

# Inconsistent species interactions across replicated systems may hinder generalisation of dynamical processes

Willem Bonnaffé<sup>1</sup> and Tim Coulson<sup>2</sup>

<sup>1</sup>University of Oxford

<sup>2</sup>Oxford

April 05, 2024

## Abstract

Generalisation of dynamical processes across natural systems is difficult because of their complexity and unobserved variables. The hope is that generalisation may be achieved if we model adequately the complexity of systems, and observe them in sufficient detail. Yet, there is still limited support for this claim. 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?

# Inconsistent species interactions across replicated systems may hinder generalisation of dynamical processes

Willem Bonnaffé<sup>1,2</sup> and Tim Coulson<sup>2</sup>

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; ben.sheldon@zoo.ox.ac.uk; tim.coulson@zoo.ox.ac.uk

**Running title:** Inconsistent dynamics prevent generalisation

**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;

**Article type:** *Letters*

**Specifications:** 149 words in abstract; 4957 words in text; 29 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; Tim Coulson led investigations, provided input for the manuscript, commented on the manuscript.

## **Abstract**

Generalisation of dynamical processes across natural systems is difficult because of their complexity and unobserved variables. The hope is that generalisation may be achieved if we model adequately the complexity of systems, and observe them in sufficient detail. Yet, there is still limited support for this claim. 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

Dynamical processes, such as ecological interactions, drive population dynamics and generate dynamical patterns. Generalising dynamical processes across biological systems is hard. This is due to biological contingencies, stemming from the complexity of biological systems (esp. ecological interactions), differences in environmental contexts (i.e. unobserved variables), and observation error (De Meester et al. 2019). These contingencies alter dynamical processes by, for instance, modifying the strength of species interactions, which ultimately results in variation in dynamics across systems with the same species composition and population structure. This prevents the identification of global dynamical models, and therefore hinders generalisation, repeatability, and transfer of knowledge across systems and studies (Lawton 1999).

This problem has been repeatedly identified in natural systems. For instance, different access to seed supplies can modify the strength of the interaction between a plant and its herbivore, leading to either stable or oscillatory dynamics (Bonsall, Van Der Meijden, and Crawley 2003). Differences in temperature can alter the ecological interaction structure of entire ecosystems (Shurin et al. 2012; Bonnaffé et al. 2021). Vital rates are often found to be inconsistent in time (Gross, Ives, and Nordheim 2005; Adamson and Morozov 2013), and space (e.g. Gamelon et al. 2019). Attempts at identifying a single population dynamics model in two mesocosms have led to partial fits, as the model could not accomodate the dynamics of the two different systems (Demyanov, Wood, and Kedwards 2006). A growing body of evidence showing that generalisation of dynamical processes across similar systems often fails (Lawton 1999, e.g. Kendall et al. 2005; Ezard, Côté, and Pelletier

21 2009).

22 A burning question is to what extent we would be able to generalise dynamics across systems if  
23 we were to properly account for contingencies, by appropriately modelling the complexity of the  
24 system, the structure of the environment, and reduce observation errors. This is difficult to assess  
25 in practice, especially in a natural setting (Brunner et al. 2019; De Meester et al. 2019). However,  
26 it may be possible in an artificial setting, where contingencies can be virtually eliminated. In spite  
27 of this there are few studies that have attempted to characterise the generalisability of dynamics  
28 across replicated systems in a laboratory setting. In such a setting, idiosyncrasies in population dy-  
29 namics can arise from (1) variations in ecological interactions and individual processes, as a result  
30 of evolution (e.g. Yoshida et al. 2003), or stochasticity (Dallas et al. 2021), (2) variations in initial  
31 conditions due to the experimental setting (Yoshida et al. 2003; De Meester et al. 2019), and (3) the  
32 complexity of the system which can lead to large changes in system dynamics with small changes  
33 in the system state and structure (Adamson and Morozov 2013). Two studies, one in aphids and  
34 the other in rotifers, found substantial variation in vital rates across replicated populations, by fit-  
35 ting a stage-structured population ODE model to population dynamics time series data (Bruijning,  
36 Jongejans, and Turcotte 2019; Rosenbaum et al. 2019). These studies hint that generalisability of  
37 population dynamical processes may not be possible because of intrinsic population structure and  
38 evolution, even in virtually identical populations hosted in artificial environments.

39 We identified three gaps in the literature. First, this kind of evidence remains scarce, due in part  
40 to the fact that dynamical modelling approaches guided by empirical data are still not widespread

41 (Pontarp, Brännström, and Petchey 2019). Second, most of these studies relied on parametric  
42 frameworks, which impose arbitrary pre-determined forms for the dynamical processes at play, so  
43 that their model may not capture properly the complexity of the dynamics of these populations (Jost  
44 and Ellner 2000; Adamson and Morozov 2013; Bonnaffé, Sheldon, and Coulson 2021). Finally,  
45 most studies usually analyse dynamics in single-species systems, but not multi-species systems,  
46 such as those with intraguild predation, which are more biologically realistic scenarios (Hiltunen  
47 et al. 2013). Further studies are consequently required to investigate the consistency of dynamical  
48 processes in simple multi-species and well-observed systems, to conclude about the generalisability  
49 of population dynamics across systems.

50 Our aim in this study is to provide an assessment of the consistency of dynamical processes in a  
51 simple multi-species system hosted in a controlled environment. We do this by quantifying the  
52 direction, strength, and consistency of interactions in time and across replicates of a simple bio-  
53 logical system in an experimental setting. We hypothesise that if the system is (1) simple enough,  
54 (2) well-observed, (3) in a controlled environment, then dynamical effects/interactions should be  
55 broadly consistent in time and across replicates, hence allowing for generalisation of dynamics  
56 across systems. We consider three replicates of a three-species system, consisting in a prey (al-  
57 gae), intermediate-predator (flagellate), and top-predator (rotifer). The algae is consumed by the  
58 flagellate and rotifer, and the flagellate is consumed by the rotifer. We use three replicated system  
59 runs from a study by Hiltunen and colleagues which feature sequential oscillations of the den-  
60 sity of the three species (Hiltunen et al. 2013). We analyse the time series with neural ordinary

61 differential equations (Bonnaffé, Sheldon, and Coulson 2021), which allows us to approximate  
62 non-parametrically population growth rates, and quantify the direction, strength, and consistency  
63 of inter- and intra-specific effects on the growth of each population. We find that the interaction  
64 between the rotifer and algae is consistent throughout time and across replicates, while the inter-  
65 action between the flagellate and the two other species is not. Our study suggests that dynamical  
66 processes may sometimes not be consistent and generalisable across systems, even when they are  
67 as close to identical as experimentation permits. We discuss these results and hint at the underlying  
68 impact of evolution driving differences in these systems.

## 69 **2 Material and Methods**

### 70 **2.1 Method overview**

71 The aim of the study is to determine the extent to which the strength and direction of interactions  
72 between three species in a tri-trophic prey-predator system are consistent in time and across repli-  
73 cates, as a way to assess the generalisability of dynamics across simple and well-observed systems.  
74 To do this we approximate the dynamics of each species by fitting neural ordinary differential equa-  
75 tions (NODEs, Bonnaffé, Sheldon, and Coulson 2021) to replicated time series data of changes in  
76 prey and predator densities. We then derive the interactions between species by looking at the sen-  
77 sitivity of the dynamics to a change in the density of each species, and assess their consistency in  
78 time and across replicates.

## 79 2.2 System

80 We consider a three-species laboratory microcosm consisting of an algal prey (*Chlorella autroph-*  
81 *ica*), a flagellate intermediate predator (*Oxyrrhis marina*), and a rotifer top predator (*Brachionus*  
82 *plicatilis*). The algal prey is consumed by the intermediate and top predator, the top predator also  
83 consumes the intermediate predator. The dynamics of this system, here the daily change in the  
84 density of each species, were recorded in three replicated time series experiments performed by  
85 Hiltunen and colleagues (Hiltunen et al. 2013, Fig. 1). The aim of their experiment was to deter-  
86 mine which type of population dynamics would arise in a system with two predators competing  
87 for the same resource (the algae), where one predator (the rotifer) would also be able to consume  
88 its competitor (the flagellate). According to their expectations, they found prey-predator oscilla-  
89 tions, where the lag between the density peaks of each species reflected their position in the food  
90 web. Namely that the peak of algae preceded the flagellate peak, which itself preceded the rotifer  
91 peak.

92 Their microcosms are close to true replicates in that environmental conditions, namely temperature,  
93 salinity, and nutrient influx, were maintained constant, and initial conditions, that is the initial  
94 density of each species were shared across all replicates. In spite of that, they still found evidence  
95 for algae evolution in some parts of the time series, which resulted in a shift of the dynamics from  
96 fast prey-predator cycles to slower oscillations, similar to those documented in previous studies on  
97 similar systems (Yoshida et al. 2003), even in lineages where genetic variation in predator defense  
98 traits was eliminated at the start of the experiment. Consequently, the time series that they reported

99 are the ones that did not present evidence of evolution, and therefore displayed purely ecological  
100 dynamics.

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

### 106 **2.3 Model specifications**

107 The aim of the modelling approach is to infer the drivers of the dynamics of each species from  
108 the time series data. More specifically, we want to quantify the effect of a change in the density  
109 of one species on the dynamics of the other species. In this way we can understand which, and  
110 to what extent, species interactions drive population dynamics. To do this we use neural ordi-  
111 nary differential equation (NODEs), which is a novel methodology allowing us to infer dynamical  
112 processes non-parametrically from time series data (Bonnaffé, Sheldon, and Coulson 2021). We  
113 choose this methodology over traditional approaches because it offers two advantages. The first  
114 lies in the fact that NODEs approximate the dynamics of populations non-parametrically, and are  
115 therefore not subject to incorrect model specifications (Jost and Ellner 2000; Adamson and Moro-  
116 zov 2013). This is important as it offers an unbiased estimation of the inter-dependences between  
117 state variables, and hence a reliable assessment of whether a species is contributing to the dynamics

118 of another. The second advantage is that it is a dynamical systems approach, which means that the  
 119 effects are estimated in a dynamically consistent system of ODEs (Bonnaffé, Sheldon, and Coulson  
 120 2021). This is useful because it accounts for the dynamical nature of the system, so that it includes  
 121 lag effects, not just direct correlations between variables.

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

$$\begin{aligned}
 \frac{dR}{dt} &= r_R(R, G, B, \beta_R)R \\
 \frac{dG}{dt} &= r_G(R, G, B, \beta_G)G \\
 \frac{dB}{dt} &= r_B(R, G, B, \beta_B)B
 \end{aligned} \tag{1}$$

123 where  $dR/dt$ ,  $dG/dt$ , and  $dB/dt$  denote the change in rotifer ( $R$ ), algae ( $G$ ), and flagellate ( $B$ )  
 124 density in continuous time. The per-capita growth rates  $r_R$ ,  $r_G$ , and  $r_B$  are non-parametric functions  
 125 of the density  $R$ ,  $G$ ,  $B$  of each species. The shapes of the non-parametric functions are controlled  
 126 by the parameter vectors  $\beta_R$ ,  $\beta_G$ , and  $\beta_B$ . Fitting the NODE system (1) amounts to finding the  
 127 parameter vectors, and thereby the per-capita growth rates, that best describe the changes in density  
 128 observed in the time series data.

129 Each non-parametric functions is an artificial neural network (ANN). ANNs are powerful math-  
 130 ematical objects that can be trained to approximate the shape of dynamical processes (Funahashi  
 131 and Nakamura 1993). For the sake of simplicity, we consider the simplest form of an ANN which  
 132 contains a single hidden layer, namely a single layer peceptron (SLP)

$$r_R = \sum_{i=1}^N \beta_i f_{\sigma}(\beta_{i0} + \beta_{i1}R + \beta_{i2}G + \beta_{i3}B) \quad (2)$$

133 which takes as input the density of each species  $R$ ,  $G$ , and  $B$ , and output the corresponding per-  
 134 capita growth rate. The parameter vector  $\beta_R$ ,  $\beta_G$ ,  $\beta_B$ , contain the weight of the connections in the  
 135 ANNs. The SLP can be viewed as a weighted sum of basis functions  $f_{\sigma}$  of the state variables of  
 136 the system. In this study we consider sigmoid basis functions, as they are commonly used and  
 137 their capacity to approximate any continuous function is well established theoretically (Funahashi  
 138 and Nakamura 1993). The number of units in the hidden layer  $N$  is chosen to be 10, as this is  
 139 a commonly used number for systems of that size (e.g. Wu, Fukuhara, and Takeda 2005). More  
 140 details regarding these models can be found in our previous work (Bonnaffé, Sheldon, and Coulson  
 141 2021).

## 142 **2.4 Model fitting**

143 This section describes how to recover the parameters  $\beta$  of the NODE system given the time series  
 144 data at hand. In a previous study, we developed an approach to fit NODE systems to time series data  
 145 (Bonnaffé, Sheldon, and Coulson 2021). The technique relied on simulating the NODE system over  
 146 the times covered by the time series, and then computing the difference between the predictions  
 147 of the NODE model, and the observations of the time series. The model is fitted to the time  
 148 series by adjusting the parameter vectors until temporal dynamics of the state variables matched

the observations as closely as possible. There are two caveats with this approach that we solve in this study by opting for a different fitting approach. The first caveat is that the fitting process previously described is computationally expensive, because the NODE system has to be simulated over the entire range of the data at every step of the optimisation. Second, the simulation prevents the computation of gradients of the posterior distribution of the model, and thereby prevents the use of efficient gradient descent approaches. Overall, this makes it hard to avoid getting trapped in local maxima.

Instead, we propose an alternative fitting approach which relies on data interpolation to approximate populations state and dynamics. In this way we avoid the simulation step as the NODE system can then be directly compared to the interpolated dynamics. We proceed in two steps, (1) we interpolate the time series data to estimate the states and dynamics of each variable, and (2) we fit the NODE system directly to these estimated dynamics.

### **Interpolating the data**

We interpolate the time series and differentiate it with respect to time in order to approximate the dynamics of the system. The interpolation is found via a non-parametric regression of the interpolating function on the time series data

$$Y(t_i) = \tilde{Y}(t_i, \Omega) + \varepsilon_i \quad (3)$$

where  $Y(t_i)$  is the observed value of the variable at time  $t_i$  and  $\tilde{Y}(t_i, \Omega)$  is the value predicted by

166 the interpolating function, up to an error  $\varepsilon_i \sim \mathcal{N}(0, \sigma_1)$ . In the present case, the variables are  
 167 either the rotifer density  $R$ , algae density  $G$ , or flagellate density  $B$ , depending on which time  
 168 series is interpolated. The interpolating function is chosen to be an SLP with sinusoid activation  
 169 functions

$$\tilde{Y}(t, \Omega) = \sum_{i=1}^N \omega_{i0} \sin(\pi(\omega_{i1} + \omega_{i2}t)) \quad (4)$$

170 where  $\tilde{Y}(t, \Omega)$  is the interpolated state variable, either  $R$ ,  $G$ , or  $B$ , and is determined to be a weighted  
 171 sum of sinusoid functions of time. The interpolation parameter vector  $\Omega$  contains the weights  $\omega_{i0}$ ,  
 172  $\omega_{i1}$ , and  $\omega_{i2}$  which control the amplitude, shift, and frequency of the oscillations in the time series,  
 173 respectively. We found sinusoid activation functions to be most efficient for interpolating popula-  
 174 tion dynamics compared to other functions (such as sigmoid, hyperbolic). Following this approach  
 175 we obtain directly an approximation of the dynamics of the state variable by differentiating the SLP  
 176 with respect to time

$$\frac{\partial}{\partial t} \tilde{Y}(t) = \sum_{i=1}^N \omega_{i0} \pi \omega_{i2} \cos(\pi(\omega_{i1} + \omega_{i2}t)) \quad (5)$$

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

$$\tilde{r}_Y = \frac{1}{\tilde{Y}} \frac{\partial \tilde{Y}}{\partial t} \quad (6)$$

Overall, interpolating the data amounts to finding the parameter vector  $\Omega$  that minimises the error in equation (3).

### **Fitting NODEs to the interpolated data**

The second step is to match the NODE system to the interpolated dynamics, given the interpolated state variables. Thanks to the interpolation of the data, this simply amounts to performing a regression of the non-parametric approximation of the per-capita growth rate (equation 2) on the interpolated per-capita growth rate (equation 6)

$$\tilde{r}_Y(t_i) = r_Y(t_i, \beta) + \eta_i \quad (7)$$

up to an error term  $\eta_i \sim \mathcal{N}(0, \sigma_2)$ . For instance, the per-capita growth rate of the rotifer writes as  $\tilde{r}_R(t_i) = r_R(\tilde{R}(t_i), \tilde{G}(t_i), \tilde{B}(t_i), \beta_R) + \eta_i$ . Fitting the NODE per-capita growth rate hence amounts to finding the parameter vector  $\beta$  that minimises the error in equation 7, given the interpolation.

### **Statistical modelling approach**

The following section describes how to recover the parameters for the interpolation and NODE system that best describe the time series, while controlling for overfitting. The fitting of the models is performed in a Bayesian framework, considering normal error structure for the residuals, and

193 normal prior density distributions on the parameters

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

194 where  $\theta$  is the parameter vector of the model, and  $\mathcal{D}$  the evidence, namely the data that the model  
 195 is fitted to. In the case of the interpolation, the evidence is the population densities, either  $R(t)$ ,  
 196  $G(t)$ , or  $B(t)$ , and the parameters are the weights  $\Omega$  in the sinusoid SLPs. In the case of fitting  
 197 the NODE model to the interpolated data, the evidence is the interpolated per-capita growth rate  
 198 of each population, either  $\tilde{r}_R$ ,  $\tilde{r}_G$ , or  $\tilde{r}_B$ , and the parameters are the weights  $\beta_R$ ,  $\beta_G$ , and  $\beta_B$  in the  
 199 non-parametric per-capita growth rates  $r_R$ ,  $r_G$ , and  $r_B$ .

200 Assuming a normal likelihood for the residuals given the evidence we get

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

201 where  $e_i(\mathcal{D}, \theta)$  are the residuals of the model given the parameters. In the case of the interpolation,  
 202 the residuals are  $Y(t_i) - \tilde{Y}(t_i)$ , and in the case of the NODE model,  $\tilde{r}_Y(t_i) - r_Y(t_i)$ . The dispersion  
 203 term  $\sigma$  in the likelihood is measured by the parameters  $\sigma_1$  in the case of the interpolation, and  $\sigma_2$   
 204 in the case of the NODE fitting.  $I$  is the number of data points, either observations in the case of  
 205 the interpolation, or interpolated points in the case of the NODE fitting.

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

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

where  $J$  is the number of parameters in the models. The parameter  $\delta$  controls the dispersion of the priors, and thereby the complexity/level of constraint of the model. Low values of  $\delta$  will lead to underfitting, as it constrains the model to be simple, while high values of  $\delta$  will lead to overfitting, by allowing for more complex shapes.

There is no standard approach for choosing  $\delta$ . To account for overfitting, we opt for a regularisation approach by optimising the models on the second-level of inference. This means that we are finding the optimal value of  $\delta$ , in addition to optimising the model parameters. We do this by optimising the marginal posterior density of the parameters, obtained by averaging out  $\delta$  following a modification of the approach developed by Cawley and Talbot (Cawley and Talbot 2007). This yields the following expression for the marginal log posterior density of the parameters

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

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

Finally, we estimate uncertainty in parameter values through anchor sampling, which produces

221 approximate Bayesian estimates of the posterior distribution of the parameters (Pearce et al. 2018).  
222 The technique is simple in that it requires sampling a parameter vector from the prior distribu-  
223 tions, and then optimising the posterior distribution from this starting point. By repeatedly taking  
224 samples, the sampled distribution approaches the posterior distribution and provides estimates and  
225 error around the quantities that can be derived from the models. The expectation of the quantities  
226 can then be approached by computing the mean of the approximated posterior distributions. The  
227 great strength of this approach is that it is unlikely to get stuck in local maxima and provides a  
228 more robust optimisation of the posterior.

## 229 **2.5 Model analysis**

230 We analyse the shape of the per-capita growth rates to recover the interaction between the three  
231 species in the system. In particular, we look at the effect and contribution of each species to the  
232 dynamics of the other. The effect is computed as the sensitivity (i.e. the gradient) of the per-capita  
233 growth rate of a given species with respect to the density of the other species. The contribution is  
234 computed following the Geber method (Hairston et al. 2005), which comes down to multiplying  
235 the dynamics of a variable by its effects of the other variables. We further compute the importance  
236 of a species in driving the dynamics of another by computing its relative contribution compared to  
237 other species at each time step. More details on how to recover these quantities can be found in our  
238 previous study (Bonnaiffé, Sheldon, and Coulson 2021).

### 239 **3 Results**

240 We analyse sequentially the dynamics of each species, focussing on the amount of variation in  
241 per-capita growth rates explained by the NODE model, the overall direction, consistency, and im-  
242 portance of ecological interactions, and differences across replicates.

#### 243 **Drivers of top predator dynamics**

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

#### 253 **Drivers of the prey dynamics**

254 The per-capita growth rate of the algae is well explained by the NODE approximation (Fig. 3,  
255 A2, B2, C2,  $r^2 > 0.8$ , Table 1). Overall, rotifers have a negative impact on the growth of algae  
256 in all replicates (Fig. 3, A3, B3, C3, red line). We find evidence for negative density-dependence  
257 in replicate A and positive density-dependence in replicate B, but not in replicate C (Fig. 3, A3,

258 B3, C3, green line). The intermediate predator has an overall negative effect on Algae only in  
 259 replicate B (Fig. 3, B3, blue line). The main driver of algae dynamics is the rotifer population,  
 260 which accounts for 58%, 43%, and 90% of the change in algae per-capita growth rate across the  
 261 three replicates. Density dependence however plays a role in replicate A and B, with 40% and 24%  
 262 of total change in growth, respectively (Table 1). The intermediate predator contributes only to  
 263 algae growth in replicate B, accounting for 32% change in growth (Table 1). Overall, effects are  
 264 found to be consistent throughout the time series except in replicate B (Fig. 1, B3), where effects  
 265 vary in complicated ways, leading to a period in the time series where the algae is mostly driven by  
 266 the intermediate predator and positive density-dependence, and less impacted by the top predator  
 267 (Fig. 3, B5, from time 3 to 7.5).

#### 268 **Drivers of the intermediate predator dynamics**

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

## 276 4 Discussion

277 Our ability to generalise dynamical processes and patterns across populations and communities is  
278 limited by the complexity of the dynamics, differences in environments, and incomplete and/or  
279 erroneous observations. Yet it remains unclear to what extent generalisation would be possible if  
280 we overcame these limitations. We tackle this question by looking at the consistency of dynamical  
281 patterns across three replicated runs of a simple three-species community, hosted in identical  
282 environmental conditions in the lab, thus featuring limited observation error. We expected to find  
283 consistency in the drivers of population dynamics, both in time and across replicates, and thereby  
284 demonstrate that generalisation of dynamical processes may be possible if the system states were  
285 well-observed and environmental conditions were known. To verify this expectation we (1) characterised  
286 the amount of variation in per-capita growth rates that is explainable deterministically, (2)  
287 quantified the direction, strength, and importance of ecological interactions for the growth of each  
288 population, and (3) described how these varied in time and across replicates. Our results are summarised  
289 in Figure 5. We find that only the effect of algae on rotifer ( $G \rightarrow R$ ), and that of rotifer on  
290 algae ( $R \rightarrow G$ ) and flagellate ( $R \rightarrow B$ ) are conserved across the replicates. We find strong variation  
291 in the direction and importance of intra-specific density-dependence in rotifer ( $R \rightarrow R$ ) and algae  
292 ( $G \rightarrow G$ ) growth across the three replicates. The role played by the intermediate predator in the  
293 system was also different in all replicates, in that it only contributed substantially to the dynamics  
294 of the algae in replicate B ( $B \rightarrow G$ ), and was either negatively, positively, or not affected by the algae  
295 ( $G \rightarrow B$ ). Overall, this shows that the dominant interactions are conserved across replicates, but

296 that minor interactions vary substantially in importance and effect. Furthermore, we find that these  
297 dynamical processes are more consistent in time within a system, than across replicates. Our results  
298 demonstrate that because of partially generalisable dynamical processes, dynamical patterns may  
299 not be generalisable across systems, even with limited observation error and when environmental  
300 conditions and community structure are conserved.

301 What might be the drivers of differences in the dynamical processes across these three replicates?  
302 One of the main source of variation in dynamics may be differences in the intrinsic structure of  
303 populations, such as variation in traits influencing intra- and inter-specific interactions which may  
304 lead to different dynamics (Yoshida et al. 2003; Yoshida et al. 2007; De Meester et al. 2019;  
305 Bruijning, Jongejans, and Turcotte 2019). Differences in the phenotypic structure may be due to  
306 unaccounted variation in initial conditions, or variation that developed throughout time as a result  
307 of evolution (e.g. Yoshida et al. 2003; Yoshida et al. 2007). In particular, the algae in this system  
308 is prone to evolve a predator defence behaviour, by forming clumps, which reduce predation risk  
309 (Hiltunen et al. 2013). In their original paper the authors limited the initial genotypic diversity in  
310 the algae and focussed on replicates which did not display evidence of evolution, in an attempt  
311 to limit the impact of initial variation in phenotypic structure, and of evolution on the dynamics  
312 (Hiltunen et al. 2013). In spite of that evolution cannot be eliminated completely, thus variation in  
313 traits governing the interactions between the species in the system may still have developed during  
314 the experiment, and led to changes in the dynamical processes across replicates. Our study hence  
315 reinforces the idea that rapid evolution may prevent generalisation of dynamical processes (Ezard,

316 Côté, and Pelletier 2009; De Meester et al. 2019), and further suggests that this may also be the  
317 case in simple systems with limited environmental variation and opportunity for evolution.

318 Alternatively, other driving factors could be demographic stochasticity as is it often regarded as a  
319 driver of differences across systems (Dallas et al. 2021). Yet, we find that the dynamics of the three  
320 species are well-explained by relatively simple deterministic effects between the state variables,  
321 which means that that though dynamical processes differ across replicates, they are consistent  
322 in time within each system. This suggests that stochasticity plays a minor role in driving the  
323 system. Finally, we cannot exclude the potential contribution of unobserved variables that were not  
324 monitored during the experiment, such as variation in nutrient levels in the chemostat, and which  
325 may also lead to differences in the predation and intra-specific interactions across systems (e.g.  
326 Bonsall, Van Der Meijden, and Crawley 2003; Fussmann and Blasius 2005; Posey, Alphin, and  
327 Cahoon 2006).

328 Should we expect limited generalisability of dynamics across systems, even if the complexity of  
329 the process is properly captured, environmental conditions known, and the system well-observed?  
330 A similar study, that inferred dynamical processes consistency from replicated time series of a  
331 simple rotifer system, found substantial variation in vital rates across replicates (Rosenbaum et al.  
332 2019), also pointing at a low generalisability of dynamical processes. Yet the level of replication  
333 of the time series of their studies was not as stringent as that of the time series we considered,  
334 which leaves room for variability in dynamics to be caused by differences in experimental setup,  
335 population history, initial densities. Bruijning and colleagues also found substantial variation in

336 vital rates across clones in a replicated system of aphids, showing that slight phenotypic variations  
337 can change the population dynamics, all else being equal (Bruijning, Jongejans, and Turcotte 2019).  
338 This phenomenon is likely to be even more important in more complicated systems and in a natural  
339 setting where most variables are unobserved, which poses a problem for the generalisation of results  
340 across studies and systems (De Meester et al. 2019). How can we expect to generalise dynamics  
341 across real systems if we are not able to do so in artificial systems? Overall, our study reinforces  
342 the view that general inferences should not be drawn from a single system, and that more efforts  
343 are required to distinguish dynamical patterns that are conserved across systems from idiosyncratic  
344 ones.

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

355 What alternatives do we have then? We propose that NODEs are a suitable framework to study

dynamical processes, as they produce inferences that are free of model assumption and facilitate comparison across studies and systems (Bonnaiffé, Sheldon, and Coulson 2021). For instance, this study already provides a more accurate and reliable 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. 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

376 series of a three-species system, using the novel framework of NODEs. We found that only the  
377 dominant interactions were conserved across the three time series, namely that between the algae  
378 and the rotifer, while the role of the intermediate predator varied substantially. Our results hence  
379 suggest that generalisation may not seem possible, even in simple system with no environmental  
380 variation. Given previous work in this system, the main cause of differences across replicates may  
381 be evolution in prey defence traits. We conclude that more work is required, using NODEs, to  
382 identify dynamical patterns that are conserved and those that are idiosyncratic across a wider range  
383 of systems.

#### 384 **Acknowledgments**

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

#### 389 **Data accessibility**

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

#### 391 **Statement of authorship**

392 Willem Bonnaff  designed the method, performed the analysis, wrote the manuscript; Tim Coulson  
393 led investigations, provided input for the manuscript, commented on the manuscript.

## References

- Adamson, M. W. and A. Y. Morozov (2013). “When can we trust our model predictions? Unearthing structural sensitivity in biological systems”. In: *Proc. R. Soc. A Math. Phys. Eng. Sci.* 469.2149, pp. 1–19.
- Bonnaiffé, Willem, Ben C. Sheldon, and Tim Coulson (2021). “Neural ordinary differential equations for ecological and evolutionary time series analysis”. In: *Methods Ecol. Evol.* 2, pp. 1–46.
- Bonnaiffé, Willem, Stéphane Legendre, Alain Danet, and Eric Edeline (2021). “Comparison of size-structured and species-level trophic networks reveals antagonistic effects of temperature on vertical trophic diversity at the population and species level”. In: *Oikos*, pp. 1–14.
- Bonsall, M B, E. Van Der Meijden, and M J Crawley (2003). “Contrasting dynamics in the same plant-herbivore interaction”. In: *Proc. Natl. Acad. Sci. U. S. A.* 100.25, pp. 14932–14936.
- Bruijning, Marjolein, Eelke Jongejans, and Martin M. Turcotte (2019). “Demographic responses underlying eco-evolutionary dynamics as revealed with inverse modelling”. In: *J. Anim. Ecol.* 88.5, pp. 768–779.
- Brunner, Franziska S., Jacques A. Deere, Martijn Egas, Christophe Eizaguirre, and Joost A.M. Raeymaekers (2019). “The diversity of eco-evolutionary dynamics: Comparing the feedbacks between ecology and evolution across scales”. In: *Funct. Ecol.* 33.1, pp. 7–12.
- Cawley, Gavin C. and Nicola L.C. Talbot (2007). “Preventing over-fitting during model selection via bayesian regularisation of the hyper-parameters”. In: *J. Mach. Learn. Res.* 8, pp. 841–861.

414 Dallas, Tad, Brett A. Melbourne, Geoffrey Legault, and Alan Hastings (2021). “Initial abundance  
415 and stochasticity influence competitive outcome in communities”. In: *J. Anim. Ecol.* Pp. 1–26.

416 De Meester, Luc et al. (2019). “Analysing eco-evolutionary dynamicsThe challenging complexity  
417 of the real world”. In: *Funct. Ecol.* 33.1, pp. 43–59.

418 Demyanov, V., S. N. Wood, and T. J. Kedwards (2006). “Improving ecological impact assessment  
419 by statistical data synthesis using process-based models”. In: *J. R. Stat. Soc. Ser. C Appl. Stat.*  
420 55.1, pp. 41–62.

421 Ellner, Stephen P., Yodit Seifu, and Robert H. Smith (2002). “Fitting Population Dynamic Models  
422 to Time-Series Data by Gradient Matching”. In: *Ecology* 83.8, p. 2256.

423 Ezard, Thomas H.G., Steeve D. Côté, and Fanie Pelletier (2009). “Eco-evolutionary dynamics:  
424 Disentangling phenotypic, environmental and population fluctuations”. In: *Philos. Trans. R. Soc.*  
425 *B Biol. Sci.* 364.1523, pp. 1491–1498.

426 Funahashi, Ken ichi and Yuichi Nakamura (1993). “Approximation of dynamical systems by con-  
427 tinuous time recurrent neural networks”. In: *Neural Networks* 6.6, pp. 801–806.

428 Fussmann, Gregor F. and Bernd Blasius (2005). “Community response to enrichment is highly  
429 sensitive to model structure”. In: *Biol. Lett.* 1.1, pp. 9–12.

430 Gamelon, Marlène et al. (2019). “Accounting for interspecific competition and age structure in  
431 demographic analyses of density dependence improves predictions of fluctuations in population  
432 size”. In: *Ecol. Lett.* 22.5, pp. 797–806.

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

435 Hairston, Nelson G., Stephen P. Ellner, Monica A. Geber, Takehito Yoshida, and Jennifer A. Fox  
 436 (2005). “Rapid evolution and the convergence of ecological and evolutionary time”. In: *Ecol.*  
 437 *Lett.* 8.10, pp. 1114–1127.

438 Hiltunen, Teppo, L. E. Jones, S. P. Ellner, and Jr G. Hairston Nelson G. (2013). “Temporal dynam-  
 439 ics of a simple community with intraguild predation: an experimental test”. In: *Ecology* 94.4,  
 440 pp. 773–779.

441 Jost, C. and S. P. Ellner (2000). “Testing for predator dependence in predator-prey dynamics: A  
 442 non-parametric approach”. In: *Proc. R. Soc. B Biol. Sci.* 267.1453, pp. 1611–1620.

443 Kendall, Bruce E. et al. (2005). “Population cycles in the pine looper moth: Dynamical tests of  
 444 mechanistic hypotheses”. In: *Ecol. Monogr.* 75.2, pp. 259–276.

445 Lawton, John H (1999). “Nordic Society Oikos Are There General Laws in Ecology ?” In: *Oikos*  
 446 84.2, pp. 177–192.

447 Pearce, Tim, Felix Leibfried, Alexandra Brintrup, Mohamed Zaki, and Andy Neely (2018). “Un-  
 448 certainty in Neural Networks: Approximately Bayesian Ensembling”. In: *arXiv*, pp. 1–10.

449 Pontarp, Mikael, Åke Brännström, and Owen L. Petchey (2019). “Inferring community assembly  
 450 processes from macroscopic patterns using dynamic eco-evolutionary models and Approximate  
 451 Bayesian Computation (ABC)”. In: *Methods Ecol. Evol.* 10.4, pp. 450–460.

452 Posey, Martin H., Troy D. Alphin, and Lawrence Cahoon (2006). “Benthic community responses  
 453 to nutrient enrichment and predator exclusion: Influence of background nutrient concentrations  
 454 and interactive effects”. In: *J. Exp. Mar. Bio. Ecol.* 330.1, pp. 105–118.

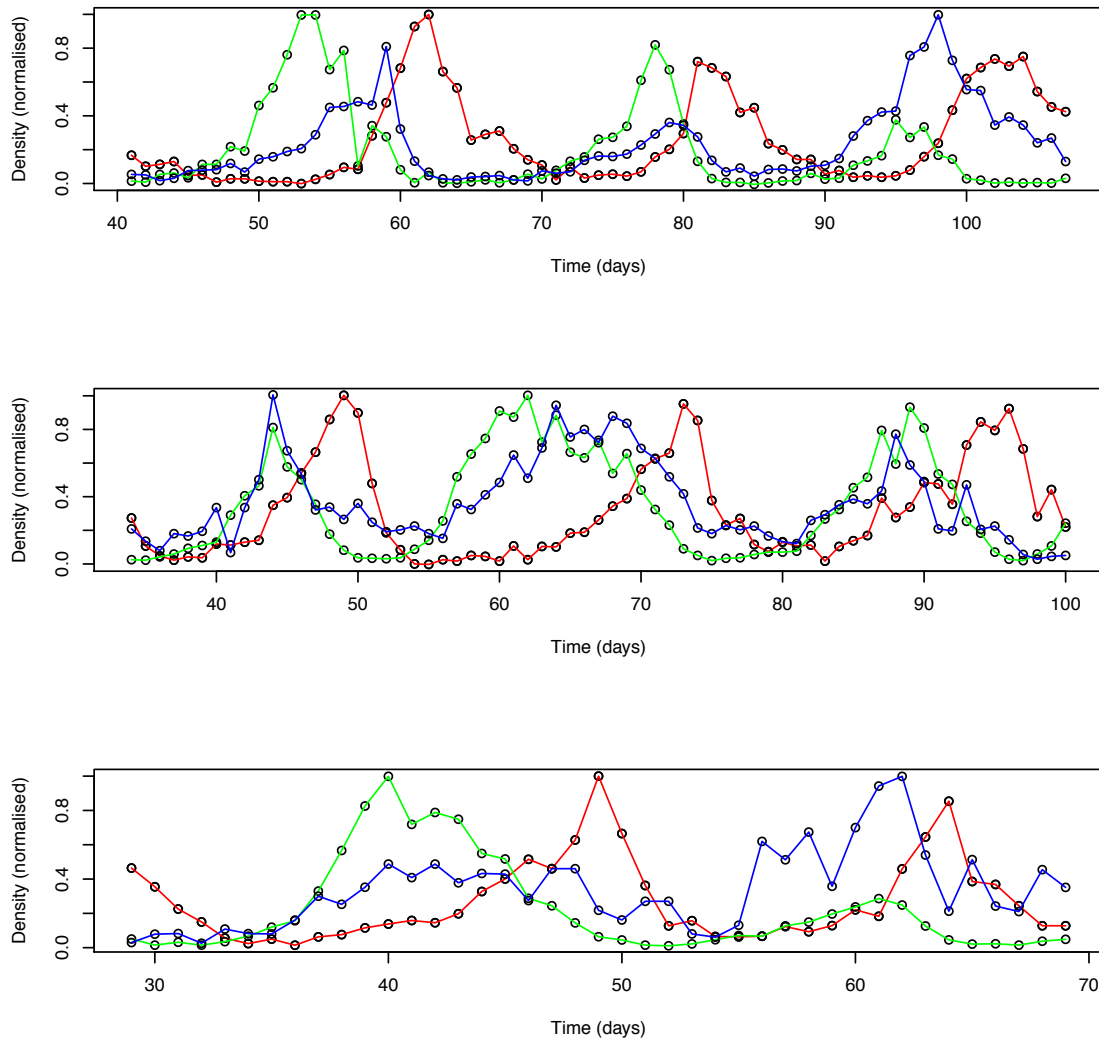
455 Rosenbaum, Benjamin, Michael Raatz, Guntram Weithoff, Gregor F. Fussmann, and Ursula Gaedke  
 456 (2019). “Estimating parameters from multiple time series of population dynamics using bayesian  
 457 inference”. In: *Front. Ecol. Evol.* 6.234, pp. 1–14.

458 Shurin, Jonathan B., Jessica L. Clasen, Hamish S. Greig, Pavel Kratina, and Patrick L. Thomp-  
 459 son (2012). “Warming shifts top-down and bottom-up control of pond food web structure and  
 460 function.” In: *Philos. Trans. R. Soc. Lond. B. Biol. Sci.* 367.1605, pp. 3008–17.

461 Wu, Jun, Makoto Fukuhara, and Tatsuoki Takeda (2005). “Parameter estimation of an ecological  
 462 system by a neural network with residual minimization training”. In: *Ecol. Modell.* 189.3-4,  
 463 pp. 289–304.

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

466 Yoshida, Takehito, Stephen P. Ellner, Laura E. Jones, Brendan J.M. Bohannan, Richard E. Lenski,  
 467 and Nelson G. Hairston (2007). “Cryptic population dynamics: Rapid evolution masks trophic  
 468 interactions”. In: *PLoS Biol.* 5.9, pp. 1868–1879.



**Figure 1: Time series of the prey algae (green), flagellate intermediate predator (blue), and top predator (red).** Each time series correspond to one out of three replicates. Black open dots correspond to observed densities, scaled between 0 and 1. Time series were obtained from the study by Hiltunen and colleagues (Hiltunen et al. 2013).

Figure 2: Drivers of rotifer dynamics. Row (1) presents the log density of rotifers in time, the dots are the observed densities while the red line is the interpolated data. Row (2) presents the change in log density at each time step, where the black line is the per-capita growth rate derived from the interpolated data and the red line is the NODE approximation. Row (3) presents the effect of each species on the dynamics/growth of rotifers, obtained by computing the sensitivity of the NODE approximation of the rotifer per-capita growth rate with respect to each species density. Row (4) presents the contribution of each species to the growth of the rotifer, obtained by multiplying its effect by its net change (i.e. row 2 and 3). Row (5) is the same as fourth but expressed as % of total change at each time step explained by each species. For all figures, green, blue, and red refer respectively to algae, agellate, and rotifer.

**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      |
|---------------------------------|-------|--------|---------|--------|
| <b>Replicate A</b>              |       |        |         |        |
|                                 | $r^2$ | 0.968  | 0.839   | 0.831  |
| <b>Mean effects</b>             | on R  | 1.180  | 1.723   | 0.632  |
|                                 | on G  | -1.888 | -1.286  | -0.208 |
|                                 | on B  | -1.476 | -0.371  | 0.131  |
| <b>% of total contributions</b> | to R  | 0.356  | 0.552   | 0.091  |
|                                 | to G  | 0.583  | 0.399   | 0.016  |
|                                 | to B  | 0.785  | 0.194   | 0.020  |
| <b>Replicate B</b>              |       |        |         |        |
|                                 | $r^2$ | 0.744  | 0.997   | 0.765  |
| <b>Mean effects</b>             | on R  | 0.058  | 1.146   | 0.113  |
|                                 | on G  | -1.505 | 0.707   | -1.287 |
|                                 | on B  | -1.000 | 0.505   | -0.046 |
| <b>% of total contributions</b> | to R  | 0.035  | 0.934   | 0.029  |
|                                 | to G  | 0.439  | 0.242   | 0.317  |
|                                 | to B  | 0.624  | 0.373   | 0.001  |
| <b>Replicate C</b>              |       |        |         |        |
|                                 | $r^2$ | 0.923  | 0.962   | 0.726  |
| <b>Mean effects</b>             | on R  | -0.106 | 0.7814  | 0.498  |
|                                 | on G  | -1.234 | -0.1840 | 0.146  |
|                                 | on B  | -0.659 | 0.0912  | 0.014  |
| <b>% of total contributions</b> | to R  | 0.080  | 0.743   | 0.175  |
|                                 | to G  | 0.900  | 0.068   | 0.030  |
|                                 | to B  | 0.913  | 0.084   | 0.001  |

## 469 5 Supplementary

### 470 A Bayesian regularisation

471 In this section we describe how to derive the modified model selection criteria developed by Caw-  
472 ley and Talbot (Cawley and Talbot 2007). Bayesian regularisation simply amounts to constraining  
473 the values of the parameters in the model to be close to a desired value. Usually, parameters are  
474 constrained by choosing normal priors centered about 0. In this case, the standard deviation of the  
475 normal priors governs the range of values that the parameters can take, and hence constrains more  
476 or less strongly the behaviour of the model (Cawley and Talbot 2007). Performing inference on the  
477 second level means that we are trying to find the appropriate value of the dispersion of the priors,  
478 in other words, the appropriate level of constraint on the model. In practice, choosing the level of  
479 constraint is difficult, Cawley and Talbot hence developed a criterion to perform model selection  
480 on the second level of inference. They proposed to optimise the marginal posterior distribution by  
481 averaging out the dispersion of the priors. With an appropriate choice of prior, the dispersion can  
482 be integrated out, leaving us with a formula for the posterior that only depends on the parameters  
483 of the model,

$$\log P(\theta|\mathcal{D}) \propto \frac{I}{2} \log \left( \sum_{i=1}^I e_i(\mathcal{D}, \theta)^2 \right) + \frac{J}{2} \log \left( \sum_{j=1}^J \theta_j^2 \right) \quad (12)$$

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

485 number of data points and parameters, respectively,  $e_i$  denote the residuals of the model, and  $\theta$   
 486 denote the parameters of the model. The construction is elegant because it is not sensitive to  
 487 the choice of prior hyperparameters, and simple as it amounts to optimising the log of the sum of  
 488 squares, rather than the sum of squares (in the case of normal ordinary least square).  
 489 The issue with this formula is that the marginal posterior density is infinity when the parameters  
 490 are 0, which leads to underfitting. In this paper we use a modified criterion, which corrects for that  
 491 problem

$$\log P(\theta|\mathcal{D}) \propto \frac{I}{2} \log \left( 1 + \sum_{i=1}^I e_i(\mathcal{D}, \theta)^2 \right) + \frac{J}{2} \log \left( 1 + \sum_{j=1}^J \theta_j^2 \right) \quad (13)$$

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