

Inferring ecological interactions from time series data using neural ordinary differential equations fitted by gradient matching

Willem Bonnaffé^{1,2}, Ben Sheldon¹, & Tim Coulson²

1. Edward Grey Institute of Field Ornithology, Department of Zoology, Oxford University, Zoology Research and Administration Building, 11a Mansfield Road, Oxford OX1 3SZ

2. Ecological and Evolutionary Dynamics Lab, Department of Zoology, Oxford University, Zoology Research and Administration Building, 11a Mansfield Road, Oxford OX1 3SZ

Emails: willem.bonnafe@stx.ox.ac.uk; tim.coulson@zoo.ox.ac.uk

Running title: Repeatable interactions and dynamics

Keywords: Artificial Neural Networks; Ecological Dynamics; Ecological interactions; Geber Method; Neural Ordinary Differential Equations; Ordinary Differential Equations; Prey-predator dynamics; Time series analysis; Rotifers; Microcosm;

Specifications: 140 words in abstract; 7071 words in text; 40 references; 5 figures; 1 table

Contact: Willem Bonnaffé, 61 St Giles, Pusey House, St Cross College, Oxford, OX1 3LZ, UK (w.bonnafe@gmail.com)

Statement of authorship: Willem Bonnaffé designed the method, performed the analysis, wrote the manuscript; Ben Sheldon provided input for the manuscript, commented on the manuscript. Tim Coulson led investigations, provided input for the manuscript, commented on the manuscript.

Abstract

Generalisation of dynamical processes across natural systems is difficult because they are complex and hard to observe. The hope is that generalisation may be achieved by adequately modelling the complexity of systems, and observe them in sufficient detail. We investigate this by looking at the consistency of ecological interactions across three replicates of a three-species prey-predator system, well-observed in an artificial environment, using neural ordinary differential equations. We find that dominant interactions are consistent across the replicates, while weaker interactions are not, leading to different dynamical patterns across replicated systems. Our study hence suggests that generalisation of dynamical processes across systems may not be possible, even in simpler systems in ideal monitoring conditions. This is a problem because if we are not able to make generalisations in a simple artificial system, how can we make generalisation in the real world?

1 Introduction

The repeatability of ecological and evolutionary dynamics varies widely across systems and species. Sticklebacks from different lakes in Canada have independently evolved to a similar river morph phenotype (Thompson, Taylor, and Mcphail 1997). In guppies, four replicated populations located in different streams in Trinidad evolved the same low-predation phenotype (Reznick, Bryga, and Endler 1990). Multiple studies in experimental microcosms, particularly in rotifer populations, have shown that population dynamics were broadly repeatable (Yoshida et al. 2003; Yoshida et al. 2007; Becks et al. 2010; Becks et al. 2012; Hiltunen et al. 2013). Overall, this demonstrated that ecological and evolutionary dynamics may be repeatable across different instances of the same system, at least qualitatively. This was a fascinating finding given the complexity of the mechanisms involved and the subtle variations in environmental conditions across the different populations.

These systems hinted at the possibility for identifying global, generalisable, dynamical models. In practice, however, generalising dynamics and dynamical processes (i.e. functional representations describing which and how state variables affect each other and determine system dynamics) across natural systems has proven difficult (Lawton 1999). First, even if the dynamical patterns, and their outcomes, may appear to be conserved across similar systems, they may be underpinned by different processes. For instance, the evolution of the sticklebacks to highly similar river-adapted phenotypes has been shown to be underpinned by radically different genetic alterations (Raeymaekers et al. 2017). Second, it is often unclear whether quantitative differences across replicated systems

21 arise from pure stochasticity (Dallas et al. 2021), observation error (De Meester et al. 2019), or
22 deterministic changes in the dynamical processes. Finally, the complexity of biological processes
23 themselves (Adamson and Morozov 2013), differences in genetic and environmental contexts, may
24 prevent the identification of a suitable dynamical model. For example, Becks and colleagues found
25 that differences in the initial amount of genetic variation in otherwise identical rotifer populations
26 led to subtle changes to the dynamics (Becks et al. 2010). Different access to seed supplies can
27 modify the strength of the interaction between a plant and its herbivore, leading to either stable or
28 oscillatory dynamics (Bonsall, Van Der Meijden, and Crawley 2003). Differences in temperature
29 can alter the ecological interaction structure of entire ecosystems (Shurin et al. 2012; Bonnaffé
30 et al. 2021). Because of this, vital rates are often found to be inconsistent in time (Gross, Ives, and
31 Nordheim 2005; Adamson and Morozov 2013), and space (e.g. Gamelon et al. 2019). Overall, a
32 growing body of evidence shows that generalisation of dynamical processes across similar natural
33 systems often fails (Lawton 1999, e.g. Kendall et al. 2005; Demyanov, Wood, and Kedwards 2006;
34 Ezard, Côté, and Pelletier 2009).

35 So how could repeatable dynamics arise across multiple instances of the same system? We would
36 expect dynamics to be repeatable if the components of the system (e.g. species), as well as interac-
37 tions between components, are conserved. For this, populations should have similar distributions
38 for the traits that underpin these interactions, and should further share the same environmental con-
39 ditions, across instances. While this is unreasonable to expect from a natural system, it may be
40 achievable in an artificial setting. In such a setting, it is possible to understand the structure of

41 the system, to control the environment, and to reduce observation error. So if we fail to identify
42 and generalise dynamical models in natural systems, perhaps we may be able to do so in artificial
43 systems.

44 In spite of this there are few studies that have attempted to characterise the generalisability of
45 dynamics across replicated systems in a laboratory setting. In such a setting, idiosyncrasies in pop-
46 ulation dynamics can arise from (1) variations in ecological interactions and individual processes,
47 as a result of evolution (e.g. Yoshida et al. 2003), or stochasticity (Dallas et al. 2021), (2) variations
48 in initial conditions due to the experimental setting (Yoshida et al. 2003; Becks et al. 2010; De
49 Meester et al. 2019), and (3) the complexity of the system which can lead to large changes in sys-
50 tem dynamics with small changes in the system state and structure (Adamson and Morozov 2013).

51 Two studies, one in aphids and the other in rotifers, found substantial variation in vital rates across
52 replicated populations, by fitting a stage-structured population ODE model to population dynamics
53 time series data (Bruijning, Jongejans, and Turcotte 2019; Rosenbaum et al. 2019). These studies
54 hint that generalisability of population dynamical processes may not be possible because of in-
55 trinsic population structure and evolution, even in virtually identical populations hosted in artificial
56 environments.

57 We identified three gaps in the literature. First, this kind of evidence remains scarce, due in part
58 to the fact that dynamical modelling approaches guided by empirical data are still not widespread
59 (Pontarp, Brännström, and Petchey 2019). Second, most of these studies relied on parametric
60 frameworks, which impose arbitrary pre-determined forms for the dynamical processes at play, so

61 that their model may not capture properly the complexity of the dynamics of these populations (Jost
62 and Ellner 2000; Adamson and Morozov 2013; Bonnaffé, Sheldon, and Coulson 2021). Finally,
63 most studies usually analyse dynamics in single-species systems, but not multi-species systems,
64 such as those with intraguild predation, which are more biologically realistic scenarios (Hiltunen
65 et al. 2013). Further studies are consequently required to investigate the consistency of dynamical
66 processes in simple multi-species and well-observed systems, to conclude about the generalisability
67 of population dynamics across systems.

68 Our aim in this study is to provide an assessment of the repeatability of dynamical processes across
69 different instances of a realistic multi-species system hosted in a well-observed environment. We
70 do this by quantifying the direction, strength, and consistency of interactions in time and across
71 replicates of a three-species microcosm in an experimental setting. We hypothesise that if the
72 system is (1) simple enough, (2) well-observed, (3) in a controlled environment, then dynamical
73 effects/interactions should be broadly consistent in time and across replicates, hence allowing for
74 generalisation of dynamics across systems. We consider three replicates of a three-species system,
75 consisting in a prey (algae), intermediate-predator (flagellate), and top-predator (rotifer). The algae
76 is consumed by the flagellate and rotifer, and the flagellate is consumed by the rotifer. We use
77 three replicated system runs from a study by Hiltunen and colleagues which feature sequential
78 oscillations of the density of the three species (Hiltunen et al. 2013). We analyse the time series
79 with neural ordinary differential equations (Bonnaffé, Sheldon, and Coulson 2021), which allows
80 us to approximate non-parametrically population growth rates, and quantify the direction, strength,

81 and consistency of inter- and intra-specific effects on the growth of each population. We find that
82 the interaction between the rotifer and algae is consistent throughout time and across replicates,
83 while the interaction between the flagellate and the two other species is not. Our study suggests
84 that dynamical processes may sometimes not be consistent and generalisable across systems, even
85 when they are as close to identical as experimentation permits. We discuss these results and hint at
86 the underlying impact of evolution driving differences in these systems.

87 **2 Material and Methods**

88 **2.1 Method overview**

89 We aim to provide a nonparametric method for estimating ecological interactions from time series
90 data of species density. We do this by approximating the dynamics of each species with neural
91 ordinary differential equations (NODEs, Bonnaffé, Sheldon, and Coulson 2021). We then compute
92 ecological interactions as the sensitivity of these dynamics to a change in the respective species
93 densities.

94 **2.2 Neural ordinary differential equation**

95 A NODE is a class of ordinary differential equation (ODE) that is partly or entirely defined as an
96 artificial neural network (ANN). They are useful to infer dynamical processes non-parametrically
97 from time series data (Bonnaffé, Sheldon, and Coulson 2021). We choose NODEs over standard
98 statistical approaches because they offer two advantages. The first is that NODEs approximate

the dynamics of populations non-parametrically. NODEs are therefore not subjected to incorrect model specifications (Jost and Ellner 2000; Adamson and Morozov 2013). This provides a more objective estimation of the inter-dependences between state variables. The second advantage is that it is a dynamical systems approach. So that the approach includes lag effects through interacting state variables, not only direct effects between them.

For the sake of simplicity, we first consider a two-species NODE system:

$$\begin{cases} \frac{dR}{dt} = r_R(R, N, \beta_R)R \\ \frac{dN}{dt} = r_N(R, N, \beta_N)N \end{cases} \quad (1)$$

where dR/dt and dN/dt denote the change in the density of prey ($R \in \mathcal{R}^+$), and predator ($N \in \mathcal{R}^+$), in continuous time. The per-capita growth rates r_R and r_N are non-parametric functions of the density of each species. The shape of the non-parametric functions is controlled by the parameter vectors β_R and β_N .

Traditionally, each non-parametric function is defined as an ANN function of the state variables. The most common class of ANN used for NODEs are single layer perceptrons (SLPs):

$$r_N(R, N, \beta_N) := \beta_0 + \sum_{i=1}^I \beta_i f_{\sigma}(\beta_{i0} + \beta_{i1}R + \beta_{i2}N) \quad (2)$$

which feature a single layer that maps the inputs, here the species densities R and N , to a single output, the per-capita growth rate. The parameter vector β_R and β_N are the weights of the connections in the SLPs. SLPs can be viewed as weighted sums of basis functions f_{σ} . More details regarding

114 these models can be found in our previous work (Bonnaffé, Sheldon, and Coulson 2021).

115 **2.3 Fitting NODEs by gradient matching**

116 This section describes how to estimate the parameters β of the NODE system given a set of time
117 series. In previous work, we developed a simulation-based approach to fit NODE systems to time
118 series data (Bonnaffé, Sheldon, and Coulson 2021). We would first simulate the NODE system over
119 the entire time series. Then we would compute the error between the predictions of the NODE
120 model and the observations. Finally, we would change the weights of the NODEs to minimise
121 this error. There are two caveats with this approach. The first caveat is that the NODE system
122 has to be simulated over the entire range of the data at every step of the optimisation. This step is
123 computationally expensive to perform. Second, the numerical integration prevents the computation
124 of gradients of the posterior distribution of the model. This prevents the use of efficient gradient
125 descent approaches.

126 Instead, we propose to use a *gradient matching* approach to fit NODEs, which relies on data in-
127 terpolation to approximate states and dynamics. The method we propose here is derived from the
128 *gradient matching* approach that Ellner and colleagues developed to fit ODEs (Ellner, Seifu, and
129 Smith 2002; Wu, Fukuhara, and Takeda 2005). We proceed in three steps, presented graphically in
130 Fig. 1 and detailed in the following sections. First, we interpolate the time series data and dynamics
131 of each species in the system. Second, we train each NODE to satisfy the interpolated dynamics,
132 thereby avoiding the simulation step of the previous method. This amounts to fitting ANNs that

133 take the density of each population as input to the interpolated per-capita growth rates.

134 **Data interpolation**

135 We interpolate the time series and differentiate it with respect to time in order to approximate
136 the dynamics of the system. We perform the interpolation via non-parametric regression of the
137 interpolating function on the time series data

$$\varepsilon_{N,t}^{(o)} = N_t - \tilde{N}(t, \Omega_N) \quad (3)$$

138 where $\varepsilon_{N,t}^{(o)}$, the observation error, is defined as the difference between N_t , the observed value of the
139 variable at time t , and $\tilde{N}(t, \Omega_N)$, the value predicted by the interpolating function. The interpolating
140 function is chosen to be a SLP with sinusoid activation functions

$$\tilde{N}(t, \Omega_N) = \exp \left\{ \sum_{i=1}^I \omega_{i0} \sin(\pi(\omega_{i1} + \omega_{i2}t)) \right\} \quad (4)$$

141 where $\tilde{N}(t, \Omega_N)$ is the interpolated state variable, defined as a weighted sum of sinusoid functions of
142 time. The interpolation parameter vector Ω_N contains the weights ω_{i0} , ω_{i1} , and ω_{i2} which control
143 the amplitude, shift, and frequency of the oscillations in the time series, respectively. We found
144 sinusoid activation functions to be most efficient for interpolating population dynamics compared
145 to other functions (such as sigmoid, hyperbolic). Following this approach we obtain directly an
146 approximation of the dynamics of the state variable by differentiating the SLP with respect to
147 time

$$\frac{\partial}{\partial t} \tilde{N}(t, \Omega_N) = \sum_{i=1}^I \omega_{i0} \pi \omega_{i2} \cos(\pi(\omega_{i1} + \omega_{i2}t)) \exp \left\{ \sum_{i=1}^I \omega_{i0} \sin(\pi(\omega_{i1} + \omega_{i2}t)) \right\} \quad (5)$$

148 as well as an analytical expression of the interpolated per-capita growth rate of the populations, by
 149 combining equation (4) and (5)

$$\tilde{r}_N(t, \Omega_N) := \frac{1}{\tilde{N}} \frac{\partial \tilde{N}}{\partial t} = \sum_{i=1}^I \omega_{i0} \pi \omega_{i2} \cos(\pi(\omega_{i1} + \omega_{i2}t)) \quad (6)$$

150 **Fitting NODEs to interpolated data**

151 In a second step, we train the NODE system (1) to satisfy the interpolated dynamics, given the
 152 interpolated state variables. Thanks to the interpolation step, this simply amounts to performing a
 153 non-parametric regression of the NODE-approximated per-capita growth rate (equation 2) on the
 154 interpolated per-capita growth rate (equation 6)

$$\varepsilon_{N,t}^{(p)} = \tilde{r}_N(t, \Omega_N) - r_N(\tilde{R}, \tilde{N}, \beta_N) \quad (7)$$

155 where the process error, $\varepsilon_{N,t}^{(p)}$, is defined as the difference between the interpolated growth rate \tilde{r}_N
 156 and the NODE approximation r_N .

157 **Statistical modelling approach**

158 Overall, fitting the NODE system (1) comes down to two steps. First, we interpolate the data by
 159 finding the parameter vector Ω that minimises the observation error $\varepsilon^{(o)}$ in equation (3). Second, we

160 fit the NODE per-capita growth rate by finding the parameter vector β that minimises the process
 161 error $\varepsilon^{(p)}$ in equation (7).

162 The fitting of the models is performed in a Bayesian framework, considering normal error structure
 163 for the residuals, and normal prior density distributions on the parameters

$$p(\theta|\mathcal{D}) \propto p(\mathcal{D}|\theta)p(\theta) \quad (8)$$

164 where θ is the parameter vector of the model, and \mathcal{D} the evidence, namely the data that the model
 165 is fitted to. Assuming a normal likelihood for the residuals given the evidence we get

$$p(\mathcal{D}|\theta) = \prod_{i=1}^I \frac{1}{\sqrt{2\pi\sigma^2}} \exp \left\{ -\frac{e_i(\mathcal{D}, \theta)^2}{2\sigma^2} \right\} \quad (9)$$

166 where $e_i(\mathcal{D}, \theta)$ are the residuals of the model given the parameters, and the evidence. In the case of
 167 the interpolation, the residuals correspond to the observation error $\varepsilon^{(o)}$ (equation 3). In the case of
 168 the NODE approximation, they correspond to the process error $\varepsilon^{(p)}$ (equation 7). I is the number
 169 of data points, either observations in the case of the interpolation, or interpolated points in the case
 170 of the NODE fitting.

171 The prior probability density functions for the parameters are given by

$$p(\theta) = \prod_{j=1}^J \frac{1}{\sqrt{2\pi\delta_j^2}} \exp \left\{ -\frac{\theta_j^2}{2\delta_j^2} \right\} \quad (10)$$

172 where J is the number of parameters in the models. The parameter δ_j controls the dispersion of the

173 priors, and thereby the complexity/level of constraint of the model.

174 There is no standard approach for choosing δ . Low values of dispersion may increase constraint
 175 on parameters too drastically, which would lead to underfitting, and result in a reduction of the
 176 variance of parameter estimates and bias mean estimates towards 0. In contrast, too high values of
 177 dispersion may lead to overfitting, by allowing for more complex shapes. To account for this, we
 178 optimise the models on the second-level of inference. This means that we are finding the optimal
 179 value of δ , in addition to optimising the model parameters. We do this by optimising the marginal
 180 posterior density of the parameters, obtained by averaging out δ following a modification of the
 181 approach developed by Cawley and Talbot (Cawley and Talbot 2007). This yields the following
 182 expression for the marginal log posterior density of the parameters

$$\log P(\theta|\mathcal{D}) \propto -\frac{I}{2} \log \left(1 + \sum_{i=1}^I e_i(\mathcal{D}, \theta)^2 \right) - \frac{J}{2} \log \left(1 + \sum_{j=1}^J \theta_j^2 \right) \quad (11)$$

$$\log P(\Omega|\mathcal{D}) \propto -\frac{I}{2} \log \left(1 + \sum_{i=1}^I \left(\varepsilon_i^{(o)} \right)^2 \right) - \frac{J}{2} \log \left(1 + \sum_{j=1}^J \Omega_j^2 \right) \quad (12)$$

$$\log p(\beta|\Omega) \propto -\frac{1}{2} \sum_{i=1}^I \left(\frac{\varepsilon_i^{(p)}}{\sigma} \right)^2 - \frac{1}{2} \sum_{j=1}^J \left(\frac{\beta_j}{\delta_j} \right)^2 \quad (13)$$

183 which amounts to optimising the log of the sum of squared residuals rather than the sum of squared
 184 residuals. $P(\theta|\mathcal{D})$ designates the marginal posterior distribution. More details on how to derive this
 185 expression from equation (8) can be found in a supplementary file (See supplementary A).

186 Finally, we estimate uncertainty in parameter values through anchor sampling, which produces ap-
187 proximate Bayesian estimates of the posterior distribution of the parameters (Pearce et al. 2018).
188 The technique is simple in that it requires sampling a parameter vector from the prior distribu-
189 tions, and then optimising the posterior distribution from this starting point. By repeatedly taking
190 samples, the sampled distribution approaches the posterior distribution and provides estimates and
191 error around the quantities that can be derived from the models. The expectation of the quantities
192 can then be approached by computing the mean of the approximated posterior distributions. The
193 great strength of this approach is that it is unlikely to get stuck in local maxima and provides a
194 more robust optimisation of the posterior. In this study, we took 100 posterior samples for each
195 time series, namely a hundred samples for the interpolation, and another hundred for the fitting
196 of the NODE. The initial value of the parameters were picked from a random normal distribution
197 with parameters $\sigma \geq 0.4$, which prevented underfitting the time series. We insured that there was
198 moderate temporal autocorrelation and normality by visualising the residuals of the models. We
199 also insured that the results were repeatable by running the entire fitting process a second time. We
200 did not perform cross validation of results as we were only interested in estimating effects within
201 the time series considered, rather than predicting future time steps.

202 **2.4 Model analysis**

203 We analyse the shape of the per-capita growth rates to recover the interaction between the three
204 species in the system. In particular, we look at the effect and contribution of each species to the
205 dynamics of the other. The effect is computed as the sensitivity (i.e. the gradient) of the per-capita

206 growth rate of a given species with respect to the density of the other species. The contribution is
207 computed following the Geber method (Hairston et al. 2005), which comes down to multiplying
208 the dynamics of a variable by its effects on the other variables. We further compute the importance
209 of a species in driving the dynamics of another by computing its relative contribution compared to
210 other species at each time step. More details on how to recover these quantities can be found in our
211 previous study (Bonnaffé, Sheldon, and Coulson 2021).

212 **3 Case study**

213 **3.1 System**

214 We consider a three-species laboratory microcosm consisting of an algal prey (*Chlorella autroph-*
215 *ica*), a flagellate intermediate predator (*Oxyrrhis marina*), and a rotifer top predator (*Brachionus*
216 *plicatilis*). The algal prey is consumed by the intermediate and top predator, the top predator also
217 consumes the intermediate predator (Arndt 1993). The dynamics of this system, here the daily
218 change in the density of each species, were recorded in three replicated time series experiments
219 performed by Hiltunen and colleagues (Hiltunen et al. 2013, Fig. 1). The aim of their experiment
220 was to determine which type of population dynamics would arise in a system with two predators
221 competing for the same resource (the algae), where one predator (the rotifer) would also be able to
222 consume its competitor (the flagellate). According to their expectations, they found prey-predator
223 oscillations, where the lag between the density peaks of each species reflected their position in the
224 food web. Namely that the peak of algae preceded the flagellate peak, which itself preceded the

225 rotifer peak.

226 Their microcosms are close to true replicates in that environmental conditions, namely temperature,
227 salinity, and nutrient influx, were maintained constant, and initial conditions, that is the initial
228 density of each species, were shared across all replicates. In spite of that, they still found evidence
229 for algae evolution in some parts of the time series, which resulted in a shift of the dynamics from
230 fast prey-predator cycles to slower oscillations, similar to those documented in previous studies on
231 similar systems (Yoshida et al. 2003; Becks et al. 2010), even in lineages where genetic variation in
232 predator defense traits was eliminated at the start of the experiment. Consequently, the time series
233 that they reported are the ones that did not present evidence of evolution, and therefore displayed
234 purely ecological dynamics.

235 We use their time series because they describe a simple yet biologically realistic ecosystem, and
236 because the quality of the replication of their microcosm reduces as much as possible observational
237 and experimental error, and rules out environmental variation (Hiltunen et al. 2013). We digitised
238 these time series by extracting by hand the coordinates of every points in the referential of the axis
239 of the graph of the original study, and analysed them.

240 **3.2 Model specifications**

241 The aim of the modelling approach is to infer the drivers of the dynamics of each species from
242 the time series data. More specifically, we want to quantify the effect of a change in the density
243 of one species on the dynamics of the other species. In this way we can understand which, and

244 to what extent, species interactions drive population dynamics. To do this we use neural ordi-
 245 nary differential equation (NODEs), which is a novel methodology allowing us to infer dynamical
 246 processes non-parametrically from time series data (Bonnaffé, Sheldon, and Coulson 2021). We
 247 choose this methodology over traditional approaches because it offers two advantages. The first
 248 lies in the fact that NODEs approximate the dynamics of populations non-parametrically, and are
 249 therefore not subject to incorrect model specifications (Jost and Ellner 2000; Adamson and Moro-
 250 zov 2013). This is important as it offers an objective estimation of the inter-dependences between
 251 state variables, and hence a reliable assessment of whether a species is contributing to the dynamics
 252 of another. The second advantage is that it is a dynamical systems approach, which means that the
 253 effects are estimated in a dynamically consistent system of ODEs (Bonnaffé, Sheldon, and Coulson
 254 2021). This is useful because it accounts for the dynamical nature of the system, so that it includes
 255 lag effects, not just direct correlations between variables.

256 We define a simple NODE system for the three-species system described previously

$$\begin{aligned}
 \frac{dR}{dt} &= r_R(R, G, B, \beta_R)R \\
 \frac{dG}{dt} &= r_G(R, G, B, \beta_G)G \\
 \frac{dB}{dt} &= r_B(R, G, B, \beta_B)B
 \end{aligned}
 \tag{14}$$

257 where dR/dt , dG/dt , and dB/dt denote the change in rotifer (R), algae (G), and flagellate (B)
 258 density in continuous time. The per-capita growth rates r_R , r_G , and r_B are non-parametric functions
 259 of the density R , G , B of each species. The shapes of the non-parametric functions are controlled

260 by the parameter vectors β_R , β_G , and β_B . Fitting the NODE system (1) amounts to finding the
 261 parameter vectors, and thereby the per-capita growth rates, that best describe the changes in density
 262 observed in the time series data.

263 Each non-parametric functions is an artificial neural network (ANN). ANNs are powerful mathe-
 264 matical objects that can be trained to approximate the shape of dynamical processes (Funahashi and
 265 Nakamura 1993; Chen and Chen 1993). For the sake of simplicity, we consider the simplest form
 266 of an ANN which contains a single hidden layer, namely a single layer peceptron (SLP)

$$r_R = \sum_{i=1}^N \beta_i f_{\sigma}(\beta_{i0} + \beta_{i1}R + \beta_{i2}G + \beta_{i3}B) \quad (15)$$

267 which takes as input the density of each species R , G , and B , and output the corresponding per-
 268 capita growth rate. The parameter vector β_R , β_G , β_B , contain the weight of the connections in the
 269 ANNs. The SLP can be viewed as a weighted sum of basis functions f_{σ} of the state variables of
 270 the system. In this study we consider sigmoid basis functions, as they are commonly used and
 271 their capacity to approximate any continuous function is well established theoretically (Funahashi
 272 and Nakamura 1993). The number of units in the hidden layer N is chosen to be 10, as this is
 273 a commonly used number for systems of that size (e.g. Wu, Fukuhara, and Takeda 2005). More
 274 details regarding these models can be found in our previous work (Bonnaffé, Sheldon, and Coulson
 275 2021).

276 **4 Results**

277 We analyse sequentially the dynamics of each species, focussing on the amount of variation in
278 per-capita growth rates explained by the NODE model, the overall direction, consistency, and im-
279 portance of ecological interactions, and differences across replicates. Results are summarised in
280 Table 1 and described in details for each species in the following section.

281 **Drivers of top predator dynamics**

282 Figure 2 presents the drivers of the dynamics of rotifer. The NODE approximation of the per-capita
283 growth rate fits quite well the interpolated per-capita growth rate across all replicates (Fig. 2, A2
284 B2 and C2, $r^2 > 0.7$, Table 1). The analysis of effects reveals overall a positive effect of algae on
285 rotifer growth in all replicates (Fig. 2, A3, B3, C3, green line). The intermediate predator has a
286 positive effect on rotifer growth in replicates A and C only (Fig. 2, A3, B3, C3, blue line). We find
287 positive intra-specific density-dependence in the first replicate only (Fig. 2, A3, red line). Overall,
288 all effects are consistent throughout the time series. The algae is the dominant driver of rotifer
289 dynamics as it accounts for 55%, 93%, and 74% of the change in per-capita growth rates across the
290 three replicates (Table 1, Fig. 2, A5, B5, C5, green line).

291 **Drivers of the prey dynamics**

292 The per-capita growth rate of the algae is well explained by the NODE approximation (Fig. 3,
293 A2, B2, C2, $r^2 > 0.8$, Table 1). Overall, rotifers have a negative impact on the growth of algae
294 in all replicates (Fig. 3, A3, B3, C3, red line). We find evidence for negative density-dependence

295 in replicate A and positive density-dependence in replicate B, but not in replicate C (Fig. 3, A3,
296 B3, C3, green line). The intermediate predator has an overall negative effect on Algae only in
297 replicate B (Fig. 3, B3, blue line). The main driver of algae dynamics is the rotifer population,
298 which accounts for 58%, 44%, and 90% of the change in algae per-capita growth rate across the
299 three replicates. Density-dependence, however, plays a role in replicate A and B, with 40% and
300 24% of total change in growth, respectively (Table 1). The intermediate predator contributes only
301 to algae growth in replicate B, accounting for 32% change in growth (Table 1). Overall, effects are
302 found to be consistent throughout the time series except in replicate B (Fig. 3, B3), where effects
303 vary in complicated ways, leading to a period in the time series where the algae is mostly driven by
304 the intermediate predator and positive density-dependence, and less impacted by the top predator
305 (Fig. 3, B5, from time 3 to 7.5).

306 **Drivers of the intermediate predator dynamics**

307 The per-capita growth rate of the intermediate predator is quite well captured by the NODE approx-
308 imation (Fig. 4, A2, B2, C2, $r^2 > 0.7$, Table 1). The intermediate predator is mainly negatively
309 affected by the rotifer population (Fig. 4, A3, B3, C3, red line). The algae has a negative effect
310 on flagellate growth in replicate A, and a positive one in replicate B (Fig. 4, A3, B3, green line).
311 The rotifer predator dynamics accounts for 78%, 62%, 91% of the change in the flagellate growth
312 rate, and the algae 20% and 37% in replicate A and B, respectively (Table 1, Fig. 4, A5, B5, C5).
313 Overall, effects are consistent throughout the time series.

314 5 Discussion

315 Our ability to generalise dynamical processes and patterns across populations and communities is
316 limited by the complexity of the processes, differences in environments, and incomplete and/or
317 erroneous observations. It remains unclear to what extent generalisation would be possible if we
318 overcame these limitations. We tackle this question by looking at the consistency of dynamical
319 patterns across three replicated runs of a simple three-species community, hosted in identical en-
320 vironmental conditions in the lab. We expected to find consistency in the drivers of population
321 dynamics, both in time and across replicates, and thereby demonstrate that generalisation of dy-
322 namical processes may be possible if the system states were well-observed and environmental
323 conditions were known. To verify this expectation we (1) characterised the amount of variation in
324 per-capita growth rates that is explainable deterministically, (2) quantified the direction, strength,
325 and importance of ecological interactions for the growth of each population, and (3) described how
326 these varied in time and across replicates. Our results are summarised in Figure 5. We find that
327 only the effect of algae on rotifer ($G \rightarrow R$), and that of rotifer on algae ($R \rightarrow G$) and flagellate
328 ($R \rightarrow B$) are conserved across the replicates. We find strong variation in the direction and impor-
329 tance of intra-specific density-dependence in rotifer ($R \rightarrow R$) and algae ($G \rightarrow G$) growth across the
330 three replicates. The role played by the intermediate predator in the system was also different in
331 all replicates, in that it only contributed substantially to the dynamics of the algae in replicate B
332 ($B \rightarrow G$), and was either negatively, positively, or not affected by the algae ($G \rightarrow B$). Overall, this
333 shows that the dominant interactions are conserved across replicates, but that minor interactions

334 vary substantially in importance and effect. Furthermore, we find that these dynamical processes
335 are more consistent in time within a system, than across replicates. Our results demonstrate that
336 because of partially generalisable dynamical processes, dynamical patterns may not be generalis-
337 able across systems, even with limited observation error and when environmental conditions and
338 community structure are conserved.

339 Overall, our results are consistent with the biology of the system. The rotifer top-predator is found
340 to have a strong negative impact on the two other species, in spite of variation in prey preference
341 across replicates. This is consistent with previous study which have established the importance
342 of rotifers for top-down control of flagellate and algal populations (Arndt 1993; Hiltunen et al.
343 2013). What is more suprising is the positive intra-specific density-dependence in the growth rate
344 of the rotifer population in replicate A. This implies that the population of rotifer grows more at
345 high density. This might be explained by various biological mechanisms, such as cannibalism
346 (Gilbert 1976), though evidence remains limited in the *Brachionus* genus, or higher mating success
347 at high density (Snell and Garman 1986). Similarly, the algae shows signs of positive intra-specific
348 density-dependence in replicate B, though this effect remains confined to a brief period in the time
349 series. This may be due to a higher chance of evading predators at high-density. This shows that the
350 NODEs approach used here recovers results consistent with existing knowledge, but also identify
351 subtle, more intriguing dynamical processes.

352 What might be the drivers of differences in the dynamical processes across these three replicates?

353 One of the main source of variation in dynamics may be differences in the intrinsic structure of

354 populations, such as variation in traits influencing intra- and inter-specific interactions which may
355 lead to different dynamics (Yoshida et al. 2003; Yoshida et al. 2007; De Meester et al. 2019;
356 Bruijning, Jongejans, and Turcotte 2019). Differences in the phenotypic structure may be due to
357 unaccounted variation in initial conditions (Becks et al. 2010), or variation that developed through-
358 out time as a result of evolution (e.g. Yoshida et al. 2003; Yoshida et al. 2007). In particular, the
359 algae in this system is prone to evolve a predator defence behaviour, by forming clumps, which
360 reduce predation risk (Yoshida et al. 2003; Hiltunen et al. 2013). In their original paper, the authors
361 limited the initial genetic diversity in the algae and focussed on replicates which did not display
362 evidence of evolution, in an attempt to limit the impact of initial variation in phenotypic structure,
363 and of evolution, on the dynamics (Hiltunen et al. 2013). In spite of that, evolution may not be
364 eliminated completely, thus variation in traits governing the interactions between the species in
365 the system may still have developed during the experiment, and led to changes in the dynamical
366 processes across replicates. This would further be consistent with results from Yoshida and col-
367 leagues, who showed that evolution of prey defense could lead to ecological dynamics inconsistent
368 with the known trophic interactions (Yoshida et al. 2007). Becks and colleagues also showed that
369 small changes in the initial genotypic diversity could lead to drastically different eco-evolutionary
370 dynamics (Becks et al. 2010). Our study hence reinforces the idea that rapid evolution may prevent
371 generalisation of dynamical processes (Ezard, Côté, and Pelletier 2009; De Meester et al. 2019),
372 and further suggests that this may also be the case in simple systems with limited environmental
373 variation and opportunity for evolution.

374 Alternatively, stochasticity may be a major driver of differences across systems (Dallas et al. 2021).
375 First, stochasticity in initial conditions, arising from the sampling of the communities of each
376 replicate, could introduce differences in the interactions between the three populations. Second,
377 stochasticity in the population dynamics themselves may result in different changes in density lev-
378 els in communities that are otherwise identical. Because our modelling approach is deterministic,
379 it does not directly provide an estimate of the total variation explained by stochasticity. Our mod-
380 elling approach decomposes the variation in the data into observation and process error (Calder et
381 al. 2003). First, the interpolation step introduces residual observation error, namely variation that
382 is not captured by the interpolation. Second, the fitting of the NODE to the interpolation introduces
383 residual process error, which is variation in the observation model that is not explained by the pro-
384 cess modelled by the NODE. Stochasticity in the dynamics could explain the observation and pro-
385 cess residual error (Calder et al. 2003), while stochasticity in initial conditions can only influence
386 differences across replicates. Yet, we find relatively small process and observation error ($> 70\%$
387 of variance explained). So that, the dynamics of the three species are well explained by relatively
388 simple linear deterministic effects between the state variables, which means that though dynamical
389 processes differ across replicates they are reasonably consistent in time within each system. This
390 suggests that stochasticity in dynamics plays a minor role in driving differences in dynamics across
391 replicates, compared to stochasticity in initial conditions. In order to quantify this, we would need
392 to estimate the influence of stochasticity directly. This can be done by modelling explicitly the
393 random distribution of model parameters that underpin the dynamics of populations, which would
394 then inform us about the importance of stochasticity driven by variation at the individual-level (Fox

395 and Kendall 2002). Additionally, we could model stochasticity explicitly in the model with neural
396 stochastic differential equations, which would allow us to separate the amount of change explain-
397 able by the deterministic part of the model, from demographic stochasticity, at each time step (Jia
398 and Benson 2019).

399 Finally, we cannot exclude the potential contribution of unobserved variables that were not moni-
400 tored during the experiment, such as variation in nutrient levels in the chemostat, and which may
401 also lead to differences in the predation and intra-specific interactions across systems (e.g. Bonsall,
402 Van Der Meijden, and Crawley 2003; Fussmann and Blasius 2005; Posey, Alphin, and Cahoon
403 2006).

404 Should we expect limited generalisability of dynamics across systems, even if the complexity of
405 the process is properly captured, environmental conditions known, and the system well-observed?
406 A similar study, that inferred dynamical processes consistency from replicated time series of a
407 simple rotifer system, found substantial variation in vital rates across replicates (Rosenbaum et al.
408 2019), also pointing at a low generalisability of dynamical processes. Yet, the level of replication
409 of the time series of their studies was not as stringent as that of the time series we considered,
410 which leaves room for variability in dynamics to be caused by differences in experimental setup,
411 population history, initial densities. Bruijning and colleagues also found substantial variation in
412 vital rates across clones in a replicated system of aphids, showing that slight phenotypic variations
413 can change the population dynamics, all else being equal (Bruijning, Jongejans, and Turcotte 2019).
414 This phenomenon is likely to be even more important in more complicated systems and in a natural

415 setting where most variables are unobserved, which poses a problem for the generalisation of results
416 across studies and systems (De Meester et al. 2019). How can we expect to generalise dynamics
417 across real systems if we are not able to do so in artificial systems? Overall, our study reinforces
418 the view that general inferences should not be drawn from a single system, and that more efforts
419 are required to distinguish dynamical patterns that are conserved across systems from idiosyncratic
420 ones.

421 Can we trust our models then if they are doomed to provide partly idiosyncratic answers? Our
422 study demonstrates that processes can vary substantially across replicates, so that there may hence
423 not be a single suitable functional form and parametrisation to model them (Lawton 1999). Yet,
424 most of the work to date has involved fitting parametric models to time series data (e.g. Bruijning,
425 Jongejans, and Turcotte 2019; Pontarp, Brännström, and Petchey 2019; Rosenbaum et al. 2019),
426 which provide a very narrow view of the range of possible functions to describe the biological
427 processes at play (Jost and Ellner 2000; Adamson and Morozov 2013). These models are subjective
428 by nature (Jost and Ellner 2000; Adamson and Morozov 2013), and hence not generalisable, so that
429 they greatly reduce our chance at identifying dynamical processes that are idiosyncratic, and those
430 that are transferable.

431 What alternatives do we have then? We propose that NODEs are a suitable framework to study
432 dynamical processes, as they produce inferences that are free of model assumption and facilitate
433 comparison across studies and systems (Bonnaffé, Sheldon, and Coulson 2021). In this sense, our
434 study already provides a potentially more objective depiction of dynamical processes than previous

work with parametric models. Furthermore, in this paper we overcame the practical challenges of implementing NODEs by providing a computationally efficient fitting procedure, relying on time series interpolation, and developed a model selection criterion robust to overfitting. Similar approaches have been proposed in the past, for instance Ellner and colleagues developed a method called gradient matching where they interpolated the data with cubic splines to which they fitted the differential equations (Jost and Ellner 2000; Ellner, Seifu, and Smith 2002). Wu and colleagues also relied on data interpolation of the data with ANNs to fit non-parametric approximations of population vital rates (Wu, Fukuhara, and Takeda 2005). But the approaches were too challenging and cumbersome to be implemented routinely, and were not used to tackle ecological interactions. Overall, our work demonstrates the usefulness of NODEs for inferring ecological interactions from count time series, which could readily be applied to a substantial pool of time series data.

Conclusion

Generalising dynamics across biological systems is hard because of the complexity of the dynamical processes (e.g. ecological interactions), differences in environmental context, and monitoring limitations. It remains unclear whether we could generalise dynamics if we properly modelled complexity, controlled for environmental effects, and observed systems precisely. We addressed this question by looking at the generalisability of dynamical processes across three replicated time series of a three-species system, using the novel framework of NODEs. We found that only the dominant interactions were conserved across the three time series, namely that between the algae and the rotifer, while the role of the intermediate predator varied substantially. Our results hence

455 suggest that generalisation may not seem possible, even in simple system with no environmental
456 variation. Given previous work in this system, the main cause of differences across replicates may
457 be evolution in prey defence traits. We conclude that more work is required, using NODEs, to
458 identify dynamical patterns that are conserved and those that are idiosyncratic across a wider range
459 of systems.

460 **Acknowledgments**

461 We thank warmly the Ecological and Evolutionary Dynamics Lab and Sheldon Lab Group at the
462 department of Zoology for their feedback and support. We thank Ben Sheldon for insightful sug-
463 gestions on early versions of the work. The work was supported by the Oxford-Oxitec scholarship
464 and the NERC DTP.

465 **Data accessibility**

466 All data and code will be made fully available at <https://github.com/WillemBonnafe/NODER/rotifer>.

467 **Statement of authorship**

468 Willem Bonnaffé designed the method, performed the analysis, wrote the manuscript; Tim Coulson
469 led investigations, provided input for the manuscript, commented on the manuscript.

470 References

- 471 Adamson, M. W. and A. Y. Morozov (2013). “When can we trust our model predictions? Un-
472 earthing structural sensitivity in biological systems”. In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 469.2149, pp. 1–19.
- 473
- 474 Arndt, H. (1993). “Rotifers as predators on components of the microbial web (bacteria, heterotrophic
475 flagellates, ciliates) - a review”. In: *Hydrobiologia* 255-256.1, pp. 231–246.
- 476 Becks, L., S. P. Ellner, L. E. Jones, and N. G. J. Hairston (2010). “Reduction of adaptive ge-
477 netic diversity radically alters eco-evolutionary community dynamics”. In: *Ecology Letters* 13.8,
478 pp. 989–997.
- 479 – (2012). “The functional genomics of an eco-evolutionary feedback loop: Linking gene expres-
480 sion, trait evolution, and community dynamics”. In: *Ecology Letters* 15.5, pp. 492–501.
- 481 Bonnaiffé, W., A. Danet, S. Legendre, and E. Edeline (2021). “Comparison of size-structured and
482 species-level trophic networks reveals antagonistic effects of temperature on vertical trophic
483 diversity at the population and species level”. In: *Oikos* 130.8, pp. 1297–1309.
- 484 Bonnaiffé, W., B. C. Sheldon, and T. Coulson (2021). “Neural ordinary differential equations for
485 ecological and evolutionary time series analysis”. In: *Methods in Ecology and Evolution* 2, pp. 1–
486 46.
- 487 Bonsall, M. B., E. Van Der Meijden, and M. J. Crawley (2003). “Contrasting dynamics in the same
488 plant-herbivore interaction”. In: *Proceedings of the National Academy of Sciences of the United*
489 *States of America* 100.25, pp. 14932–14936.

490 Bruijning, M., E. Jongejans, and M. M. Turcotte (2019). “Demographic responses underlying eco-
 491 evolutionary dynamics as revealed with inverse modelling”. In: *Journal of Animal Ecology* 88.5,
 492 pp. 768–779.

493 Calder, C., M. Lavine, P. Müller, and J. S. Clark (2003). “Incorporating multiple sources of stochas-
 494 ticity into dynamic population models”. In: *Ecology* 84.6, pp. 1395–1402.

495 Cawley, G. C. and N. L. C. Talbot (2007). “Preventing over-fitting during model selection via
 496 bayesian regularisation of the hyper-parameters”. In: *Journal of Machine Learning Research* 8,
 497 pp. 841–861.

498 Chen, T. and H. Chen (1993). “Approximations of Continuous Functionals by Neural Networks
 499 with Application to Dynamic Systems”. In: *IEEE Transactions on Neural Networks* 4.6, pp. 910–
 500 918.

501 Dallas, T., B. A. Melbourne, G. Legault, and A. Hastings (2021). “Initial abundance and stochas-
 502 ticity influence competitive outcome in communities”. In: *Journal of Animal Ecology*, pp. 1–
 503 26.

504 De Meester, L. et al. (2019). “Analysing eco-evolutionary dynamics—The challenging complexity
 505 of the real world”. In: *Functional Ecology* 33.1, pp. 43–59.

506 Demyanov, V., S. N. Wood, and T. J. Kedwards (2006). “Improving ecological impact assessment
 507 by statistical data synthesis using process-based models”. In: *Journal of the Royal Statistical*
 508 *Society. Series C: Applied Statistics* 55.1, pp. 41–62.

509 Ellner, S. P., Y. Seifu, and R. H. Smith (2002). “Fitting Population Dynamic Models to Time-Series
 510 Data by Gradient Matching”. In: *Ecology* 83.8, p. 2256.

511 Ezard, T. H. G., S. D. Côté, and F. Pelletier (2009). “Eco-evolutionary dynamics: Disentangling
512 phenotypic, environmental and population fluctuations”. In: *Philosophical Transactions of the*
513 *Royal Society B: Biological Sciences* 364.1523, pp. 1491–1498.

514 Fox, G. A. and B. E. Kendall (2002). “Demographic stochasticity and the variance reduction ef-
515 fect”. In: *Ecology* 83.7, pp. 1928–1934.

516 Funahashi, K.-i. and Y. Nakamura (1993). “Approximation of dynamical systems by continuous
517 time recurrent neural networks”. In: *Neural Networks* 6.6, pp. 801–806.

518 Fussmann, G. F. and B. Blasius (2005). “Community response to enrichment is highly sensitive to
519 model structure”. In: *Biology Letters* 1.1, pp. 9–12.

520 Gamelon, M. et al. (2019). “Accounting for interspecific competition and age structure in demo-
521 graphic analyses of density dependence improves predictions of fluctuations in population size”.
522 In: *Ecology Letters* 22.5, pp. 797–806.

523 Gilbert, J. J. (1976). “Selective cannibalism in the rotifer *Asplanchna sieboldi*: Contact recognition
524 of morphotype and clone”. In: *Proceedings of the National Academy of Sciences* 73.9, pp. 3233–
525 3237.

526 Gross, K., A. R. Ives, and E. V. Nordheim (2005). “Estimating fluctuating vital rates from time-
527 series data: A case study of aphid biocontrol”. In: *Ecology* 86.3, pp. 740–752.

528 Hairston, N. G. J., S. P. Ellner, M. A. Geber, T. Yoshida, and J. A. Fox (2005). “Rapid evolution and
529 the convergence of ecological and evolutionary time”. In: *Ecology Letters* 8.10, pp. 1114–1127.

530 Hiltunen, T., L. E. Jones, S. P. Ellner, and N. G. J. Hairston (2013). “Temporal dynamics of a simple
531 community with intraguild predation: an experimental test”. In: *Ecology* 94.4, pp. 773–779.

532 Jia, J. and A. R. Benson (2019). “Neural jump stochastic differential equations”. In: *Advances in*
533 *Neural Information Processing Systems* 32.NeurIPS.

534 Jost, C. and S. P. Ellner (2000). “Testing for predator dependence in predator-prey dynamics: A
535 non-parametric approach”. In: *Proceedings of the Royal Society B: Biological Sciences* 267.1453,
536 pp. 1611–1620.

537 Kendall, B. E. et al. (2005). “Population cycles in the pine looper moth: Dynamical tests of mech-
538 anistic hypotheses”. In: *Ecological Monographs* 75.2, pp. 259–276.

539 Lawton, J. H. (1999). “Are There General Laws in Ecology ?” In: *Oikos* 84.2, pp. 177–192.

540 Pearce, T., F. Leibfried, A. Brintrup, M. Zaki, and A. Neely (2018). “Uncertainty in Neural Net-
541 works: Approximately Bayesian Ensembling”. In: *arXiv*, pp. 1–10.

542 Pontarp, M., Å. Brännström, and O. L. Petchey (2019). “Inferring community assembly processes
543 from macroscopic patterns using dynamic eco-evolutionary models and Approximate Bayesian
544 Computation (ABC)”. In: *Methods in Ecology and Evolution* 10.4, pp. 450–460.

545 Posey, M. H., T. D. Alphin, and L. Cahoon (2006). “Benthic community responses to nutrient en-
546 richment and predator exclusion: Influence of background nutrient concentrations and interactive
547 effects”. In: *Journal of Experimental Marine Biology and Ecology* 330.1, pp. 105–118.

548 Raeymaekers, J. A. M. et al. (2017). “Adaptive and non-adaptive divergence in a common land-
549 scape”. In: *Nature Communications* 8.1, pp. 1–8.

550 Reznick, D. N., H. Bryga, and J. A. Endler (1990). “Experimentally induced life-history evolution
551 in a natural population”. In: *Nature* 346.6282, pp. 357–359.

552 Rosenbaum, B., M. Raatz, G. Weithoff, G. F. Fussmann, and U. Gaedke (2019). “Estimating param-
553 eters from multiple time series of population dynamics using bayesian inference”. In: *Frontiers*
554 *in Ecology and Evolution* 6.234, pp. 1–14.

555 Shurin, J. B., J. L. Clasen, H. S. Greig, P. Kratina, and P. L. Thompson (2012). “Warming shifts
556 top-down and bottom-up control of pond food web structure and function.” In: *Philosophical*
557 *transactions of the Royal Society of London. Series B, Biological sciences* 367.1605, pp. 3008–
558 17.

559 Snell, T. W. and B. L. Garman (1986). “Encounter probabilities between male and female rotifers”.
560 In: *Journal of Experimental Marine Biology and Ecology* 97.3, pp. 221–230.

561 Thompson, C. E., E. B. Taylor, and J. D. Mcphail (1997). “Parallel Evolution of Lake-Stream
562 Pairs of Threespine Sticklebacks (*Gasterosteus*) Inferred from Mitochondrial DNA Variation”.
563 In: *Evolution* 51.6, pp. 1955–1965.

564 Wu, J., M. Fukuhara, and T. Takeda (2005). “Parameter estimation of an ecological system by a
565 neural network with residual minimization training”. In: *Ecological Modelling* 189.3-4, pp. 289–
566 304.

567 Yoshida, T., S. P. Ellner, L. E. Jones, B. J. M. Bohannan, R. E. Lenski, and N. G. J. Hairston (2007).
568 “Cryptic population dynamics: Rapid evolution masks trophic interactions”. In: *PLoS Biology*
569 5.9, pp. 1868–1879.

570 Yoshida, T., L. E. Jones, S. P. Ellner, G. F. Fussmann, and N. G. J. Hairston (2003). “Rapid evo-
571 lution drives ecological dynamics in a predator – prey system”. In: *Nature* 424.July, pp. 303–
572 306.

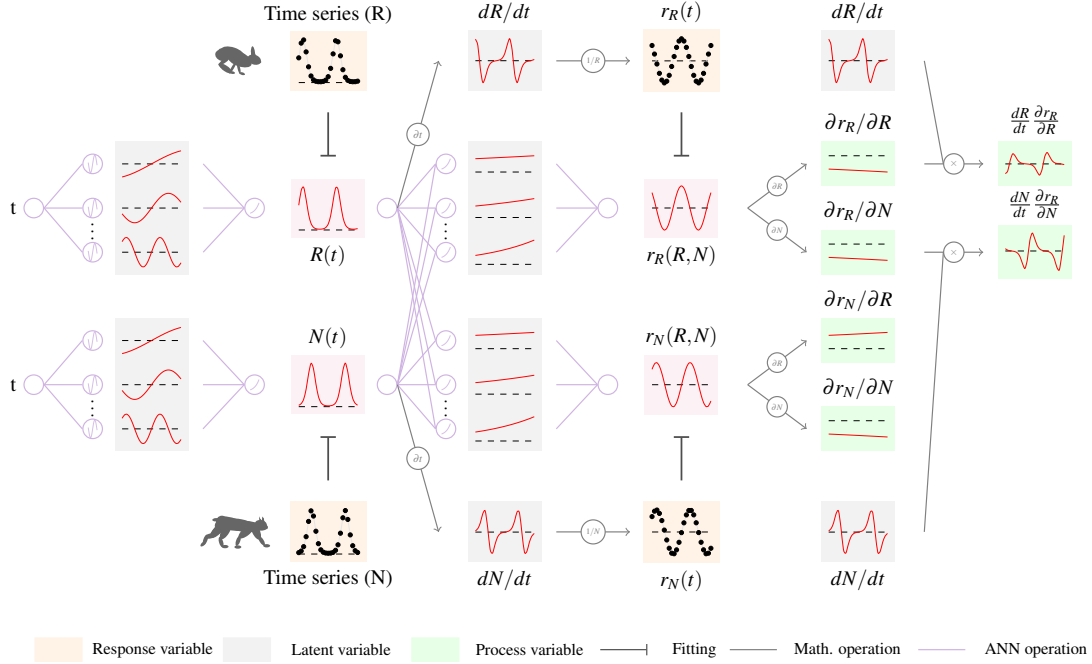


Figure 1: Overview of fitting neural ordinary differential equations by gradient matching

The first step is to compute a continuous time approximation (interpolation) of each state variables (e.g. resource $R(t)$ and predator $N(t)$). To do that we fit an ANN, that takes time as input, to each time series. Dynamics of populations can then be computed by taking the derivative of the ANN with respect to time, dR/dt and dN/dt . This provides an interpolation of the per-capita growth rate of each population, e.g. $r_R(t) = 1/R dR/dt$. In a second step, we approximate non-parametrically the per-capita growth rates with respect to the density of each populations, $r_R = s(R, N)$. To do that we fit an ANN, which takes as input the interpolated variables $R(t)$ and $N(t)$, to the interpolated per-capita growth rates $r_R(t)$ and $r_N(t)$. In a final step, we approximate the ecological interactions, by computing the sensitivity of the per-capita growth rates with respect to the density of each population, e.g. $E : N \rightarrow R = \partial r_R / \partial N$. We also compute the contribution of each species to the dynamics of the other by multiplying the dynamics of each variable with its effect on the growth rates (i.e. the Geber method), e.g. $C : N \rightarrow R = dN/dt \times \partial r_R / \partial N$.

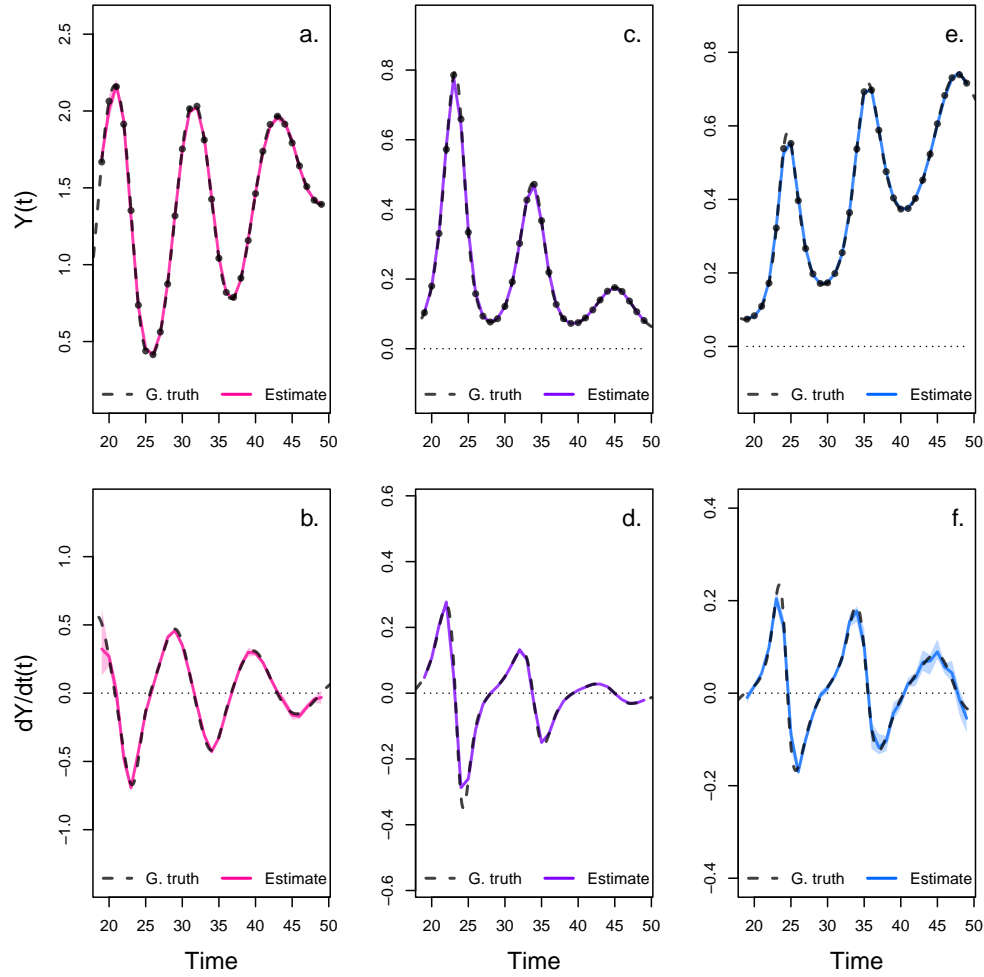


Figure 2: Interpolated density and dynamics of algae, flagellate, and rotifer in the artificial system. This figure corresponds to the first step in the overview figure. It shows the accuracy of the interpolated densities of algae (a.), flagellate (c.), and rotifer (e.). We obtain interpolated densities by fitting observed densities (black dots) with ANNs that take time as input. The observed densities were obtained by sampling a tri-trophic prey-predator ODE model at regular time steps. We then derive interpolated dynamics (b., d., f.) by computing the temporal derivative of the interpolated densities with respect to time. In all graphs, the dashed line represents the ground truth, namely trajectories generated by the ODE model. The solid lines correspond to the interpolations. The shaded area shows the 90% confidence interval, obtained by approximately sampling the marginal posterior distributions.

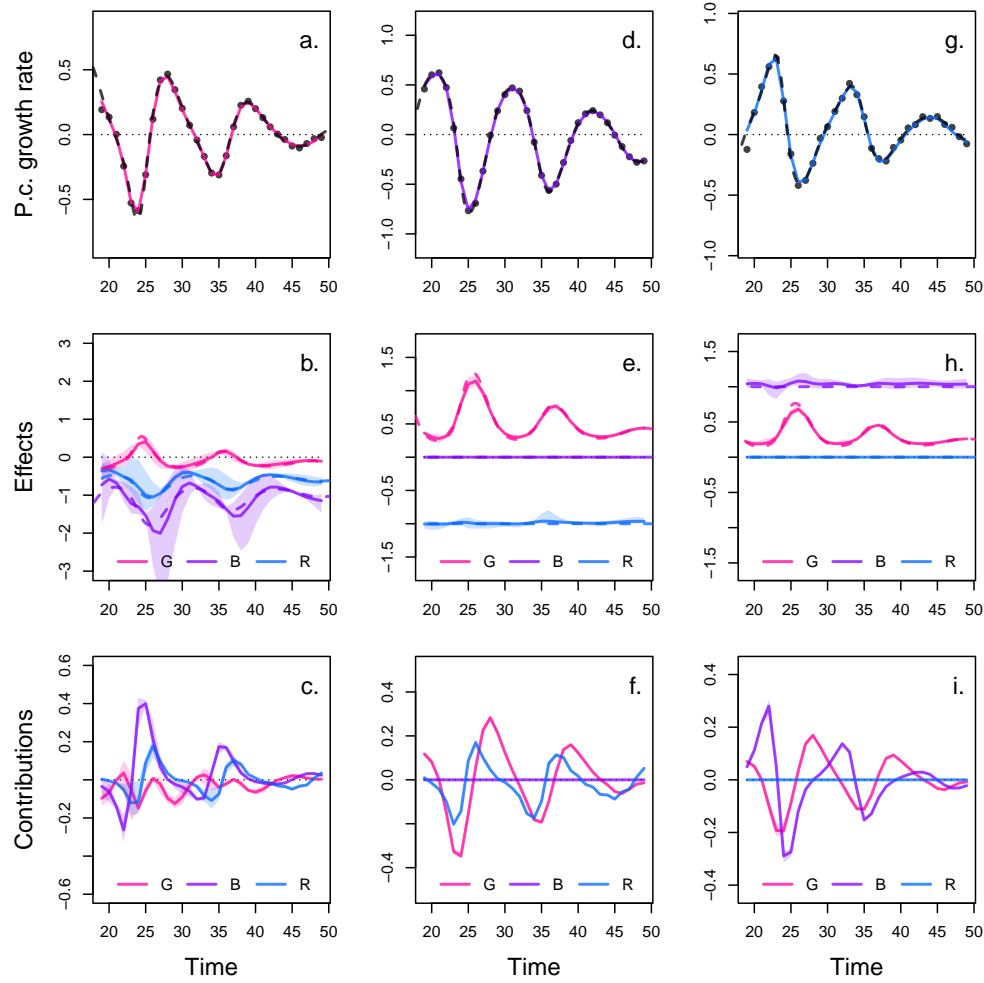


Figure 3: Drivers of dynamics of algae, flagellate, and rotifer in the artificial system. This figure corresponds to the second step in the overview figure. It displays the NODE non-parametric approximations of the per-capita growth rate of algae (a., b., c.), flagellate (d., e., f.), and rotifer (g., h., i.). We obtain the NODE approximations (a., d., g., solid line) by fitting the interpolated per-capita growth rates (black dots) with ANNs that take population densities as input. We then estimate the direction of ecological interactions (effects, b., e., h.) by computing the derivative of the NODE approximations with respect to each density. Finally, we compute the strength of ecological interactions (contributions, c., f., i.) by multiplying the interpolated dynamics of each population (fig. 1, b., d., f.) with its effects. Dashed lines correspond to ground truth, obtained from the original trajectories of the tri-trophic ODE model. The shaded area shows the 90% confidence interval, obtained by approximately sampling the posterior distributions.

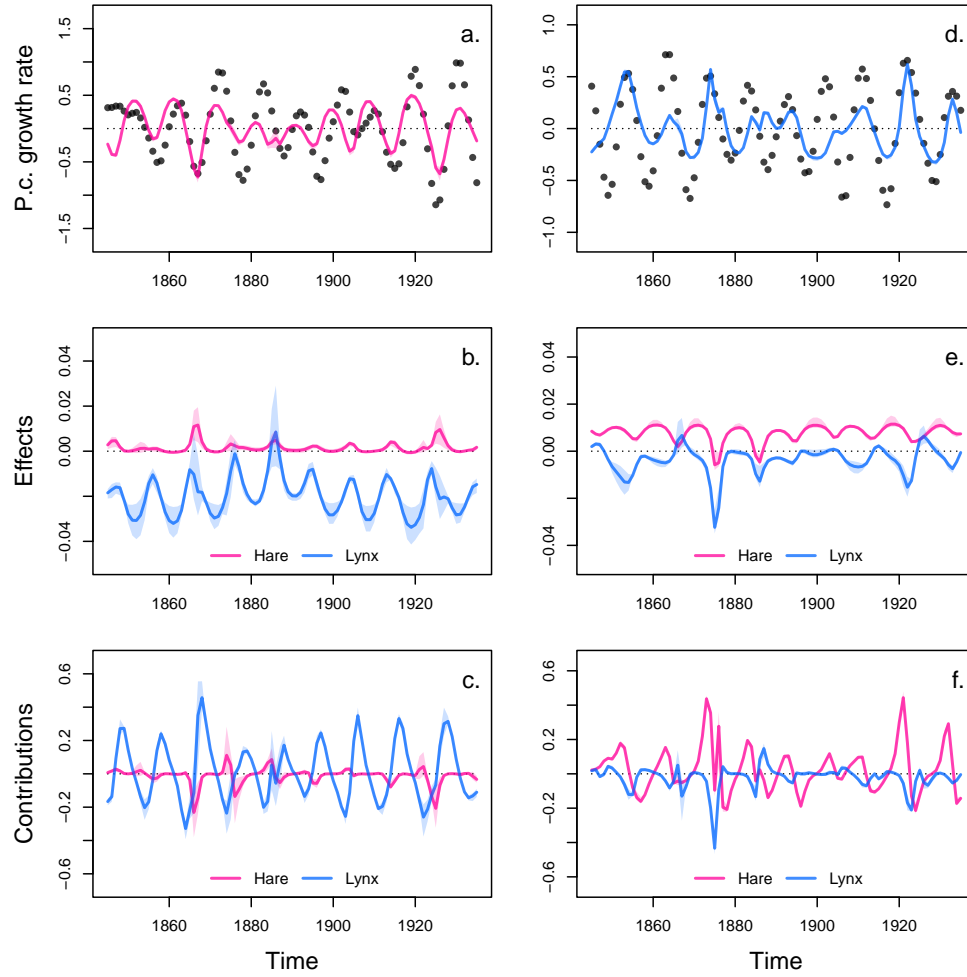


Figure 4: Drivers of dynamics of hare and lynx in the Odum and Barrett pelt count time series. This figure displays the NODE non-parametric approximations of the per-capita growth rate of hare (a., b., c.), and lynx (d., e., f.). We obtain the NODE approximations (a., d., solid line) by fitting the interpolated per-capita growth rates (black dots) with ANNs that take population densities as input. We then estimate the direction of ecological interactions (effects, b., e.) by computing the derivative of the NODE approximations with respect to each density. Finally, we compute the strength of ecological interactions (contributions, c., f.) by multiplying the interpolated dynamics of each population with its effects. The shaded area shows the 90% confidence interval, obtained by approximately sampling the posterior distributions.

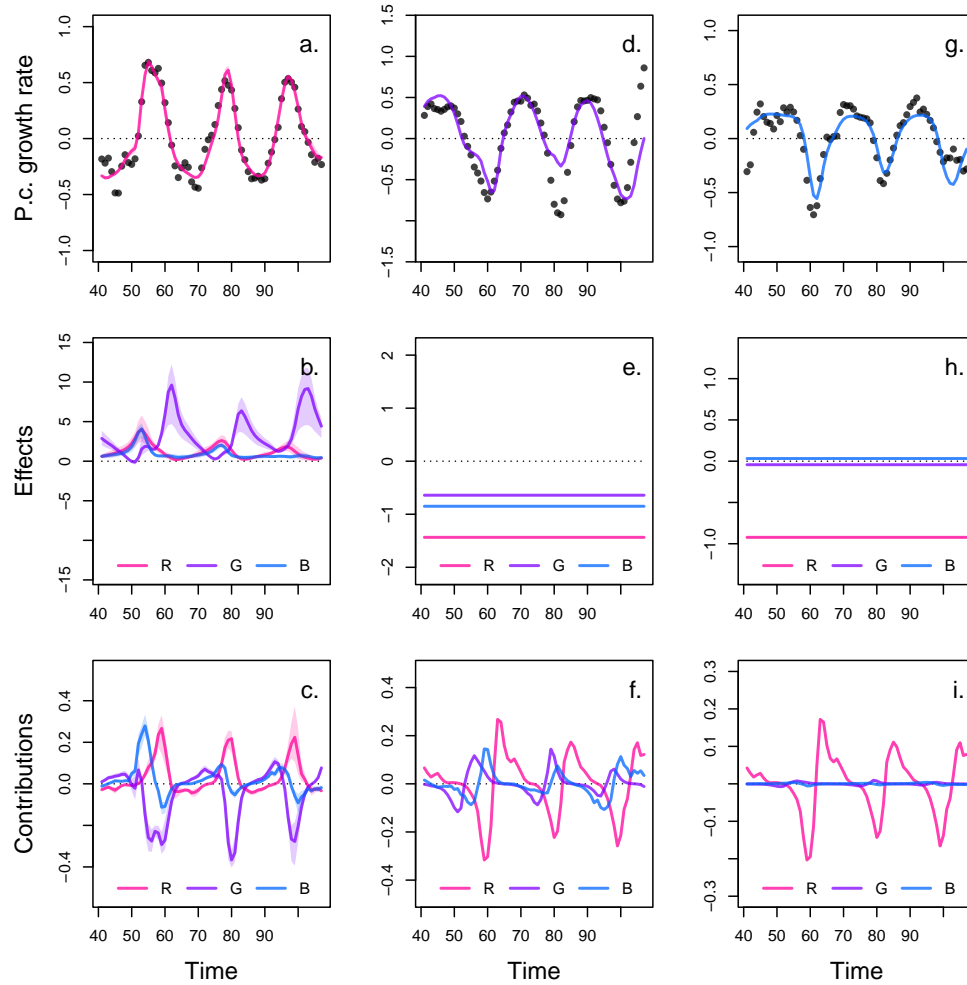


Figure 5: Drivers of dynamics of algae, flagellate, and rotifer in replicate A. This figure displays the NODE non-parametric approximations of the per-capita growth rate of algae (a., b., c.), flagellate (d., e., f.), and rotifer (g., h., i.). We obtain the NODE approximations (a., d., g., solid line) by fitting the interpolated per-capita growth rates (black dots) with ANNs that take population densities as input. We then estimate the direction of ecological interactions (effects, b., e., h.) by computing the derivative of the NODE approximations with respect to each density. Finally, we compute the strength of ecological interactions (contributions, c., f., i.) by multiplying the interpolated dynamics of each population with its effects. The shaded area shows the 90% confidence interval, obtained by approximately sampling the posterior distributions. The replicated time series were obtained by digitising the time series in Hiltunen et al. (2013).

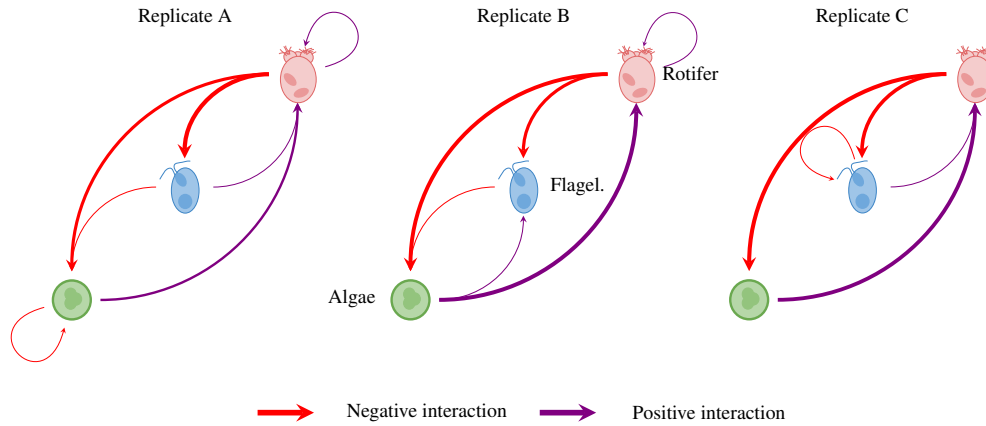


Figure 6: Interaction networks inferred from 3 replicated time series of algae, flagellate, and rotifers. This figure shows the direction and strength of ecological interactions inferred from 3 replicated sets of time series of algae, flagellate, and rotifer, using NODEs fitted by gradient matching. The replicates B and C were analysed in the same way as replicate A (see fig. 5 for details). Red and purple arrows correspond to negative or positive mean effects. We estimated mean effects by averaging effects (i.e. derivative of NODE approximated per-capita growth rates with respect to each population density) across the time series. The width of the arrows is proportional to the relative strength of the ecological interaction. We compute the relative strength as the % of total contributions attributable to either algae, flagellate, or rotifer, obtained from summing the square of contributions of each species throughout the time series. For instance in replicate A, the relative strength of the effect of rotifer on algae is found by summing the square of the red line in fig. 5 f., and computing the % of total contributions that it accounts for. We provide the value of the mean effects and relative strengths in Table 1. The replicated time series were obtained by digitising the time series in Hiltunen et al. (2013).

Table 1: Summary analysis. r^2 corresponds to the r squared of the NODE non-parametric approximation of the pre-capita growth rate compared to the interpolated per-capita growth rate for each of the three species. Mean effects are obtained by averaging the effect of one species on the growth rate of another throughout the time series. The % of total contributions is obtained by summing the square of contributions of one species density to the growth of the other at each time step throughout the time series, then by computing the proportion of total change that it accounts for.

		R	G	B
replicate A				
Mean effects	on R	0.27	0.77	0.97
	on G	-1.17	-0.44	-0.85
	on B	-0.78	0.04	0.03
% of total contributions	to R	0.08	0.48	0.44
	to G	0.75	0.08	0.17
	to B	1	0	0
replicate B				
Mean effects	on R	0.08	0.59	0.22
	on G	-1	0.05	-0.48
	on B	-0.47	0.14	-0.02
% of total contributions	to R	0.02	0.93	0.05
	to G	0.9	0	0.1
	to B	0.9	0.1	0
replicate C				
Mean effects	on R	-0.1	0.45	0.93
	on G	-1.76	-0.13	-0.12
	on B	-0.76	0.01	0.08
% of total contributions	to R	0.01	0.31	0.67
	to G	0.99	0.01	0
	to B	0.99	0	0.01

573 6 Supplementary

574 A Bayesian regularisation

575 In this section we describe how to derive the modified model selection criteria developed by Caw-
576 ley and Talbot (Cawley and Talbot 2007). Bayesian regularisation simply amounts to constraining
577 the values of the parameters in the model to be close to a desired value. Usually, parameters are
578 constrained by choosing normal priors centered about 0. In this case, the standard deviation of the
579 normal priors governs the range of values that the parameters can take, and hence constrains more
580 or less strongly the behaviour of the model (Cawley and Talbot 2007). Performing inference on the
581 second level means that we are trying to find the appropriate value of the dispersion of the priors,
582 in other words, the appropriate level of constraint on the model. In practice, choosing the level of
583 constraint is difficult, Cawley and Talbot hence developed a criterion to perform model selection
584 on the second level of inference. They proposed to optimise the marginal posterior distribution by
585 averaging out the dispersion of the priors. With an appropriate choice of prior, the dispersion can
586 be integrated out, leaving us with a formula for the posterior that only depends on the parameters
587 of the model,

$$\log P(\theta|\mathcal{D}) \propto \frac{I}{2} \log \left(\sum_{i=1}^I e_i(\mathcal{D}, \theta)^2 \right) + \frac{J}{2} \log \left(\sum_{j=1}^J \theta_j^2 \right) \quad (16)$$

588 where $P(\theta|\mathcal{D})$ denotes the marginal posterior density, \mathcal{D} denotes the evidence, I and J denote the

number of data points and parameters, respectively, e_i denote the residuals of the model, and θ denote the parameters of the model. The construction is elegant because it is not sensitive to the choice of prior hyperparameters, and simple as it amounts to optimising the log of the sum of squares, rather than the sum of squares (in the case of normal ordinary least square).

The issue with this formula is that the marginal posterior density is infinity when the parameters are 0, which leads to underfitting. In this paper we use a modified criterion, which corrects for that problem

$$\log P(\theta|\mathcal{D}) \propto \frac{I}{2} \log \left(1 + \sum_{i=1}^I e_i(\mathcal{D}, \theta)^2 \right) + \frac{J}{2} \log \left(1 + \sum_{j=1}^J \theta_j^2 \right) \quad (17)$$

where the marginal posterior density depends only on the residuals of the model when the parameters are equal to 0, and otherwise depends on both the parameters and the residuals. This construction can be obtained simply by assuming a gamma prior for the parameters $p(\xi) \propto \frac{1}{\xi} \exp\{-\xi\}$, where ξ is the regularisation parameter, instead of the improper Jeffreys' prior that Cawley and Talbot used in their original study, namely $p(\xi) \propto \frac{1}{\xi}$. The details of the integration of the posterior distribution over ξ can be found in Cawley and Talbot's original paper.