Sprott, J. C. (2006). Chaos in low-
1223 dimensional Lotka–Volterra models of competition. *Nonlinearity*, 19(10), 2391–2404.
1224 doi:10.1088/0951-7715/19/10/006
- 1225 Varah, J. M. (1982). A spline least squares method for numerical parameter estimation in differential
1226 equations. *SIAM Journal on Scientific and Statistical Computing*, 3(1), 28–46. doi:10.1137/0903003
- 1227 Vilela, M., Borges, C. C. H., Vinga, S., Vasconcelos, A. T. R., Santos, H., Voit, E. O., & Almeida, J. S. (2007).
1228 Automated smoother for the numerical decoupling of dynamics models. *BMC Bioinformatics*, 8(1),
1229 305. doi:10.1186/1471-2105-8-305
- 1230 Voit, E. O. (2000). Canonical Modeling: Review of Concepts with Emphasis on Environmental Health.
1231 *Environmental Health Perspectives*, 108(s5), 895–909. doi:10.1289/ehp.00108s5895
- 1232 Voit, E. O. (2013). Biochemical systems theory: a review. *ISRN Biomathematics*, 2013, 1–53.
1233 doi:10.1155/2013/897658
- 1234 Voit, E. O. (2017). *A First Course in Systems Biology* (Second edi). Garland Science.
1235 doi:10.1201/9780203702260
- 1236 Voit, E. O., & Almeida, J. (2003). Dynamic Profiling and Canonical Modeling. In *Metabolic Profiling: Its*
1237 *Role in Biomarker Discovery and Gene Function Analysis* (pp. 257–276). Boston, MA: Springer US.
1238 doi:10.1007/978-1-4615-0333-0_14
- 1239 Voit, E. O., & Almeida, J. (2004). Decoupling dynamical systems for pathway identification from
1240 metabolic profiles. *Bioinformatics*, 20(11), 1670–1681. doi:10.1093/bioinformatics/bth140
- 1241 Voit, E. O., & Chou, I.-C. (2010). Parameter estimation in canonical biological systems models.
1242 *International Journal of Systems and Synthetic Biology*, 1(June), 1–19.
- 1243 Voit, E. O., Davis, J. D., & Olivença, D. V. (2021). Inference and validation of the structure of Lotka-
1244 Volterra models. *BioRxiv*. doi:10.1101/2021.08.14.456346
- 1245 Voit, E. O., Marino, S., & Lall, R. (2005). Challenges for the Identification of Biological Systems from in
1246 vivo Time Series Data, 5(December 2004), 83–92.

- 1247 Voit, E. O., & Savageau, M. A. (1982a). Power-law approach to modeling biological systems; II.
1248 Application to ethanol production. *J. Ferment. Technol.*, 60(3), 229–232.
- 1249 Voit, E. O., & Savageau, M. A. (1982b). Power-law approach to modeling biological systems; III. Methods
1250 of analysis. *J. Ferment. Technol.*, 60(3), 233–241.
- 1251 Voit, E. O., & Savageau, M. A. (1986). Equivalence between S-systems and Volterra systems.
1252 *Mathematical Biosciences*, 78(1), 47–55. doi:10.1016/0025-5564(86)90030-1
- 1253 Vucetich, J. A. (2021). Wolves and moose of Isle Royale. Retrieved from <https://isleroyalewolf.org/>
- 1254 Zhou, Y., & Chen, B. (2006). Analysis of multi-ISPs game based on lotka-volterra model. In *CIMCA 2006:
1255 International Conference on Computational Intelligence for Modelling, Control and Automation,
1256 Jointly with IAWTIC 2006: International Conference on Intelligent Agents Web Technologies ... IEEE
1257 Computer Society*. doi:10.1109/CIMCA.2006.44
- 1258
- 1259