Received: 18 August 2022 | Accepted: 14 April 2023

DOI: 10.1111/2041-210X.14121

Methods in Ecology and Evolution

RESEARCH ARTICLE

Fast fitting of neural ordinary differential equations by Bayesian neural gradient matching to infer ecological interactions from time-series data

Willem Bonnaffé[1,2] | Tim Coulson[2]

[1]Big Data Institute, University of Oxford, Oxford, UK

[2]Department of Biology, University of Oxford, Zoology Research and Administration Building, Oxford, UK

Correspondence
Willem Bonnaffé
Email: willem.bonnaffe@nds.ox.ac.uk

Handling Editor: Sydne Record

Abstract

1. Inferring ecological interactions is hard because we often lack suitable parametric representations to portray them. Neural ordinary differential equations (NODEs) provide a way of estimating interactions non-parametrically from time-series data. NODEs, however, are slow to fit, and inferred interactions usually are not compared with the ground truth.
2. We provide a fast NODE fitting method, Bayesian neural gradient matching (BNGM), which relies on interpolating time series with neural networks and fitting NODEs to the interpolated dynamics with Bayesian regularisation. We test the accuracy of the approach by inferring ecological interactions in time series generated by an ODE model with known interactions. We compare these results against three existing approaches for estimating ecological interactions, standard NODEs, ODE models and convergent cross-mapping (CCM). We also infer interactions in experimentally replicated time series of a microcosm featuring an algae, flagellate and rotifer population, in the hare and lynx system, and the Maizuru Bay community featuring 11 species.
3. Our BNGM approach allows us to reduce the fitting time of NODE systems to only a few seconds and provides accurate estimates of ecological interactions in the artificial system, as true ecological interactions are recovered with minimal error. Our benchmark analysis reveals that our approach is both faster and more accurate than standard NODEs and parametric ODEs, while CCM was found to be faster but less accurate. The analysis of the replicated time series reveals that only the strongest interactions are consistent across replicates, while the analysis of the Maizuru community shows the strong negative impact of the chameleon goby on most species of the community, and a potential indirect negative effect of temperature by favouring goby population growth.
4. Overall, NODEs alleviate the need for a mechanistic understanding of interactions, and BNGM alleviates the heavy computational cost. This is a crucial step availing quick NODE fitting to larger systems, cross-validation and uncertainty

quantification, as well as more objective estimation of interactions, and complex context dependence, than parametric models.

KEYWORDS
artificial neural networks, ecological dynamics, ecological interactions, Geber method, gradient matching, microcosm, neural ordinary differential equations, time-series analysis

1 | INTRODUCTION

The concepts of population and community (i.e. groups of populations) are central in ecology (Berryman, 2002). Ecologists have had a longstanding interest in finding laws that govern population and community dynamics, namely changes in the number of individuals in the populations present in a community (Lawton, 1999; Turchin, 1999). Population dynamics can be characterised by a logistic growth, or similar functional forms, limited by ecological interactions with other organisms, and by the state of the environment (Berryman, 2003; Turchin, 2001). Intra-specific interactions correspond to interactions between individuals of within and across different sex, age or size classes, belonging to the same species (Turchin, 2001). Inter-specific interactions are interactions between individuals from different species, be it competitors, prey, predators or pathogens (Berryman, 2003; Turchin, 2001). These interactions can cause populations to have lagged effects impacting their own growth, often called feedback effects, mediated by their impact on the other populations they interact with (Berryman & Turchin, 1997).

Characterising these interactions has been a longtime challenge. Ecologists started analysing time-series data with parametric models (Gross et al., 2005; Ives et al., 2003; Kendall et al., 1999; Royama, 1984), as time series of population counts are the most commonly collected long-term data in biology (Kendall et al., 1999). Initial analysis involved fitting simple auto-regressive linear models to time series of a single species, leading to contentious interpretations of interactions thereby inferred (e.g. Berryman & Turchin, 1997). For instance, Royama et al. interpreted higher-order lags as evidence of species interactions (Royama, 1984), while Lande et al. interpreted them as age-structure signatures (Lande et al., 2002). Coulson et al. showed they can even be caused by interactions between the sexes (Mysterud et al., 2002). Jonzen et al. added doubt over interpreting lags by demonstrating that autocorrelation in environmental noise could prevent altogether the reliable estimation of lag effects in single species time-series data (Jonzén et al., 2002). More recent work has investigated time series of multiple species, environmental factors, and has mechanistically modelled various ecological interactions (e.g. Adams et al., 2020; Bruijning et al., 2019; Rosenbaum et al., 2019). In these models, ecological interactions are quantified explicitly by specific parameters, rather than phenomenologically with lags. This allowed for a more thorough quantification of interactions and comparison of alternative ecological interaction network architectures.

However, ecologists still face two main obstacles when estimating ecological interactions from time-series data. The first is that interactions are highly context-dependent, so they change in time with the state of the ecosystem and of the environment (Song et al., 2020). Ecological interactions were traditionally considered linear or fixed, yet there is substantial evidence that this is not the case in nature (e.g. Bonnaffé, Legendre, et al., 2021; Bonsall et al., 2003; Bruijning et al., 2019; Gross et al., 2005; Kendall et al., 2005; Rosenbaum et al., 2019; Ushio et al., 2018). The effect of the population on itself depends on the density of individuals (e.g. Brook & Bradshaw, 2006; Lingjaerde et al., 2001; Moe et al., 2005); while predation rates can depend on the density of the predator (Jost & Ellner, 2000; Yoshida et al., 2003). Many vital rates underpinning ecological interactions are age- and size-dependent (Bonnaffé et al., 2018; Bonnaffé, Legendre, et al., 2021) and governed by environmental variables, such as temperature (Brown et al., 2004). Interactions also change following evolution of the traits that underpin them (Turchin et al., 2003; Yoshida et al., 2003). This makes it virtually impossible to model the full complexity of ecological interactions (Kendall et al., 1999; Lawton, 1999).

This leads to the second obstacle, known as structural sensitivity, namely sensitivity of the results to the structure of the model (Adamson & Morozov, 2013; Wood, 2001). Because of the complexity of the interactions, we often lack suitable mathematical representations to portray them (Ellner et al., 2002; Jost & Ellner, 2000; Wood, 2001; Wu et al., 2005). Parametric representations of the interactions are assumed a priori, which means that any interaction quantified is ultimately contingent on this arbitrary choice, and hence potentially biased (Ellner et al., 2002; Jost & Ellner, 2000; Wood, 2001; Wu et al., 2005). Parametric inference of ecological interactions from time-series data, therefore, only provides qualitative evidence, requiring further experimental verification and quantification (Kendall et al., 1999).

Non-parametric modelling provides a powerful alternative that can help solve these problems (e.g. Ellner et al., 2002; Jost & Ellner, 2000; Pasquali & Soresina, 2018; Wood, 2001; Wu et al., 2005). Non-parametric forms give more freedom to researchers wishing to model population dynamics, and allow a test of whether the linear or linearised assumption of standard models is warranted. Interactions are quantified as the sensitivity of the non-parametric approximation of the dynamics with respect to other state variables (Sugihara et al., 2012; Ushio et al., 2018). Non-parametric models require minimal assumptions regarding the mathematical nature of ecological interactions (Gross et al., 2005; Jost & Ellner, 2000), and

hence, provide interaction estimates that are more robust to model structure (Wood, 2001). In particular, artificial neural networks (ANNs) offer a promising, yet underused, non-parametric alternative to linear functional forms. In previous work, we introduced a powerful framework, relying on neural ordinary differential equations (NODEs, Chen et al., 2019) to approximate the dynamics of populations non-parametrically, from which we derive ecological interactions (Bonnaffé, Sheldon, et al., 2021). More specifically, the ANNs embedded in the ODEs learn non-parametrically the shape of the per-capita growth rate of the populations and its dependence on the state variables of the system (Bonnaffé, Sheldon, et al., 2021). Combined with the Geber method (Hairston et al., 2005), we are able to estimate the direction, strength and degree of non-linearity of interactions.

One limitation of the approach lies in the computational cost of fitting the NODEs (Bonnaffé, Sheldon, et al., 2021; Chen et al., 2019). This is due to the fact that NODEs, as with ODEs, need to be simulated over the entire range of the time series to compute the likelihood of the trajectories of the model. This can be avoided by using gradient matching, which requires interpolating the time series, and fitting the ODEs directly to the interpolated dynamics (Aarts & Veer, 2001; Ellner et al., 2002; Jost & Ellner, 2000). Although a similar approach has been proposed (see Treven et al., 2021), there are no implementations of it to fitting NODEs, in spite of its great potential for reducing computational costs. In addition, given the novelty of the framework, the accuracy and robustness of NODEs in estimating ecological interactions remain largely unexplored. Most of the work to date is concerned with the accuracy of the fitted trajectories and of the forecasts (Frank, 2022; Mai et al., 2016; Treven et al., 2021), while little attention has been given to the functional form of the processes that are producing the dynamics approximated by NODEs (but see Hu et al., 2020 for a step in this direction). It is important to understand to what extent the neural networks embedded within NODEs carry meaningful biological information (Novak & Stouffer, 2021).

In this manuscript, we first introduce a novel fitting technique for NODEs, Bayesian neural gradient matching (BNGM). The method extends gradient matching by using neural networks to interpolate the time-series data instead of splines (Ellner et al., 2002), and Bayesian regularisation to fit NODEs to the interpolated dynamics (Cawley & Talbot, 2007). This reduces the fitting time of NODEs to only a few seconds, compared with about 30 min in our previous work (Bonnaffé, Sheldon, et al., 2021), allowing for efficient cross-validation and uncertainty quantification. We then demonstrate that NODEs are highly accurate in recovering ecological interactions in an artificial three-species prey–predator system where truth is known. Finally, we conclude by characterising ecological interactions in three replicates of an experimental three-species prey–predator system with an algae, flagellate and rotifer (Hiltunen et al., 2013), in the classic hare and lynx time series (Odum & Barrett, 1972), as well as in the larger aquatic community of the Maizuru Bay in Japan (Ushio et al., 2018). We find that only main interactions, between the algae and the rotifer, are conserved across the three replicates and not the interactions of the flagellate with the other species. We also find that in most cases linear interactions are sufficient to explain the dynamics apart from non-linearity in the effect of the prey on the top predator in both the rotifer and lynx. Finally, we find that the dynamics of the aquatic species of the Maizuru Bay community are largely determined by a single species, the chameleon goby, and by an indirect effect of temperature via its impact on goby population dynamics.

2 | MATERIALS AND METHODS

2.1 | Method overview

We provide a non-parametric method for estimating ecological interactions from time-series data of species density. We do this by approximating the dynamics of each species with NODEs (Bonnaffé, Sheldon, et al., 2021). We then compute ecological interactions as the sensitivity of these dynamics to a change in the respective species densities (Bonnaffé, Sheldon, et al., 2021; Sugihara et al., 2012). We provide a novel method, BNGM, allowing us to fit NODE systems in a only a few seconds.

2.2 | Neural ordinary differential equation

A NODE is a type of ordinary differential equation (ODE) that is partly or entirely defined as an ANN (Chen et al., 2019). They are useful to infer dynamical processes non-parametrically from time-series data (Bonnaffé, Sheldon, et al., 2021). We choose NODEs over standard 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 (Adamson & Morozov, 2013; Jost & Ellner, 2000). 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.

We first consider a general NODE system,

$$\frac{dy_i}{dt} = f_p(y, \theta_i), \tag{1}$$

where dy_i / dt denotes the temporal change in the ith variable of the system, y_i, as a function of the other state variables $y = \{y_1, y_2, \ldots, y_I\}$. The function f_p is a non-parametric function of the state variables and its shape is controlled by the parameter vector θ_i. In the context of NODEs, f_p is an ANN. The most common class of ANN used in NODEs are single-layer fully connected feedforward ANNs (e.g. Wu et al., 2005), also referred to as single-layer perceptrons (SLPs, e.g. Bonnaffé, Sheldon, et al., 2021),

$$f_p(y, \theta_i) = f_\lambda\left(\theta_i^{(0)} + \sum_{j=1}^{J} \theta_{ij}^{(1)} f_\sigma\left(\theta_{ij}^{(2)} + \sum_{k=1}^{I} \theta_{ijk}^{(3)} y_k\right)\right), \tag{2}$$

which feature a single layer, containing J neurons, that maps the inputs, here the state variables y, to a single output, the dynamics of state variable i, dy_i/dt. The parameter vector θ_i contains the weights $\theta^{(l)}$ of the connections in the SLPs. SLPs can be viewed as weighted sums of activation functions f_σ, which are usually chosen to be sigmoid functions $f(x) = 1/(1+\exp(-x))$. The link function f_λ allows us to map the output of the network to a specific domain, for instance applying tanh will constrain the dynamics between −1 and 1, $dy_i/dt \in]-1,1[$. Multi-layer networks can also be used but are generally considered unnecessary since pioneering work established that a single layer is sufficient to approximate any continuous function to a desired level of error (Funahashi & Nakamura, 1993).

This general form can be changed to represent biological constraints on the state variables. In particular, for population dynamics, the state variables are strictly positive population densities, $y_i = N_i \in \mathscr{R}^+$. We could, hence, re-write Equation (1) as, $dN_i/dt = f_p(N, \theta_i)N_i$, where the SLPs approximate the per-capita growth rate of the populations. More details regarding these models can be found in our previous work (Bonnaffé, Sheldon, et al., 2021).

2.3 | Fitting neural ordinary differential equations by Bayesian neural gradient matching

In this section, we describe how to estimate the parameters θ of the NODE system given a set of time series. Fitting NODEs can be highly computationally intensive, which hinders uncertainty quantification, cross-validation, and model selection (Bonnaffé, Sheldon, et al., 2021). We solve this issue by introducing BNGM, a computationally efficient approach to fit NODEs. The approach involves two steps (Figure 1). First, we interpolate the state variables and their dynamics with neural networks (Figure 1, red boxes). Second, we train each NODE to satisfy the interpolated state and dynamics (Figure 1, blue boxes). This bypasses the costly numerical integration of the NODE system and provides a fully mathematically tractable expression for the posterior distribution of the parameter vector θ, and hence, analytical expressions for the gradients. We coin the term BNGM to emphasise two important refinements of the standard gradient matching algorithm (Ellner et al., 2002). The first is that we use neural networks as interpolation functions, and the second is that we use Bayesian regularisation to limit overfitting and estimate uncertainty around parameters (Cawley & Talbot, 2007).

2.3.1 | Interpolating the time series

The first step is to interpolate the time series and differentiate it with respect to time to approximate the state and dynamics of the variables. We perform the interpolation via non-parametric regression of the interpolating functions on the time-series data,

$$Y_{it} = \tilde{y}_i(t, \omega_i) + \varepsilon_{it}^{(o)}, \tag{3}$$

where Y_{it} is the observed value of the state variable i at time t, $\tilde{y}_i(t,\omega_i)$ is the value predicted by the interpolation function given the parameter vector ω_i, and $\varepsilon_{it}^{(o)}$ is the observation error between the observation and prediction. The interpolation function is chosen to be a neural network,

$$\tilde{y}_i(t,\omega_i) = f_\lambda\left(\omega_i^{(0)} + \sum_{j=1}^{J}\omega_{ij}^{(1)} f_\sigma\left(\omega_{ij}^{(2)} + \omega_{ij}^{(3)} t\right)\right), \tag{4}$$

where the parameter vector ω_i contains the weights $\omega^{(l)}$ of the network. We can further differentiate this expression with respect to time to obtain an interpolation of the dynamics of the state variables (Figure 1, red boxes),

$$\frac{\partial \tilde{y}_i}{\partial t}(t,\omega_i) = \sum_{j=1}^{J}\omega_{ij}^{(1)}\omega_{ij}^{(3)}\frac{\partial f_\sigma}{\partial t}\left(\omega_{ij}^{(2)} + \omega_{ij}^{(3)} t\right)\frac{\partial f_\lambda}{\partial t}\left(\omega_i^{(0)} + \sum_{k=1}^{J}\omega_{ik}^{(1)} f_\sigma\left(\omega_{ik}^{(2)} + \omega_{ik}^{(3)} t\right)\right). \tag{5}$$

2.3.2 | Fitting neural ordinary differential equations to the interpolated time series

The second step is to train the NODE system (Equation 1) to satisfy the interpolated dynamics. Thanks to the interpolation step, this simply amounts to performing a non-parametric regression of each NODE (Equation 1) on the interpolated dynamics (Equation 5),

$$\frac{\partial \tilde{y}_i}{\partial t}(t,\omega_i) = \frac{dy_i}{dt}(\tilde{y},\theta_i) + \varepsilon_{it}^{(p)}, \tag{6}$$

where $\varepsilon_{it}^{(p)}$ is the process error, namely the difference between the interpolated dynamics, $\partial\tilde{y}_i/\partial t$ and the NODE, dy_i/dt, given the interpolated state variables $\tilde{y} = \{\tilde{y}_1, \tilde{y}_2, \ldots, \tilde{y}_I\}$ (Figure 1, blue boxes).

2.3.3 | Bayesian regularisation

In the context of standard gradient matching, defining the observation model (Equation 3) and process model (Equation 6) would be sufficient to fit the NODE system (Equation 1) to the time series via optimisation (Ellner et al., 2002; Jost & Ellner, 2000; Wu et al., 2005). We could find the parameter vector ω_i and θ_i that minimise the sum of squared observation and process errors, $\varepsilon_{it}^{(o)}$ and $\varepsilon_{it}^{(p)}$ (Equations 3 and 6). However, this approach is prone to overfitting, and does not provide estimates of uncertainty around model predictions. To account for this, we introduce Bayesian regularisation, which allows us to control for overfitting by constraining parameters with prior distributions (Cawley & Talbot, 2007), and to root our interpretation of uncertainty in a Bayesian framework.

First, we define a simple Bayesian model to fit the interpolation functions (Equation 3) to the time-series data. We assume normal distributions for the observation error, $\varepsilon_{ij}^{(o)} \sim \mathcal{N}(0, \sigma_i)$, and for the parameters, $\omega_{ij} \sim \mathcal{N}(0, \gamma_{ij})$. Here, we are only interested in interpolating the time series accurately, irrespective of the value of σ_i and γ_{ij}. Therefore, we use the approach developed by Cawley

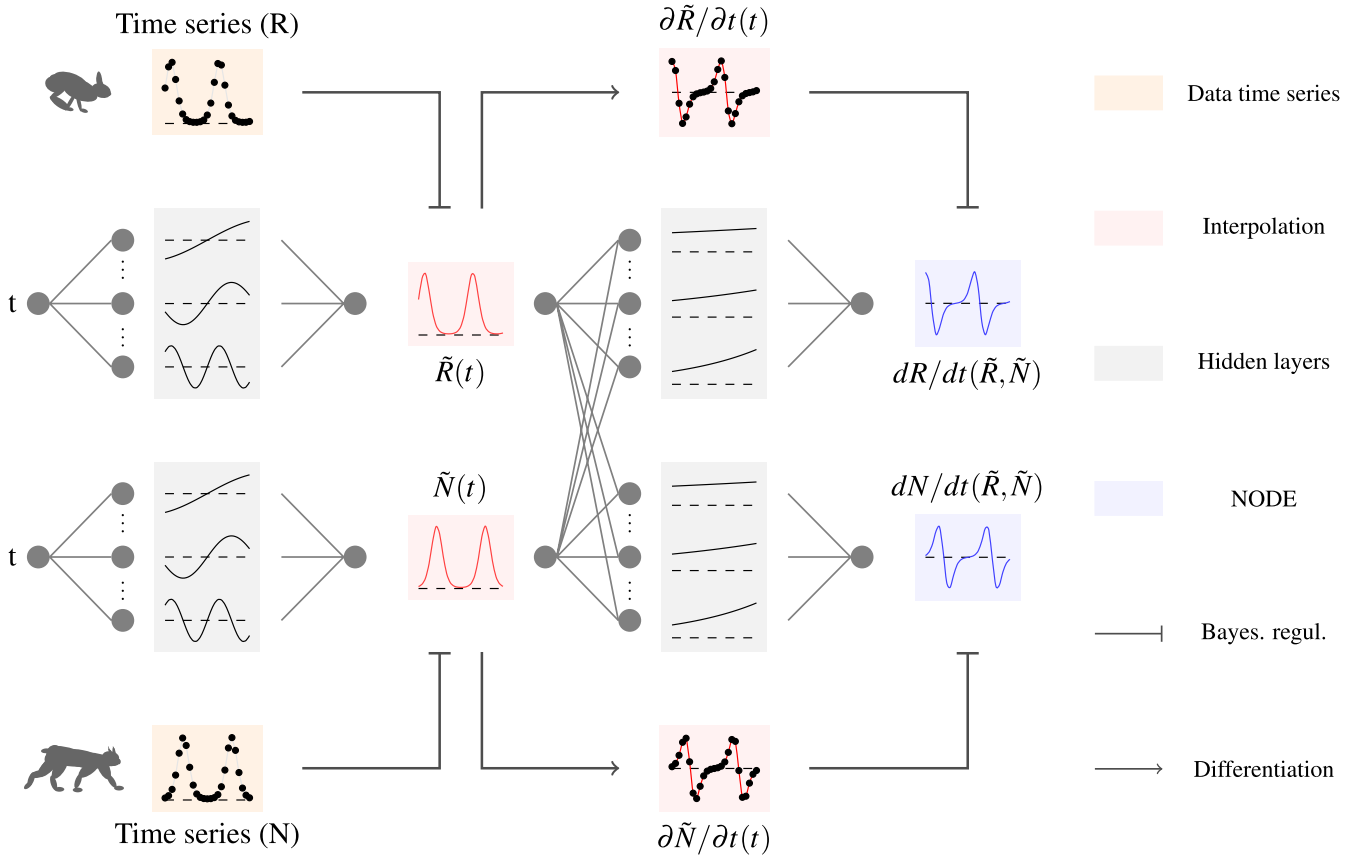

FIGURE 1 Overview of fitting neural ordinary differential equations (NODE) by Bayesian neural gradient matching (BNGM). In a first step we compute a continuous time approximation (interpolation) of each state variable, here the prey $\tilde{R}(t)$ and predator density $\tilde{N}(t)$ (red boxes). To do that we fit an ANN, that takes time as input, to each time series, via Bayesian regularisation. Interpolated dynamics of populations can then be computed by taking the derivative of the ANN with respect to time, $\partial\tilde{R}/\partial t$ and $\partial\tilde{N}/\partial t$. In a second step, we fit each NODE, dR/dt and dN/dt (blue boxes), to the interpolated dynamics. To do that we fit an ANN, which takes as input the interpolated variables $\tilde{R}(t)$ and $\tilde{N}(t)$, to the interpolated dynamics $\partial\tilde{R}/\partial t$ and $\partial\tilde{N}/\partial t$, via Bayesian regularisation.

and Talbot to average out the value of the parameters σ_i and γ_{ij} in the full posterior distribution (Cawley & Talbot, 2007), assuming hyperpriors $p(\xi) \propto \frac{1}{\xi}\exp\{-\xi/2\}$ for both parameters. This yields the following expression for the log marginal posterior density of the parameters,

$$\log P(\omega_i \mid Y_i) \propto -\frac{N^{(o)}}{2}\log\left(1+\sum_{t=1}^{N^{(o)}}\left(\varepsilon_{it}^{(o)}\right)^2\right)-\frac{M^{(o)}}{2}\log\left(1+\sum_{j=1}^{M^{(o)}}\omega_{ij}^2\right) \quad (7)$$

where P is the marginal posterior density, $\omega_i = \{\omega_{i1}, \omega_{i2}, \ldots, \omega_{iM^{(o)}}\}$ is the observation parameter vector controlling the interpolation function, $Y_i = \{Y_{i1}, Y_{i2}, \ldots, Y_{iN^{(o)}}\}$ corresponds to the sequence of observations of state variable i at time step t, $N^{(o)}$ is the total number of time steps in the time series, $\varepsilon_{it}^{(o)}$ is the observation error at time step t between the interpolated and observed value of variable i, $M^{(o)}$ is the total number of parameters. More details on how to derive this expression can be found in Appendix A.

Then, we define a simple Bayesian model to fit the NODEs to the interpolated dynamics, given the interpolated states. We assume normal distributions for the observation error, $\varepsilon_{it}^{(p)} \sim \mathcal{N}(0, \sigma_i)$, and parameters, $\theta_{ij} \sim \mathcal{N}(0, \delta_{ij})$. This gives the following expression for the log posterior density of the parameters given the interpolations,

$$\log p(\theta_i \mid \omega) \propto -\frac{1}{2}\sum_{t=1}^{N^{(p)}}\left(\frac{\varepsilon_{it}^{(p)}}{\sigma_i}\right)^2-\frac{1}{2}\sum_{j=1}^{M^{(p)}}\left(\frac{\theta_{ij}}{\delta_{ij}}\right)^2 \quad (8)$$

where $\theta_i = \{\theta_{i1}, \theta_{i2}, \ldots, \theta_{iM^{(p)}}\}$ are the NODE parameters of the ith variable, $\omega = \{\omega_1, \omega_2, \ldots, \omega_I\}$ are the interpolation parameters of each state variable, $\varepsilon_{it}^{(p)}$ is the process error of variable i at time step t between the interpolated dynamics and NODE prediction, σ_i is the standard deviation of the likelihood, $N^{(p)}$ is the total number of time steps, $M^{(p)}$ is the total number of parameters, δ_{ij} is the standard deviation of the prior distribution of parameter θ_{ij}.

This approach allows us to limit overfitting by adjusting the constraint on the parameters, which is controlled by the standard deviation of the parameter prior distributions, δ_{ij} (Bonnaffé, Sheldon, et al., 2021; Cawley & Talbot, 2007). We could set small values of δ to limit the degree of non-linearity in the response, or to eliminate specific variables from the model by constraining their parameters to be close to zero. We identify the appropriate degree of constraint δ_i on NODE parameters via cross-validation. To do that, we split the

interpolated data into a train, validation, and test set, for instance, in three thirds, and then fit the NODE model to the train set and predict the validation set. We repeat this process for increasing values of δ_i, until we find the value that maximises the log likelihood of the validation data. We can perform multiple folds of validation by swapping the train and validation set, or by varying the size of the train/validation split. Ultimately, once we have identified the appropriate value of the constraint parameters δ_i, we fit the model to both the training and validation set, and assess the accuarcy of the predictions on the test data, which is never seen by the model during training.

2.4 | Inference and uncertainty quantification

Finally, we estimate uncertainty in parameter values by anchored ensembling, which produces approximate Bayesian estimates of the posterior distribution of the parameters (Pearce et al., 2018). This involves sampling a parameter vector from the prior distributions, $\theta_i \sim \mathcal{N}(0, \delta_i)$, and then optimising the posterior distribution from this starting point, $\theta_i^* = \underset{\theta_i}{\text{argmax}} \log p(\theta_i \mid \omega)$. By repeatedly taking samples, the sampled distribution θ^* approaches the posterior distribution and provides estimates and error around the quantities that can be derived from the models. The expectation and uncertainty around derived quantities can then be obtained by computing the mean and variance of the approximated posterior distributions. The strength of this approach is that it is unlikely to get stuck in local maxima, hence providing a more thorough exploration of the parameter space.

2.5 | Analysing neural ordinary differential equations

In this study we are mainly interested in two outcomes of NODEs, namely inferring the direction (or effect) and strength (or contribution) of interactions between the state variables (Bonnaffé, Sheldon, et al., 2021). We define the direction of the interaction between variable y_i and y_j as the derivative of the dynamics of y_i with respect to y_j, and vice versa (Sugihara et al., 2012),

$$e_{ijt} = \frac{\partial}{\partial y_j} \frac{dy_i}{dt}. \tag{9}$$

Knowing the direction, however, is not sufficient to determine the importance of a variable for the dynamics of another. Given the same effects, a variable that fluctuates a lot will have a greater impact on the dynamics of a focal variable, compared with a variable that remains quasi-constant. For example, a predator can have a negative effect on the prey population, but its actual impact/contribution to the dynamics of the prey population depends on its own dynamics, that is if the predator population decreases, it has a positive contribution to the change in growth rate of the prey population. We hence compute the strength of the interaction by multiplying the dynamics of a variable y_j by its effect on the focal variable y_i, also known as the Geber method (equation 3 in Hairston et al., 2005),

$$c_{ijt} = \frac{dy_j}{dt} \frac{\partial}{\partial y_j} \frac{dy_i}{dt}. \tag{10}$$

To summarise results across the entire time series, we can compute the mean effects e_{ij} by averaging e_{ijt} across all time steps, $e_{ij} = 1 / N^{(p)} \sum_t e_{ijt}$, as well as the relative total contribution, c_{ij}, of a variable to the dynamics of another by computing the relative sum of square contributions, $c_{ij} = \left(\sum_{ijt} c_{ijt}^2 \right)^{-1} \sum_t c_{ijt}^2$. By computing the direction and strength of interactions between all the variables in the system we can build dynamically informed ecological interaction networks (e.g. Figure 6). Other metrics can be computed by analysing the NODEs, such as equilibrium states, these are discussed in our previous work (Bonnaffé, Sheldon, et al., 2021).

3 | CASE STUDIES

3.1 | Case study 1: Artificial tri-trophic prey–predator oscillations

In this first case study, we aim to demonstrate the accuracy of the NODE fitted by BNGM in inferring non-linear per-capita growth rates in a system where truth is known. Hence, we simulate a set of time series from a tri-trophic ODE model with known equations and parameters, and we compare the fitted NODEs to the actual ODEs.

3.1.1 | System

We consider a tri-trophic ODE system consisting of a prey, an intermediate predator, and a top predator. The system is built on the real tri-trophic system featuring algae, flagellates and rotifers, considered in case study 2 (Hiltunen et al., 2013),

$$\begin{aligned} \frac{dG}{dt} &= \left(\alpha \left(1 - \frac{G}{\kappa} \right) - \frac{\beta B}{1 + \delta G} - \frac{\gamma R}{1 + \delta G} \right) G \\ \frac{dB}{dt} &= \left(\frac{\beta G}{1 + \delta G} - \phi R - \mu \right) B \\ \frac{dR}{dt} &= \left(\frac{\gamma G}{1 + \delta G} + \phi B - \nu \right) R, \end{aligned} \tag{11}$$

where G, B and R, correspond to the prey, intermediate, and top predator population densities, respectively, α is the prey intrinsic growth rate, limited by a carrying capacity κ, β and γ are the predation rates by the intermediate and top predator, δ is the saturation rate of prey predation, which emulates the capacity of the algae to display predator defence at higher algal density (Hiltunen et al., 2013), ϕ is the predation rate of the intermediate predator by the top predator, μ and ν are the intrinsic mortality of the intermediate and top predator.

We simulate a case of invasion, by introducing the top predator at a low density, with a set of parameters that result in dampening

prey–predator oscillations, namely $\alpha = 1$, $\beta = 2.5$, $\gamma = 1.5$, $\kappa = 3$, $\delta = \phi = \mu = \nu = 1$. We focus on the middle section of the time series, $t \in [20, 50]$, as in the initial section the top predator is rare, and in the later section, populations have attained a fixed equilibrium point. The resulting time series are presented in Figure 2.

3.1.2 | Neural ordinary differential equation model

To non-parametrically learn the per-capita growth rate of each species, and to derive ecological interactions, we define a three-species NODE system,

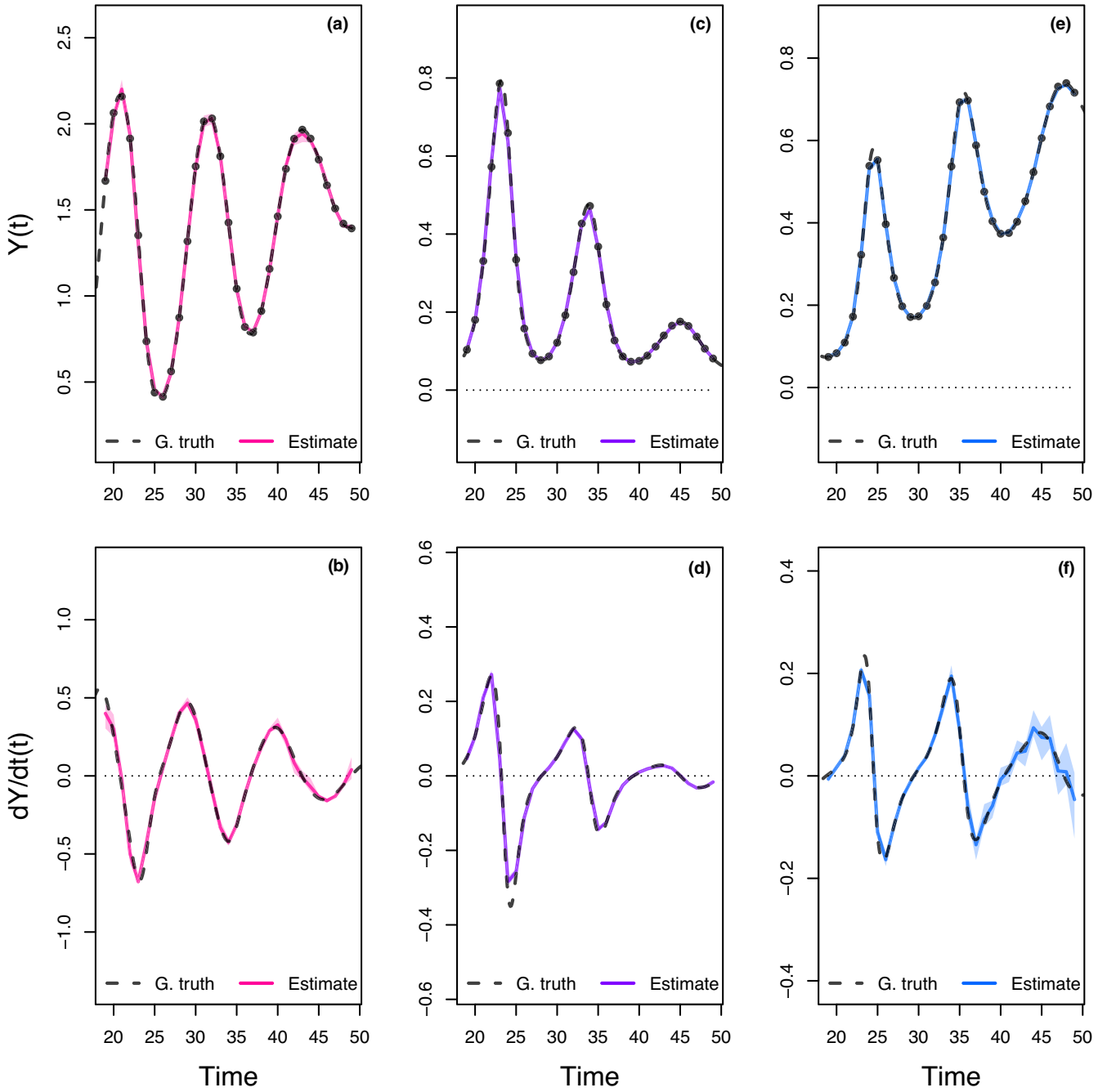

FIGURE 2 Interpolated density and dynamics of prey, intermediate, and top predators in the artificial system. This figure corresponds to the first step in the overview figure (Figure 1). It shows the accuracy of the interpolated densities of prey (a), intermediate (c), and top predators (e). We obtain 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 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.

$$\begin{aligned}\frac{dR}{dt} &= r_R(R,G,B,\beta_R)R,\\ \frac{dG}{dt} &= r_G(R,G,B,\beta_G)G,\\ \frac{dB}{dt} &= r_B(R,G,B,\beta_B)B,\end{aligned} \tag{12}$$

where the per-capita growth rates r_R, r_G and r_B are neural network functions of the density R, G, B of each species (function f_p, Equation 2). We choose a combination of linear and exponential activation functions $f_{\sigma,j\leq J/2}(x) = x$, and $f_{\sigma,j>J/2}(x) = \exp(x)$. This allows us to progressively switch from a simple linear model to a non-linear model by releasing the constraint on the parameters of the network during cross-validation. The number of units in the hidden layer J is chosen to be 10, as this is a commonly used number for systems of that size (e.g. Bonnaffé, Sheldon, et al., 2021; Wu et al., 2005).

3.1.3 | Time-series interpolation

We interpolate the time series using the neural network described in Section 2.3 (Equation 4). We set the number of neurons in the network to $J = 30$. We use sinusoid activation functions, $f_\sigma(x) = \sin(x)$, so that the weights $\omega_{ij}^{(1)}$, $\omega_{ij}^{(2)}$ and $\omega_{ij}^{(3)}$ control the amplitude, shift and frequency of the oscillations in the time series, respectively. Given that the population densities are strictly positive $R, G, B \in \mathscr{R}^+$, we use an exponential link function, $f_\lambda(x) = \exp(x)$. We then approximate the marginal posterior distribution of the interpolation parameters, and thereby of interpolated states and dynamics, by taking 100 samples from the log marginal posterior distribution (Equation 7) via anchored ensembling. In practice, the high number of parameters in the neural network equation may impede the fit of the time series, especially for short time series. We found that dividing the number of parameters $M^{(o)}$ (Equation 7) by the number of neurons in the network J (Equation 2) improved the quality of the interpolations. Interpolated states and dynamics are presented in Figure 2.

3.1.4 | Fitting neural ordinary differential equations to the interpolated time series

We fit the NODE system to the interpolated time series. In practice, we fit the NODE to the expectation of the interpolated state and dynamics, $\mathrm{E}(\tilde{y}_i)$ and $\mathrm{E}(d\tilde{y}_i/dt)$, by averaging over all sampled interpolation parameters. An alternative approach could be to consider the interpolation that maximises the log marginal posterior density, but this may decrease repeatability due to the difficulty of reliably identifying a global maximum. Averaging across multiple interpolations ensures an overall smoother and robust interpolation. In addition, we standardise the response and explanatory variables with respect to their mean and standard deviation (i.e. $Z = (Y - \mu)/\sigma$). This is to facilitate the training of the NODE by equalising the scale of the different parameters in the neural network. Then, we identify the optimal regularisation parameter δ (Equation 8) by cross-validation. To do that, we split the data in three thirds, train NODEs on the first third, and calculate the log likelihood of the validation set for increasing values of δ, from 0.1 (linear) to 1.0 (highly non-linear), by increments of 0.1. This allows us to identify the maximum degree of non-linearity, δ, in the per-capita growth rate that ensures generalisability throughout the time series. Then, we approximate the posterior distribution of the NODE parameters by taking 30 samples from the posterior distribution (Equation 8). We ensure moderate temporal autocorrelation and normality by visualising the residuals of the models. We also ensure results repeatability by running the entire fitting process a second time.

3.1.5 | Computing ecological interactions

Finally, we analyse the shape of the per-capita growth rates to recover the interaction between the three species in the system. In particular, we look at the effect and contribution of each species to the dynamics of the others. Effects are computed as the sensitivity (i.e. the gradient) of the per-capita growth rate of a given species with respect to the density of the other species (Bonnaffé, Sheldon, et al., 2021; Sugihara et al., 2012). Contributions are computed following the Geber method (Hairston et al., 2005), which consists in multiplying the dynamics of a variable by its effects on the other variables. We further compute the importance of a species in driving the dynamics of another by computing its relative total contribution compared with other species. More details on how to compute these quantities can be found in Section 2.5 and in our previous study (Bonnaffé, Sheldon, et al., 2021).

3.1.6 | Benchmark

To demonstrate the suitability of BNGM for fitting NODEs and inferring ecological interactions we compare our approach with three existing methods. For this purpose, we focus on the artificial time series, as this offers the possibility for comparing predictions to the ground truth, known from the equations that generated the time series.

We first consider a standard NODE model (Bonnaffé, Sheldon, et al., 2021), as our BNGM approach seeks to alleviate the computational cost of fitting NODEs. We define the per-capita growth rate as an ANN with a single layer, 3 inputs, 10 hidden nodes, and exponential activation functions. We use a Bayesian model, assuming log normal distributions for species density $Y_i \sim \log\mathcal{N}(y_i, \sigma_i)$, and uniform uninformative prior distributions for the network parameters $\theta_i \sim \mathcal{U}(-10, 10)$, initial densities $y_{i0} \sim \mathcal{U}(0, 10)$ and variance $\sigma_i \sim \log\mathcal{N}(0.5, 0.5)$. Our implementation of standard NODEs differs from our BNGM approach in three ways. First, the standard NODE ANN has three outputs instead of one, as variables are fitted jointly. Second, computing the posterior density of the parameters requires to solve the NODE system with a numerical ODE solver

(Runge–Kutta, package deSolve). Third, we do not constrain the parameters of the network given that the prohibitive fitting times prevent the tuning of the regularisation parameters.

We also consider a parametric ODE model, as this is the closest parametric alternative to NODEs to infer ecological interactions. This model only differs from the standard NODE model in that the per-capita growth rate is approximated by second-order polynomial functions $r_i(y,\theta_i) = \theta_i^{(0)} + \sum_j \theta_{ij}^{(1)} y_j + \sum_j \sum_k \theta_{ijk}^{(2)} y_j y_k$, instead of an ANN, which can handle simple non-linearities.

To ensure the most meaningful comparison, we implemented the NODEBNGM, standard NODE and parametric ODE models in base R, using BFGS for optimisation (function *optim*, R v4.2.0). We also followed a similar fitting procedure by independently training 30 models on the train/validation set (i.e. two-third of the time series) and predicting the test set (remaining third).

Finally, we implement CCM. This technique performs locally linear approximations of the state space of the system to estimate the sensitivity of the dynamics of a variable to a change in other variables (Sugihara et al., 2012). For this we use the package rEDM (v1.13.1; Sugihara et al., 2012), and adapt the example code provided for the three species system. We fit the CCM model on the train set and predict outcomes on the test set. We then retrieve s-map coefficients (i.e. the interactions) and approximate the population dynamics and per-capita growth using finite differences, given that the standard implementation of CCM does not provide these estimates by default.

For all four methods, we compute the runtime as the average time required to train a single model. Using the best preforming model on the train set, we then predict the population dynamics, growth rate and ecological effects for the entire time series, including the test set. We compute the accuracy of the predictions by computing the mean sum of squared error of predictions versus the truth for both the per-capita growth rate and ecological effects, on the train and test set. We also build the corresponding dynamical interaction networks, using the inferred mean effects and total contributions, and compare them with the true network of interactions. Results are shown in Figure 4 and presented in detail in Figures S1–S5.

3.2 | Case study 2: Real tri-trophic prey–predator oscillations

In this second case study, we want to assess the quality of the NODE analysis when performed on a real time series. We are further interested in comparing the direction and strength of uncovered ecological interactions across virtually identical replicated time series.

3.2.1 | System

We consider a three-species laboratory microcosm consisting of an algal prey *Chlorella autrophica*, a flagellate intermediate predator *Oxyrrhis marina* and a rotifer top predator *Brachionus plicatilis*. The algal prey is consumed by the intermediate and top predator, which also consumes the intermediate predator (Arndt, 1993). The dynamics of this system, here the daily change in the density of each species, were recorded in three replicated time-series experiments performed by Hiltunen and colleagues (Hiltunen et al., 2013). We use their time series because they describe a simple yet biologically realistic ecosystem, and because the quality of the replication of their microcosm reduces as much as possible observational and experimental error, and rules out environmental variation (Hiltunen et al., 2013). We digitised these time series by extracting by hand the coordinates of every points in the referential of the axis of the graph of the original study and analysed them.

3.2.2 | Neural ordinary differential equation analysis

We apply the same analysis as performed on the artificial tri-trophic prey–predator oscillations. This allows us to recover a non-parametric approximation of the growth rate of each species, and then derive the direction and strength of the ecological interactions that underpin their dynamics. We present detailed results of the analysis of the second time series (Figure 5), and a summary comparison of the three time series (Figure 6). Complementary results, including cross-validation plots, and detailed results for the other two replicates can be found in Appendices C–E.

3.3 | Case study 3: Real di-trophic prey–predator oscillations

We infer ecological interactions by NODEBNGM in the hare-lynx system (Odum & Barrett, 1972). This is to provide an example of a longer time series, and to offer a point of comparison with previous and future implementations of NODEs, which commonly use this time series (e.g. Bonnaffé, Sheldon, et al., 2021; Frank, 2022).

3.3.1 | System

The system is described in details in our previous work (Bonnaffé, Sheldon, et al., 2021). The data consist in a 90-year long time series of counts of hare and lynx pelts collected by trappers in the Hudson Bay area in Canada (Odum & Barrett, 1972). The time series displays characteristic 10-year long prey–predator oscillations.

3.3.2 | Neural ordinary differential equation analysis

We apply the same analysis as previously described, with the exception that the NODE system only features two variables, *H* and *L*, instead of 3. Results are presented in Figure 7.

3.4 | Case study 4: Maizuru Bay aquatic community

Finally, we demonstrate the capacity of NODEs to analyse the drivers of the dynamics of a larger community by analysing the time series of the Maizuru Bay community (Ushio et al., 2018).

3.4.1 | System

The dataset for this system consists of 12-year long time series of fortnight abundance estimates of the 15 dominant species in the Maizuru Bay, Japan. The data was collected every 2 weeks along three 200 m long and 2 m wide transects by underwater visual census conducted along the coast of the Maizuru fishery research station of Kyoto University from 2002 to 2014 (for more details see Ushio et al., 2018). Bottom sea temperature (at 10 m depth) was also recorded on each census. The dataset contains 14 dominant species of fish and 1 genus of jellyfish. Only species with more than 1000 sightings were included in the final dataset.

We focussed our analysis on the species with the least sparse records. We discarded the following species from our analysis: *Engraulis japonicus*, *Plotosus lineatus*, *Chaenogobius gulosus* and *Siganus fuscescens*. We also excluded periods that presented jellyfish blooms, as these were isolated events that could cause numerical errors in the estimation of the dynamics of species abundance. In total, we considered a time period of a hundred time steps from June 2004 to August 2008, and 11 species out of 15, namely *Aurelia* sp., *Sebastes inermis*, *Trachurus japonicus*, *Girella punctata*, *Pseudolabrus sieboldi*, *Halichoeres poecilopterus*, *Halichoeres tenuispinnis*, *Pterogobius zonoleucus*, *Tridentiger trigonocephalus*, *Sphyraena pinguis* and *Rudarius ercodes*. We included the sea bottom temperature (in °C) as an additional environmental variable.

3.4.2 | Neural ordinary differential equation analysis

We then analysed this dataset following the approach described in Section 2. We split the data into three-thirds to create a training, validation and testing set, and we followed the same procedure as described before to tune the regularisation parameters. The NODE system consisted of 11 NODEs, where the per-capita growth rate is determined by a single-layer ANN with 12 input nodes, 10 hidden nodes, a mixture of linear and exponential activation functions and 1 output node. Due to the high dimensionality of the effects obtained (11 by 12), we only present mean effects and relative total contributions, obtained by taking the mean, and the relative mean squares, respectively, of the effects and contributions across the entire time series (Figure 8). The time series of effects and contributions are presented in greater details in Appendix G.

4 | RESULTS

4.1 | Case study 1: Artificial tri-trophic system

We present the results of fitting NODEs by BNGM to the artificial tri-trophic time series in Figures 2 and 3. We find that both the interpolation of the state variables and dynamics are highly accurate (Figure 2), given that they closely match the ground truth, known from the equations of the ODE model that we used to generate the time series (Equation 11). Similarly, we find that the NODE approximation of the per-capita growth rate of each species also closely matches the ground truth (Figure 3a,d,g). We find negative non-linear effects of the two predators on the growth rate of the prey (Figure 3b, blue and purple lines). This non-linear pattern is mirrored by the effect of the prey on the growth rate of the predators (Figure 3e,h, red line). The interaction between the two predators is also well-recovered (Figure 3e, blue line, and h, purple line), in spite of a slight tendency for overestimating the degree of non-linearity of effects. The BNGM approach hence accurately recovers the dynamical characteristics of the artificial system.

4.2 | Benchmark

Figure 4 shows the performance of fitting NODEs by BNGM, compared with standard NODEs, parametric ODEs, and CCM models. We find that fitting NODEs by BNGM provides the highest estimation accuracy for growth rates and ecological effects, both on the training and test set, as well as competitive runtimes. Standard NODEs provide similar estimation accuracies but take over 15 min to train. CCM is the fastest technique, as results are obtained in under a second, but is relatively less accurate. Parametric ODEs are found to be both slow and less accurate.

We present a detailed breakdown of the runtime of fitting NODEs by BNGM for each system in Table 1. We find that it takes on average 5.35 min to fit NODEs by BNGM on the smaller systems with $I = 3$ or less species, and about 23 min to fit the larger system from the Maizuru Bay, which features $I = 12$ variables. This includes performing $100 \times I$ and $30 \times I$ full optimisations of the posterior distribution of the interpolation and NODE parameters, respectively. This amounts to about 5.37 s to sample each variable of the NODE system once in the smaller systems, and about 28 s for the Maizuru Bay community. This is over a 100 fold improvement over standard NODE models, which take on average 20 min (Figure 4).

4.3 | Case study 2: Real tri-trophic prey–predator oscillations

We present an in-depth analysis of the drivers of the dynamics of the algae, flagellate, and rotifer population in replicate B (Figure 5). We find slightly positive non-linear intra-specific density dependence

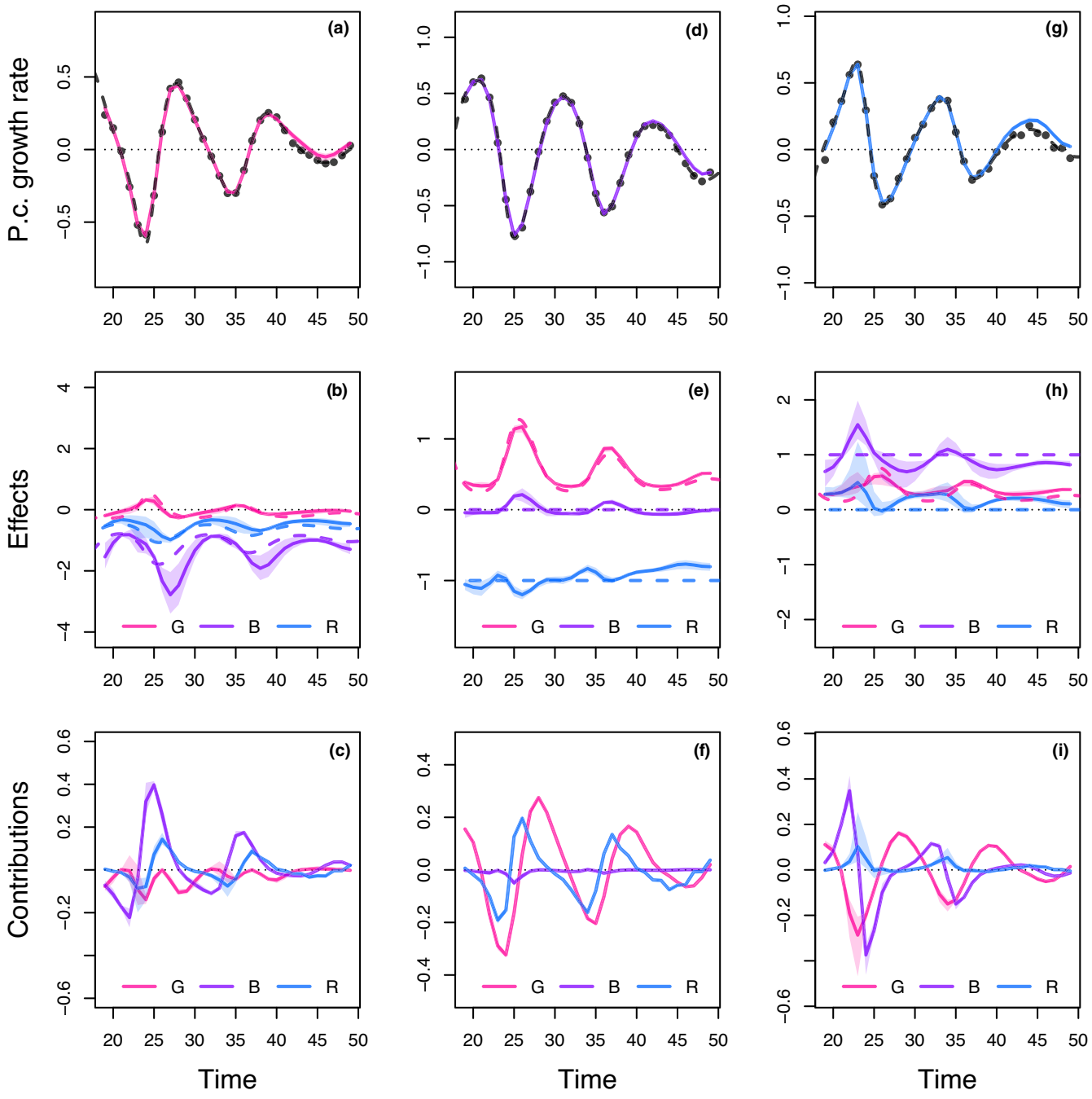

FIGURE 3 Drivers of dynamics of prey, intermediate, and top predator in the artificial system. This figure corresponds to the second step in the overview figure (Figure 1). It displays the NODE non-parametric approximations of the per-capita growth rate of prey (a–c), intermediate (d–f), and top predators (g–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 (Figure 2b,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.

in algal growth (Figure 5b, red line), and negative non-linear interspecific effects of the two predators (purple and blue line). We find that the growth rate of the flagellate is driven by a positive effect of algal density, a negative effect of predation by the rotifer and intraspecific density dependence (Figure 5e,f). The rotifer population is almost solely driven by a positive non-linear effect of algal density (Figure 5h, red line). Overall, comparing results across the three replicates reveals that the effect of the rotifer population on the flagellate and algae, and the effect of the algae on the rotifer, are the strongest and most consistent interactions (Figure 6; Table 2).

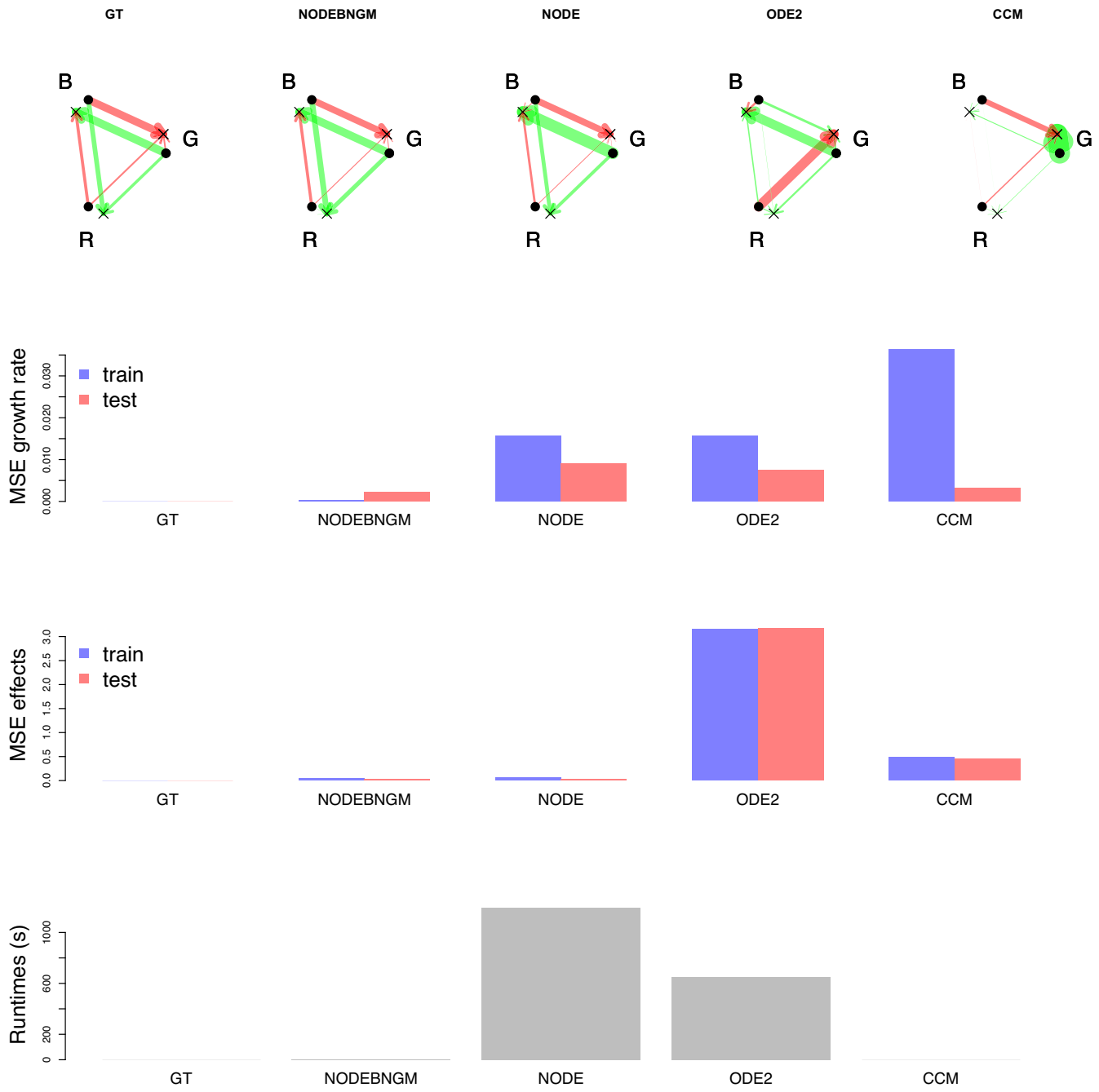

FIGURE 4 Runtimes and accuracy of NODEs fitted by BNGM compared with standard NODEs, ODEs, and CCM. The NODEBNGM method (non-parametric) involves fitting a NODE system by Bayesian neural gradient matching (BNGM). The NODE method (non-parametric) involves fitting a NODE system with an ODE solver. The ODE2 method (parametric) involves fitting an ODE system with polynomial functions of species densities with an ODE solver. The CCM method (non-parametric) involves computing locally linear approximations of the state space. For each method, we trained 30 models on the two first thirds of the artificial time series where ground truth is known (Figure 2). We computed runtimes as the mean time (in seconds) required to train a single model. Using the best identified model, we predicted the growth rate and effects on the train and test set. We computed accuracies as the mean squared error of predictions vs ground truth (known from the equations that generated the data) (see Figures S1–S5 for more details). At the top, we show the dynamical interaction network of the system predicted by the best models, where G, B, R correspond to the prey, intermediate and top predator, respectively. Green and red colours correspond to positive and negative interactions, the width of arrows is proportional to relative total contribution to the growth rate of the population computed following the Geber method (Bonnaffé, Sheldon, et al., 2021; Hairston et al., 2005).

TABLE 1 Summary of model runtimes. We measured the time required to perform 100 interpolations and 30 NODE fits to each variable in the systems. Replicate A, B, and C correspond to each replicated time series of the aglae, flagellate, and rotifer tri-trophic system (Hiltunen et al., 2013). The Hare-Lynx system correspond to the 90 years long time series of hare and lynx pelt counts (Odum & Barrett, 1972). The Ushio system corresponds to the 100 time step long time-series dataset of the 11 dominant species in the Maizuru Bay community (Ushio et al., 2018). The number of time steps (*N* steps) is given for each time series. The total time per fit is obtained by dividing the total time in seconds by the number of fits (i.e. 130). These results were obtained on a macbook pro M1 MAX 2021, in base R (v4.0.2), with non-optimised code.

			Interpolation		NODE fit			
System	*N* variables	*N* time steps	*N* fits	Time (s)	*N* fits	Time (s)	Total	Total p. fit
Replicate A	3	66	100	234.45	30	80.41	314.86	5.02
Replicate B	3	66	100	238.01	30	81.13	319.14	5.08
Replicate C	3	40	100	136.51	30	49.94	186.45	3.03
Hare-lynx	2	90	100	341.74	30	21.55	363.29	4.14
Ushio	12	100	100	806.12	30	604.42	1410.54	28.21

The interactions of the flagellate with the algae and its effect on the rotifer population varies substantially across replicates (Figure 6; Table 2).

4.4 | Case study 3: Real di-trophic prey–predator oscillations

We present the analysis of the drivers of the hare-lynx population dynamics in Figure 7. Cross-validation provides weak support for non-linear effects in the per-capita growth rate of the hare and lynx. We find that the hare population growth rate is mostly determined by a non-linear negative effect of the lynx population (Figure 7b,c, blue line), and by weak non-linear positive density dependence (red line). The lynx growth rate is determined by a positive non-linear effect of the hare (Figure 7e,f, red line), and to a lesser extent by negative non-linear intra-specific density dependence (blue line).

4.5 | Case study 4: Drivers of the Maizuru Bay community dynamics

We show the results of the NODE analysis of the drivers of the dynamics Maizuru Bay community in Figure 8. Our main finding is that the chameleon goby (*T. trigonocephalus*) has a strong negative effect on 8 of the 11 dominant species of the community. We find that *E. ercodes* also has a strong negative impact on other species in the community, although relatively smaller than that of the chameleon goby. We find a positive effect of sea bottom temperature on the growth rate of the chameleon goby. Other effects are found to be mostly positive and have a relatively smaller impact on community dynamics.

5 | DISCUSSION

Characterising ecological interactions from time-series data is challenging. This is due to the fact that interactions can be highly context-dependent processes (Song & Saavedra, 2021), making it difficult to identify parametric models that encapsulate their complexity (Wood, 2001). Interactions estimated with parametric models are contingent on the parameterisation arbitrarily chosen by the observer, and hence risk being biased (Adamson & Morozov, 2013; Wood, 2001). We provide a novel method for estimating ecological interactions non-parametrically, by using NODEs fitted with BNGM. First, we remove the cost of fitting NODEs by introducing BNGM, which allows for NODE fitting in only a few seconds. The method involves interpolating time series and dynamics with neural networks, and then fitting NODEs to interpolated dynamics with Bayesian regularisation. We further demonstrate that this approach is accurate, as NODEs approximate with minimal error the ecological interactions in artificial time series, where real interactions are known, performing better than three existing methods. Finally, we estimate the strength, direction, importance, and non-linearity of ecological interactions in three natural and experimental systems, showing variation in ecological interactions within and across the time series.

5.1 | Performance of neural ordinary differential equations fitted by Bayesian neural gradient matching compared with existing methods

Our approach relies on approximating population dynamics with NODEs and then computing their sensitivity to a change in the density of the different populations in the system (Bonnaffé, Sheldon, et al., 2021). We demonstrate that NODEs fitted by BNGM accurately recover the dynamics, strength, direction and non-linearity of ecological interactions in artificial tri-trophic prey–predator time series, where truth is known. In particular, we find that the interactions between the prey and the two predators are non-linear, and thereby oscillate throughout the time series, which is consistent with the ground truth model, that features a resistance to predation at high prey density. We also recover the interactions between the two predators, in spite of a slight tendency to overestimate the degree of non-linearity. To our knowledge, this is the

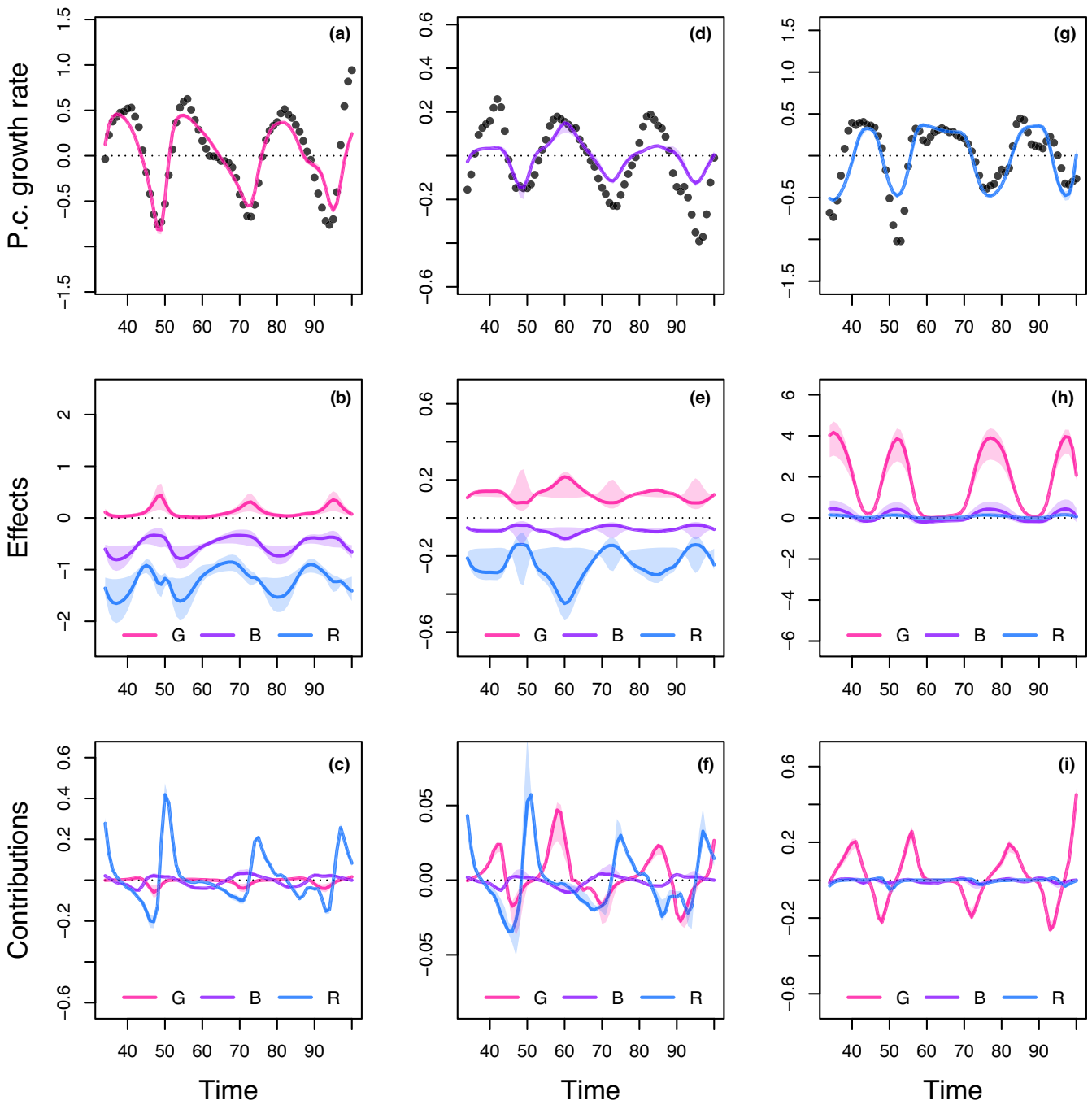

FIGURE 5 Drivers of dynamics of algae, flagellate, and rotifer in replicate B. This figure displays the NODE non-parametric approximations of the per-capita growth rate of algae (a–c), flagellate (d–f), and rotifer (g–i). We obtain the NODE approximations (a, d, g, solid line) by fitting the interpolated 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).

first assessment of the accuracy of NODEs in recovering interactions between variables from time-series data, as most of the work focuses on assessing the accuracy of the fitting and forecasting of time series (e.g. Chen et al., 2019; Frank, 2022; Mai et al., 2016; Treven et al., 2021).

We find that fitting NODEs by BNGM provides higher estimation accuracies of ecological interactions and reduces substantially fitting times compared with standard NODEs (Bonnaffé, Sheldon, et al., 2021), and parametric ODEs (Rosenbaum et al., 2019). This difference is attributable to three factors. First, BNGM alleviates

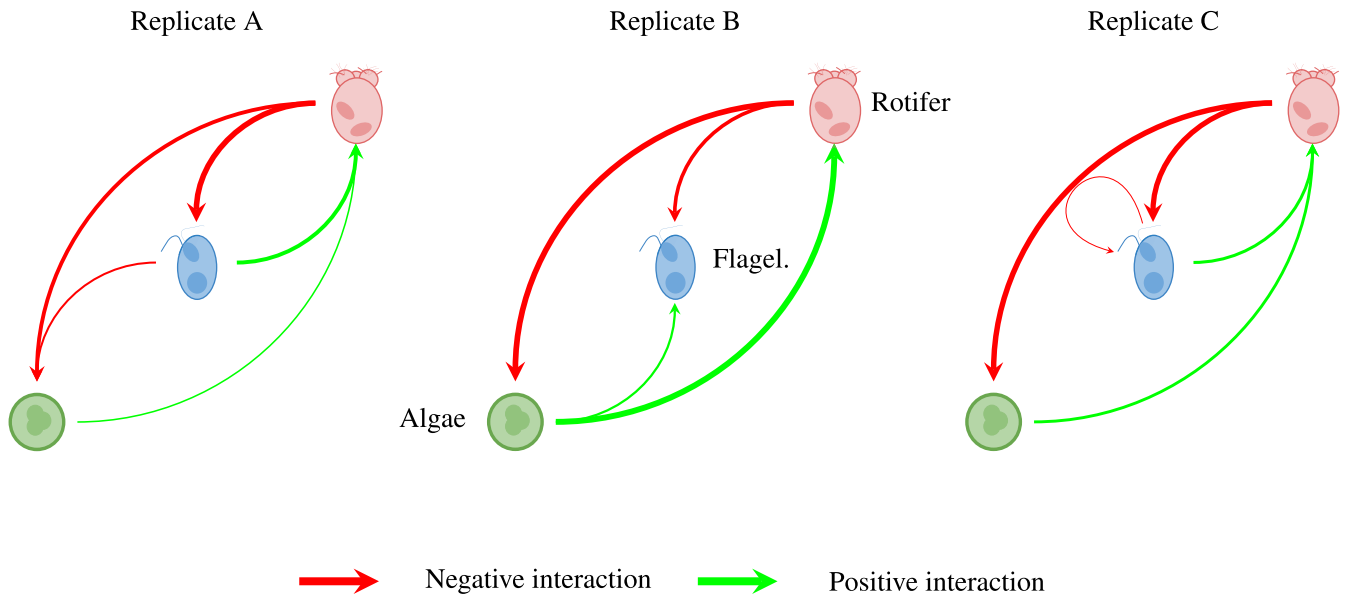

FIGURE 6 Interaction networks inferred from three replicated time series of algae, flagellate, and rotifers. This figure shows the direction and strength of ecological interactions inferred from three replicated sets of time series of algae, flagellate, and rotifer, using NODEs fitted by BNGM. The replicates A and C were analysed in the same way as replicate B (see Figure 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 blue line in Figure 5c, and comparing it with the sum of squares of all contributions (Figure 5c, red, purple and blue lines). We provide the value of the mean effects and relative strengths in Table 2. The replicated time series were obtained by digitising the time series in Hiltunen et al. (2013).

the need for numerically solving the NODE system, which makes it faster to evaluate the posterior distribution. Second, it allows for the calculation of analytical gradients of the posterior distribution, which greatly improve the speed and efficiency of the gradient descent optimisation algorithm. Finally, it makes it possible to fit each variables independently of the others, which results in a simpler optimisation problem.

CCM remains faster than our approach in recovering estimates of ecological interactions (Sugihara et al., 2012), however its accuracy is lower. A possible explanation for this comes from the fact that CCM computes the sensitivity of the total population growth rate, rather than the per-capita growth rate, which can change estimated effects. Additionally, CCM relies on piecewise linear reconstructions of the state space (Deyle et al., 2015), whereas NODEs compute a global non-linear approximation of the per-capita growth rate on the entire range covered by the data. We view the former as potentially more sensitive to local noise in the state space (Cenci et al., 2019), compared with the latter, which uses all evidence available to inform local inference.

Our BNGM approach extends standard gradient matching, by using ANNs as interpolating functions, and Bayesian regularisation to control the non-linearity of the processes (Cawley & Talbot, 2007). The use of ANNs as interpolating functions sets it apart from the initial approach of Ellner et al., who use splines to interpolate the time series before approximating the ODEs (Ellner et al., 2002). ANNs are more general and flexible than splines, as well as being easier to manipulate given that they are defined continuously on the state space, which is especially useful when handling multiple interactions between variables. Our approach is related to that of Wu et al., who use ANNs to approximate both the states and ODEs of prey–predator systems (Wu et al., 2005), as well as that of Treven and colleagues, who developed the Gaussian process equivalent (Treven et al., 2021). In both approaches, they train the interpolation functions at the same time as the NODEs, to constrain the interpolation of trajectories such that they can be achieved by the NODE system, which thereby introduces dynamical coupling between state variables. One of the downsides of this approach is that misestimating one of the state variables of the model biases the estimation of the states and dynamics of other variables. To avoid this, we fit each interpolation and NODE independently to each time series. In addition, this makes it possible to parallelise the code, resulting in potentially even faster computation.

Our approach opens new possibilities for non-parametric inference of ecological interactions from time-series data. The lower fitting times makes it possible to tackle larger systems, quick and extensive model comparison, cross-validation, and apply more thorough statistical treatments of the uncertainty of these models, for instance by implementing Markov-chain Monte-Carlo sampling.

TABLE 2 Comparison of the direction and strentgh of ecological interactions estimated by BNGM across three replicated tri-trophic microcosms. 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. The variables *G*, *B*, and *R* correspond to the population density of algae, flagellate, and rotifer respectively. r^2 corresponds to the r squared of the NODE non-parametric approximation of the per-capita growth rate of the three species.

		G	B	R
Replicate A	r^2	0.11	0.37	0.47
Mean effects	on G	−0.08	−1.14	−1.13
	on B	0.28	−0.21	−0.66
	on R	0.60	1.09	0.32
Prop. of total contributions	to G	0.01	0.34	0.65
	to B	0.02	0.04	0.93
	to R	0.26	0.66	0.08
Replicate B	r^2	0.52	0.4	0.65
Mean effects	on G	0.12	−0.53	−1.23
	on B	0.12	−0.06	−0.25
	on R	1.83	0.10	0.07
Prop. of total contributions	to G	0.02	0.03	0.95
	to B	0.38	0.02	0.61
	to R	0.99	0.00	0.01
Replicate C	r^2	0.59	0.32	0.73
Mean effects	on G	0.09	−0.49	−1.96
	on B	0.04	−0.19	−0.65
	on R	1.07	0.79	−0.01
Prop. of total contributions	to G	0.00	0.08	0.91
	to B	0.01	0.10	0.88
	to R	0.49	0.51	0.00

5.2 | Ecological interactions in real prey–predator systems

We further tested NODEs in a real setting, by inferring ecological interactions across three replicated time series of an experimental tri-trophic system of algae, flagellate and rotifer populations (Hiltunen et al., 2013). Our approach reveals that only stronger interactions, namely the negative effects of the rotifer top predator on the other species, and the positive effect of algae on the rotifer, are conserved across the three replicated time series. We also find evidence for non-linearity in the dynamics of the rotifer, as the positive effect of the algae on rotifer growth oscillates throughout the time series. This is consistent with the biology of the system, as the algae tends to form anti-predation clumps at higher density, which would dampen the positive effect of algal density on rotifer growth at high algal density (Hiltunen et al., 2013; Yoshida et al., 2003). We find it interesting that the weaker interactions with the flagellate predator are not consistent across time series, given the controlled laboratory conditions. This system is known to evolve rapidly, it is hence possible that fast evolution of the different populations from the onset of the time series may have driven the system onto different attractors (Hiltunen et al., 2013; Yoshida et al., 2003, 2007). Additionally, stochasticity in population dynamics may have a similar effect (Dallas et al., 2021). Disentangling these two sources of variation would require refining the modelling framework, for instance by explicitly including evolution in the model (e.g. with the Price equation, Ellner et al., 2011), and by using neural stochastic differential equations (i.e. Rackauckas et al., 2019) fitted with a particle filter. While these would constitute interesting developments, our method is still a useful first step, identifying differences between the time series and demonstrating a reasonable amount of deterministic consistency in the dynamics, judging by the cross-validation and fits.

We also analysed the hare-lynx time series (Odum & Barrett, 1972), as it is a common benchmark in the field of time-series analysis, and provides a comparison point with our previous implementation of NODEs (Bonnaffé, Sheldon, et al., 2021). As in our previous study, we found a predatory inter-specific interaction between lynx and hare, and negative intra-specific density dependence in the lynx. Evidence for positive density dependence in the hare was more limited than previously found. We also found stronger evidence for non-linearity, as intra- and inter-specific effects oscillated throughout the time series, as a result of density dependence. This difference with our previous study is due to the fact that our previous implementation of NODEs was based on simulating the full NODE system, and hence imposed dynamical coupling between the variables. This dynamical coupling comes at a cost—if one variable is not explained well by the model, it will bias the interactions and dynamics of other variables. Here, the time series of lynx and hare are analysed independently, each state variable is interpolated as closely as desired, its effects on the dynamics of other variables are hence even more robust to model misspecification than before.

We applied our approach to analyse the drivers of the dynamics of 11 species in the Maizuru Bay (Ushio et al., 2018). We inferred 11×11 ecological interactions and 11 dependencies on water temperature. We found that the chameleon goby had a strong negative impact on the other species of the system, showing a high competitive potential. This species is viewed as an aggressive competitor (Ushio et al., 2018) and is considered an invasive species in places where it has been introduced (Goren et al., 2009). We also find a positive effect of temperature on the growth rate of the chameleon goby, which suggests that warming could have indirect negative effects on many species on Maizuru Bay by favouring the reproduction of the goby.

Surprisingly, our results differ substantially from those obtained by Ushio et al. in their original analysis of the system with CCM (Ushio et al., 2018). This may be due to a several factors. First, we considered a different set of species, as some of the time series that

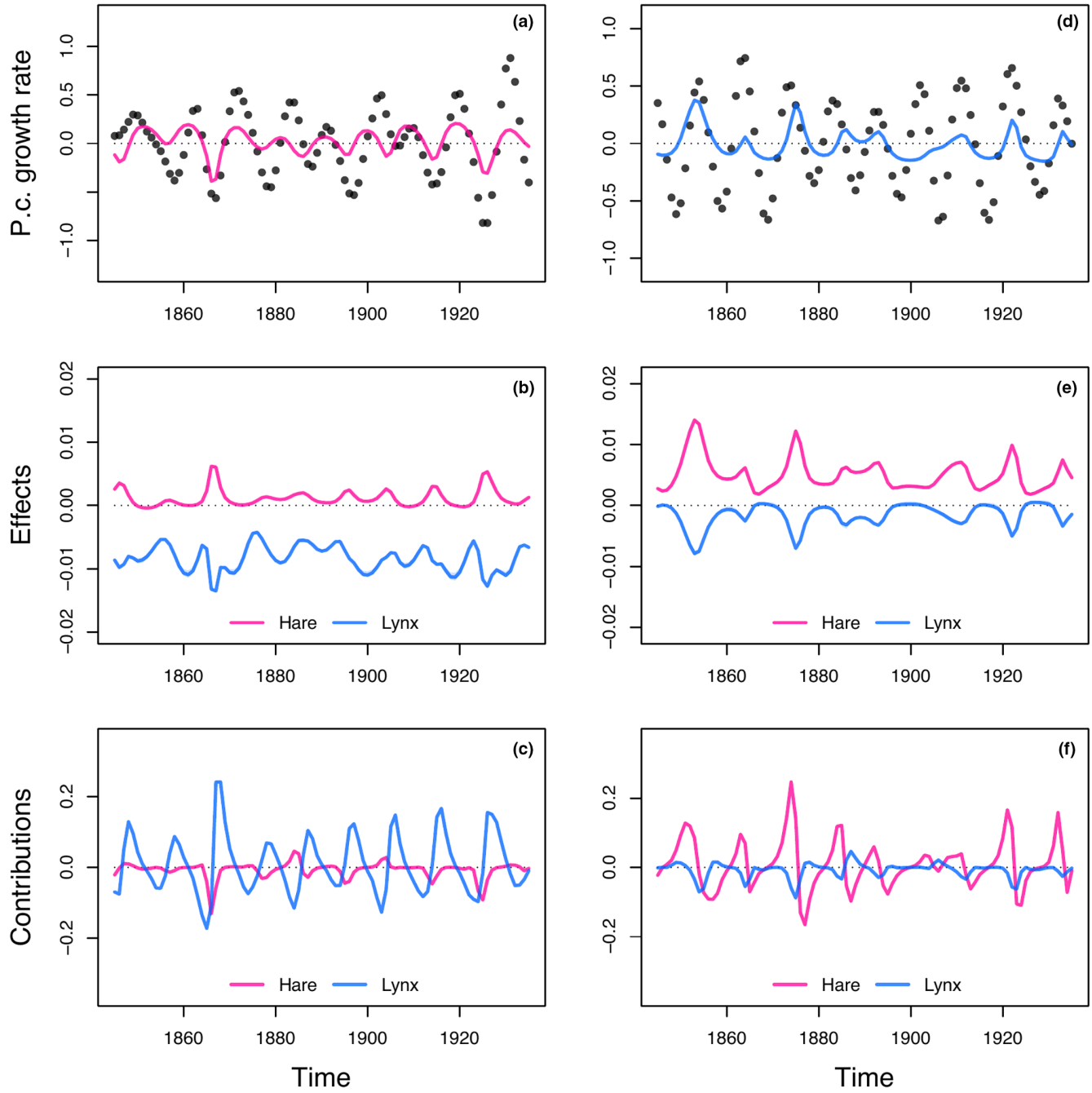

FIGURE 7 Drivers of dynamics of hare and lynx in the Odum and Barrett pelt count time series. This figure displays the NODE non-parametric approximations of the per-capita growth rate of hare (a–c), and lynx (d–f). We obtain the NODE approximations (a, d, solid line) by fitting the interpolated per-capita growth rates (black dots) with ANNs that take interpolated 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.

Ushio and colleagues used were too sparse to be suitable for our analysis. Additionally, this difference may be explained by the disparity in the estimation accuracies revealed by our benchmark analysis, relating to fundamental mathematical differences between the two approaches, as discussed previously. If nothing else, our analysis of the Maizuru community dynamics demonstrates the usefulness of our BNGM method for fitting NODEs to a larger, more realistic system.

Overall, our approach provides a novel and powerful way of estimating interactions non-parametrically from time-series data. The benefit of using NODEs is that they make no assumptions about the nature of the ecological interactions that drive

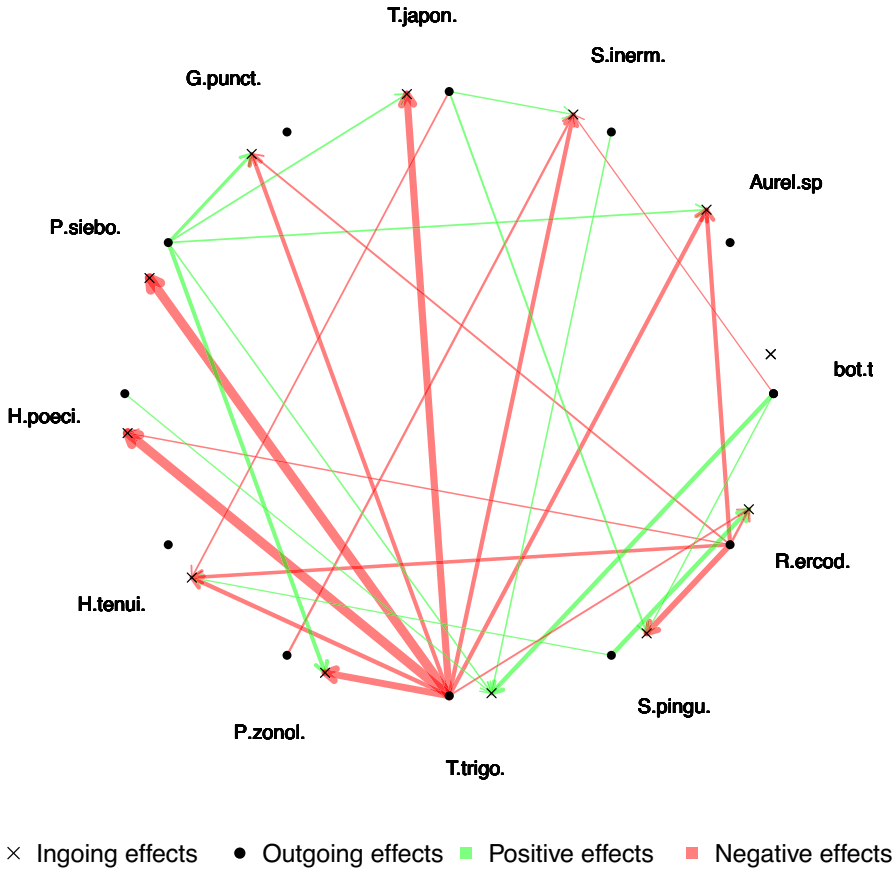

FIGURE 8 Dynamical interaction network of the Maizuru Bay community. This figure summarises the results of the NODEBNGM analysis of the Maizuru Bay community time series (Figure S18). Bot.t corresponds to the temperature at sea bottom. Species are referred to by their shortened acronym. The main species are *Aurelia* sp., *Sebastes inermis*, *Trachurus japonicus*, *Girella punctata*, *Pseudolabrus sieboldi*, *Halichoeres poecilopterus*, *Halichoeres tenuispinnis*, *Pterogobius zonoleucus*, *Tridentiger trigonocephalus*, *Sphyraena pinguis*, and *Rudarius ercodes*. Red and green arrows correspond to negative or positive mean effects, obtained by averaging the sensitivity of the growth rate of a species to the density of other species across the time series. The width of the arrows is proportional to the relative total contribution (in %) of species density to the growth rate of other species, obtained by computing the relative sum the square of contributions across the time series. For the sake of clarity, we only display interactions which contribute to more than 10% of the change in the population growth rates.

the dynamics of the species (Bonnaffé, Sheldon, et al., 2021; Chen et al., 2019). Hence, we have a better chance at estimating the actual value of the interactions, knowing that it is not subjected to potential incorrect model specifications (Adamson & Morozov, 2013; Ellner et al., 2002; Jost & Ellner, 2000; Kendall et al., 2005; Wu et al., 2005).

5.3 | Limits and prospects

One of the main difficulties in quantifying ecological interactions is to identify potential context dependences on other state variables (Song & Saavedra, 2021), for example, whether predation rates are affected by temperature. Our approach allows for the quantification of context dependence, which shows as non-linear fluctuations of interactions throughout the time series. In the present work, we only report non-linearity as evidence for context dependence in the interactions, but we do not attempt to understand what it is attributable to. For instance, we identify non-linear density dependence in the effect of the algae on the rotifer, but we do not know whether this is due to a change in the effect with algae density or rotifer density or both. To disentangle these higher-order effects we could compute the Hessian of the system, namely the second-order derivative of the dynamics with respect to the different state variables. Though this procedure is simple mathematically, it would result in 27 second-order effects to analyse for the simple three species system considered here. This type of analysis would get rapidly out of hand for larger systems. Further work should hence consider how to handle these higher-order effects, as a way to unveil context dependence in ecological interactions.

One further issue is that some interactions may depend on variables that are not observed. For instance, some population dynamics are strongly determined by their demographic state (Coulson et al., 2004; Lande et al., 2002), which would call for time series of the relevant demographic stages. In the system considered here, the dynamics of algae in the rotifer system are most likely coupled with that of nitrogen, for which no time series was available (Hiltunen

et al., 2013). Our method only accounts for observed variables, so that time series for all important variables are required, though unaccounted variables are captured to some extent by non-linear fluctuations in interactions. One interesting prospect would hence be to incorporate unobserved/latent state variables into the NODE system (Dupont et al., 2019; Frank, 2022; Zhang et al., 2019). Careful thought has to be given here as whether to use an ODE or NODE for the latent states given that they are not directly constrained by observations.

A further question is whether we could use similar approaches to analyse systems larger than the ones considered in this study. In particular, microbial communities feature thousands of species, and so potentially millions of interactions, which poses a real problem for inference with dynamical models (Ridenhour et al., 2017). Even the simplest linear ODE model would contain millions of parameters, hence entering deep learning territory. We believe that our success there is more readily limited by the availability of time series sufficiently long to identify this many interactions, rather than by our models. A possible next step could be to address the capacity of more complex NODE and ODE models to estimate interactions in large artificial communities (e.g. Coenen et al., 2020), which could inform us on the relationship between model complexity and data requirements in terms of time-series length and sampling frequency.

We consider NODEs that are only defined along the time dimension. The framework could easily be extended to any other dimension by considering partial differential equations instead (Rackauckas et al., 2019). For instance, in a spatial ecology context we could model the dynamics of populations along two additional spatial dimensions. In an evolutionary context, we could model the dynamics of populations in phenotype space, by adding phenotypic traits as additional dimensions. The BNGM method could be instrumental in fitting these models, which are notoriously expensive to stimulate.

6 | CONCLUSION

We provide a method, BNGM, which allows for NODE fitting in a matter of seconds. This is a crucial step for efficient model selection and uncertainty quantification in NODEs. We also demonstrate that NODEBNGM results in faster, more accurate estimation of the direction, strength, and non-linearity of ecological interactions than existing approaches, in a system where truth is known. Finally, we estimate ecological interactions in real prey predator systems, showing that only stronger interactions seem to be consistent across replicated time series, and that a single species can account for a large part of the changes in community dynamics. Our study allows for efficient NODE fitting, and confirms the power of NODEs in identifying dynamical coupling between populations.

AUTHOR CONTRIBUTIONS

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

ACKNOWLEDGEMENTS

We thank Ben Sheldon for insightful suggestions on early versions of the work. We thank warmly the Ecological and Evolutionary Dynamics Lab and Sheldon Lab Group at the department of Zoology for their feedback and support. We thank Johan Bonnaffé for his help in digitising the algae, flagellate, and rotifer time series. We thank Stella Felsinger for her help in revising the text of this manuscript. The work was supported by the Oxford-Oxitec Graduate Scholarship and the Natural Environment Research Council Doctoral Training Partnership (NERC DTP).

CONFLICT OF INTEREST STATEMENT

The authors have no conflicts of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://www.webofscience.com/api/gateway/wos/peer-review/10.1111/2041-210X.14121.

DATA AVAILABILITY STATEMENT

All data and code are available on GitHub at https://github.com/WillemBonnaffe/NODEBNGM, as well as on Zenodo at https://zenodo.org/record/7825866#.ZDgXh-zMIUE (https://doi.org/10.5281/zenodo.7825866).

ORCID

Willem Bonnaffé https://orcid.org/0000-0002-5053-8891

REFERENCES

Aarts, L. P., & Veer, P. V. D. (2001). Neural network method for solving partial differential equations. *Neural Processing Letters*, *14*(3), 261–271. https://doi.org/10.1023/A:1012784129883

Adams, M. P., Sisson, S. A., Helmstedt, K. J., Baker, C. M., Holden, M. H., Plein, M., Holloway, J., Mengersen, K. L., & McDonald-Madden, E. (2020). Informing management decisions for ecological networks, using dynamic models calibrated to noisy time-series data. *Ecology Letters*, *23*(4), 607–619.

Adamson, M. W., & Morozov, A. Y. (2013). When can we trust our model predictions? Unearthing structural sensitivity in biological systems. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *469*(2149), 1–19.

Arndt, H. (1993). Rotifers as predators on components of the microbial web (bacteria, heterotrophic flagellates, ciliates)—A review. *Hydrobiologia*, *255–256*(1), 231–246. https://doi.org/10.1007/BF00025844

Berryman, A., & Turchin, P. (1997). Detection of density dependence: Comment. *Ecology*, *78*(1), 318–320.

Berryman, A. A. (2002). Population: A central concept for ecology? *Oikos*, *97*(3), 439–442.

Berryman, A. A. (2003). On principles, laws and theory in population ecology. *Oikos*, *103*(3), 695–701.

Bonnaffé, W., Legendre, S., Danet, A., & Edeline, E. (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. *Oikos*, *130*, 1297–1309. https://doi.org/10.1111/oik.08173

Bonnaffé, W., Martin, M., Mugabo, M., Meylan, S., & Galliard, J. F. L. (2018). Ontogenetic trajectories of body coloration reveal its

function as a multicomponent nonsenescent signal. *Ecology and Evolution*, *8*(24), 12299–12307. https://doi.org/10.1002/ece3.4369

Bonnaffé, W., Sheldon, B. C., & Coulson, T. (2021). Neural ordinary differential equations for ecological and evolutionary time series analysis. *Methods in Ecology and Evolution*, *2*, 1–46.

Bonsall, M. B., Meijden, E. V. D., & Crawley, M. J. (2003). Contrasting dynamics in the same plant–herbivore interaction. *Proceedings of the National Academy of Sciences of the United States of America*, *100*(25), 14932–14936.

Brook, B. W., & Bradshaw, C. J. A. (2006). Strength of evidence for density dependence in abundance time series of 1198 species. *Ecology*, *87*(6), 1445–1451.

Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M., & West, G. B. (2004). Toward a metabolic theory of ecology. *Ecology*, *85*(7), 1771–1789. https://doi.org/10.1890/03-9000

Bruijning, M., Jongejans, E., & Turcotte, M. M. (2019). Demographic responses underlying eco-evolutionary dynamics as revealed with inverse modelling. *Journal of Animal Ecology*, *88*(5), 768–779.

Cawley, G. C., & Talbot, N. L. C. (2007). Preventing over-fitting during model selection via Bayesian regularisation of the hyper-parameters. *Journal of Machine Learning Research*, *8*, 841–861.

Cenci, S., Sugihara, G., & Saavedra, S. (2019). Regularized S-map for inference and forecasting with noisy ecological time series. *Methods in Ecology and Evolution*, *10*(5), 650–660.

Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. (2019). Neural ordinary differential equations. *arXiv*, 1–19.

Coenen, A. R., Hu, S. K., Luo, E., Muratore, D., & Weitz, J. S. (2020). A primer for microbiome time-series analysis. *Frontiers in Genetics*, *11*, 310. https://doi.org/10.3389/fgene.2020.00310

Coulson, T., Guinness, F., Pemberton, J., & Clutton-Brock, T. (2004). The demographic consequences of releasing a population of red deer from culling. *Ecology*, *85*(2), 411–422.

Dallas, T., Melbourne, B. A., Legault, G., & Hastings, A. (2021). Initial abundance and stochasticity influence competitive outcome in communities. *Journal of Animal Ecology*, *90*, 1–26.

Deyle, E. R., May, R. M., Munch, S. B., & Sugihara, G. (2015). Tracking and forecasting ecosystem interactions in real time. *Proceedings of the Royal Society B: Biological Sciences*, *283*, 1–9.

Dupont, E., Doucet, A., & Teh, Y. W. (2019). Augmented neural ODEs. *arXiv*, 1–11.

Ellner, S. P., Geber, M. A., & Hairston, N. G. J. (2011). Does rapid evolution matter? Measuring the rate of contemporary evolution and its impacts on ecological dynamics. *Ecology Letters*, *14*(6), 603–614. https://doi.org/10.1111/j.1461-0248.2011.01616.x

Ellner, S. P., Seifu, Y., & Smith, R. H. (2002). Fitting population dynamic models to time-series data by gradient matching. *Ecology*, *83*(8), 2256–2270.

Frank, S. A. (2022). Automatic differentiation and the optimization of differential equation models in biology. *arXiv*, 1–10.

Funahashi, K., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, *6*(6), 801–806.

Goren, M., Gayer, K., & Lazarus, N. (2009). First record of the far east chameleon goby tridentiger trigonocephalus (Gill, 1859) in the Mediterranean Sea. *Aquatic Invasions*, *4*(2), 413–415.

Gross, K., Ives, A. R., & Nordheim, E. V. (2005). Estimating fluctuating vital rates from time-series data: A case study of aphid biocontrol. *Ecology*, *86*(3), 740–752.

Hairston, N. G. J., Ellner, S. P., Geber, M. A., Yoshida, T., & Fox, J. A. (2005). Rapid evolution and the convergence of ecological and evolutionary time. *Ecology Letters*, *8*(10), 1114–1127.

Hiltunen, T., Jones, L. E., Ellner, S. P., & Hairston, N. G. J. (2013). Temporal dynamics of a simple community with intraguild predation: An experimental test. *Ecology*, *94*(4), 773–779. https://doi.org/10.1890/12-0786.1

Hu, P., Yang, W., Zhu, Y., & Hong, L. (2020). Revealing hidden dynamics from time-series data by ODENet. *arXiv*, 1–17.

Ives, A. R., Dennis, B., Cottingham, K. L., & Carpenter, S. R. (2003). Estimating community stability and ecological interactions from time-series data. *Ecological Monographs*, *73*(2), 301–330.

Jonzén, N., Lundberg, P., Ranta, E., & Kaitala, V. (2002). The irreducible uncertainty of the demography–environment interaction in ecology. *Proceedings of the Royal Society B: Biological Sciences*, *269*(1488), 221–225.

Jost, C., & Ellner, S. P. (2000). Testing for predator dependence in predator-prey dynamics: A non-parametric approach. *Proceedings of the Royal Society B: Biological Sciences*, *267*(1453), 1611–1620.

Kendall, B. E., Briggs, C. J., Murdoch, W. W., Turchin, P., Ellner, S. P., McCauley, E., Nisbet, R. M., & Wood, S. N. (1999). Why do populations cycle? A synthesis of statistical and mechanistic modeling approaches. *Ecology*, *80*(6), 1789–1805.

Kendall, B. E., Ellner, S. P., McCauley, E., Wood, S. N., Briggs, C. J., Murdoch, W. W., & Turchin, P. (2005). Population cycles in the pine looper moth: Dynamical tests of mechanistic hypotheses. *Ecological Monographs*, *75*(2), 259–276.

Lande, R., Engen, S., Saether, B.-E., Filli, F., Matthysen, E., & Weimerskirch, H. (2002). Estimating density dependence from population time series using demographic theory and life-history data. *The American Naturalist*, *159*, 321–337.

Lawton, J. H. (1999). Are there general laws in ecology ? *Oikos*, *84*(2), 177–192. https://doi.org/10.2307/3546712

Lingjaerde, O. C., Stenseth, N. C., Kristoffersen, A. B., Smith, R. H., Moe, S. J., Read, J. M., Daniels, S., & Simkiss, K. (2001). Exploring the density-dependent structure of blowfly populations by nonparametric additive modeling. *Ecology*, *82*(9), 2645–2658.

Mai, M., Shattuck, M. D., & O'Hern, C. S. (2016). Reconstruction of ordinary differential equations from time series data. *arXiv*, 1–15.

Moe, S. J., Kristoffersen, A. B., Smith, R. H., & Stenseth, N. C. (2005). From patterns to processes and back: Analysing density-dependent responses to an abiotic stressor by statistical and mechanistic modelling. *Proceedings of the Royal Society B: Biological Sciences*, *272*(1577), 2133–2142.

Mysterud, A., Coulson, T., & Stenseth, N. C. (2002). The role of males in the dynamics of ungulate populations. *Journal of Animal Ecology*, *71*, 907–915. https://doi.org/10.1046/j.1365-2656.2002.00655.x

Novak, M., & Stouffer, D. B. (2021). Geometric complexity and the information-theoretic comparison of functional-response models. *Frontiers in Ecology and Evolution*, *9*. https://doi.org/10.3389/fevo.2021.740362

Odum, E. P., & Barrett, G. W. (1972). Fundamentals of ecology. *The Journal of Wildlife Management*, *36*(4), 1372.

Pasquali, S., & Soresina, C. (2018). Estimation of the mortality rate functions from time series field data in a stage-structured demographic model for Lobesia botrana. *arXiv*, 1–15.

Pearce, T., Leibfried, F., Brintrup, A., Zaki, M., & Neely, A. (2018). Uncertainty in neural networks: Approximately Bayesian ensembling. *arXiv*, 1–10.

Rackauckas, C., Innes, M., Ma, Y., Bettencourt, J., White, L., & Dixit, V. (2019). DiffEqFlux.Jl—A Julia library for neural differential equations. *arXiv*, 1–17.

Ridenhour, B. J., Brooker, S. L., Williams, J. E., van Leuven, J. T., Miller, A. W., Dearing, M. D., & Remien, C. H. (2017). Modeling time-series data from microbial communities. *ISME Journal*, *11*(11), 2526–2537. https://doi.org/10.1038/ismej.2017.107

Rosenbaum, B., Raatz, M., Weithoff, G., Fussmann, G. F., & Gaedke, U. (2019). Estimating parameters from multiple time series of population dynamics using Bayesian inference. *Frontiers in Ecology and Evolution*, *6*(234), 1–14.

Royama, T. (1984). Population dynamics of the spruce budworm Choristoneura Fumiferana. *Ecological Monographs*, *54*(4), 429–462.

Song, C., Ahn, S. V., Rohr, R. P., & Saavedra, S. (2020). Towards a probabilistic understanding about the context-dependency of species interactions. *Trends in Ecology & Evolution*, *35*(5), 384–396.

Song, C., & Saavedra, S. (2021). Bridging parametric and nonparametric measures of species interactions unveils new insights of non-equilibrium dynamics. *Oikos*, *130*(7), 1027–1034. https://doi.org/10.1111/oik.08060

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500.

Treven, L., Wenk, P., Dorfler, F., & Krause, A. (2021). Distributional gradient matching for learning uncertain neural dynamics models. *arXiv*, 1–14.

Turchin, P. (1999). Population regulation: A synthetic view. *Oikos*, *84*(1), 153–159.

Turchin, P. (2001). Does population ecology have general laws? *Oikos*, *94*, 17–26.

Turchin, P., Wood, S. N., Ellner, S. P., Kendall, B. E., Murdoch, W. W., Fischlin, A., Casas, J., McCauley, E., & Briggs, C. J. (2003). Dynamical effects of plant quality and parasitism on population cycles of larch budmoth. *Ecology*, *84*(5), 1207–1214.

Ushio, M., Hsieh, C. H., Masuda, R., Deyle, E. R., Ye, H., Chang, C. W., Sugihara, G., & Kondoh, M. (2018). Fluctuating interaction network and time-varying stability of a natural fish community. *Nature*, *554*(7692), 360–363.

Wood, S. N. (2001). Partially specified ecological models. *Ecological Monographs*, *71*(1), 1–25.

Wu, J., Fukuhara, M., & Takeda, T. (2005). Parameter estimation of an ecological system by a neural network with residual minimization training. *Ecological Modelling*, *189*(3–4), 289–304.

Yoshida, T., Ellner, S. P., Jones, L. E., Bohannan, B. J. M., Lenski, R. E., & Hairston, N. G. J. (2007). Cryptic population dynamics: Rapid evolution masks trophic interactions. *PLoS Biology*, *5*(9), 1868–1879. https://doi.org/10.1371/journal.pbio.0050235

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

Zhang, H., Gao, X., Unterman, J., & Arodz, T. (2019). Approximation capabilities of neural ODEs and invertible residual networks. *arXiv*, 1–11.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

Appendix A. Bayesian regularisation.

Appendix B. Complementary results benchmark analysis.

Appendix C. Complementary results case study 2 replicate A.

Appendix D. Complementary results case study 2 replicate B.

Appendix E. Complementary results case study 2 replicate C.

Appendix F. Complementary results case study 3.

Appendix G. Complementary results case study 4.

How to cite this article: Bonnaffé, W., & Coulson, T. (2023). Fast fitting of neural ordinary differential equations by Bayesian neural gradient matching to infer ecological interactions from time-series data. *Methods in Ecology and Evolution*, *14*, 1543–1563. https://doi.org/10.1111/2041-210X.14121