

RESEARCH ARTICLE

Neural ordinary differential equations for ecological and evolutionary time-series analysis

Willem Bonnafe^{1,2}  | Ben C. Sheldon¹  | Tim Coulson² 

¹Edward Grey Institute of Field Ornithology,
Department of Zoology, Oxford University,
Oxford, UK

²Ecological and Evolutionary Dynamics Lab,
Department of Zoology, Oxford University,
Oxford, UK

Correspondence

Willem Bonnafe

Email: willem.bonnafe@stx.ox.ac.uk

Funding information

NERC DTP scholarship; Oxford-Oxitec
scholarship

Handling Editor: Robert Freckleton

Abstract

1. Inferring the functional shape of ecological and evolutionary processes from time-series data can be challenging because processes are often not describable with simple equations. The dynamical coupling between variables in time series further complicates the identification of equations through model selection as the inference of a given process is contingent on the accurate depiction of all other processes.
2. We present a novel method, neural ordinary differential equations (NODEs), for learning ecological and evolutionary processes from time-series data by modelling dynamical systems as ordinary differential equations and dynamical functions with artificial neural networks (ANNs). Upon successful training, the ANNs converge to functional shapes that best describe the biological processes underlying the dynamics observed, in a way that is robust to mathematical misspecifications of the dynamical model.
3. We demonstrate NODEs in a population dynamic context and show how they can be used to infer ecological interactions, dynamical causation and equilibrium points. We tested NODEs by analysing well-understood hare and lynx time-series data, which revealed that prey-predator oscillations were mainly driven by the interspecific interaction, as well as intraspecific densitydependence, and characterised by a single equilibrium point at the centre of the oscillation.
4. Our approach is applicable to any system that can be modelled with differential equations, and particularly suitable for linking ecological, evolutionary and environmental dynamics where parametric approaches are too challenging to implement, opening new avenues for theoretical and empirical investigations.

KEYWORDS

artificial neural networks, ecological dynamics, evolutionary dynamics, Geber method, neural ordinary differential equations, ordinary differential equations, prey-predator dynamics, time-series analysis

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society

1 | INTRODUCTION

Understanding and predicting dynamics is at the heart ecology and evolution, whether it is to forecast the outcome of evolutionary responses (Norberg et al., 2012), adverse consequences of climate change (Cotto et al., 2017) or to assess the population viability (Pierson et al., 2015). The complexity of biological systems hinders our ability to anticipate how they change. For instance, even a seemingly simple system of two interacting species like the hare–lynx system can lead to intriguing and unpredictable dynamics, such as hares apparently feeding on lynx (Deng, 2018). While incorrectly inferring the drivers of historical hare and lynx population dynamics might have limited consequences, incorrect predictions of environmental threats, such as effects of pests on crops, harvesting of natural resources or epidemics (e.g. Ferguson et al., 2020), may have a greater societal impact. It is hence paramount that we continue improving our capacity to understand and forecast the dynamics of natural systems.

In practice, gaining both understanding and predictability of a given system can be achieved through dynamical modelling of time-series data. The general idea is to formulate a mathematical model that describes changes in the key variables of a given biological system, and to constrain it to behave like the real system via model fitting. The analysis of the dynamical properties of the model then reveal key biological properties of the biological system such as ecological interactions (e.g. Hiltunen et al., 2013), dynamical causation (Sugihara et al., 2012), equilibrium points and stability (e.g. Cortez & Patel, 2017). Dynamical system modelling is broadly divided into parametric and nonparametric dynamical models, depending on the nature of the functions that determine the dynamics of the key variables.

Parametric dynamical models broadly consist of time difference equation systems in discrete time, such as auto-regressive models (e.g. Stenseth et al., 2002), and ordinary differential equation systems (ODEs) in continuous time (e.g. Deng, 2018). The general approach consists in translating hypotheses regarding the dynamics of the system variables into a parametric mathematical model, and then finding parameters that produce dynamic consistent with the behaviour of the system. This approach is limited by our capacity to formulate assumptions and conceptualise the system, which makes it challenging to build a suitable model for a given ecological system, and can be problematic in cases where fast action is required (e.g. dealing with invasive species or algal blooms).

The counterpart of parametric dynamical models is nonparametric dynamical models, primarily consisting of recurrent neural networks (RNNs, e.g. Lek & Gue, 1999). In practice, RNNs are fitted to the time series by predicting iteratively the state of the system at the next time step given the current state and associated covariates. These models have been applied successfully to study ecological dynamics, such as the growth of pine trees (Chon et al., 2000), changes in aquatic benthic communities (Chon et al., 2001) and the dispersal of ungulates (Dalziel et al., 2008). The main disadvantage of these approaches is that they are not embedded in a theoretical

framework and thus provide limited insights into the mechanisms at play.

Neural ordinary differential equations (NODEs), a recent breakthrough in AI research (Chen et al., 2019; Rubanova et al., 2019), offer a promising alternative, as they combine the strengths of both parametric and nonparametric approaches to time-series analysis. NODEs are defined by embedding artificial neural networks (ANNs), which are universal function approximators (Chen & Chen, 1993), in ODE systems, which are powerful mathematical models for dynamical systems, allowing ANNs to approximate key dynamical processes (Chen et al., 2019; Rubanova et al., 2019). Then by training the NODE system to replicate time series of observations of the system state, the ANNs converge to the functional shapes that best describe the underlying dynamical processes (e.g. Wu et al., 2005). NODEs hence provide a novel way to infer hidden dynamical processes from time-series data (Chen et al., 2019; Rubanova et al., 2019), as hypotheses can be specified with parametric ODEs, and unknown processes can be approximated nonparametrically by the ANNs (e.g. Wu et al., 2005).

Hybrid approaches combining ANNs and ODEs have been applied with success in the past across a range of fields. For example, physicists have recently applied approaches combining partial differential equations (PDEs) and ANNs to model fluid dynamics (Sinchev et al., 2018), and to discover governing physical equations from data (Berg & Nyström, 2019). Neuroscientists have modelled the activation dynamics of interacting populations of neurons in an Ape with a NODE-like system, faithfully reproducing the behaviour of motoneurons (Sussillo et al., 2015). Only one study in ecology, to our knowledge, successfully combined ANNs with ODEs to learn functional responses in a prey–predator system, reproducing intriguing non-stationary oscillations (Wu et al., 2005).

It is our view that the potential of NODEs to further our understanding of ecological and evolutionary system dynamic remains largely unexplored, as no study to date has formally introduced the method in an ecological dynamic context and demonstrated which, and how, biological insights could be derived from applying them to real ecological data. For instance, NODEs could be used to learn ecological interactions from time-series data, establish dynamical causality between system variables and determine equilibrium points in a community dynamic context. In doing so, NODEs are not contingent on assumptions and hence offer interesting counterbalance points to results obtained from parametric dynamical models.

Our aim in this manuscript was to introduce the NODE framework and demonstrate how it can be applied to real ecological data, with a focus on learning ecological interactions, dynamical causality and equilibrium points. To do this, we define the mathematical prerequisites and general steps to implement NODE systems for a given time series. Then, we demonstrate how the approach can be used in practice by analysing the dynamics of a di-trophic system, the hare–lynx pelt count time series (Odum & Barrett, 1972). In particular, we highlight how ANNs uncover the shape, direction and strength of ecological interactions in the time series, as well as

dynamical causality and potential equilibrium points. Finally, we discuss these results in light of the biology of the system and highlight further applications and developments of the framework in the context of ecological and evolutionary dynamics.

2 | MATERIALS AND METHODS

2.1 | Method overview

The following sections describe how to define, train and analyse a NODE system in a population dynamic context (an overview is presented in Figure 1). In particular, we use NODEs to recover nonparametric approximations of the per capita growth rate of the different species present in a community, from time-series data. From these nonparametric approximations, we demonstrate how to recover the direction and strength of species interactions, the contribution of each species dynamics to the dynamics of other species, as well as equilibrium points in the state space of the system. We have chosen to demonstrate the method in a population dynamic context as it is a subject of interest to many ecologists, and makes the workings of the method more concrete and accessible. Throughout the following sections, we encourage the reader to keep in mind that NODEs can be used to model the dynamics of other biological variables, beyond population density, such as phenotypic trait and spatial dynamics, among others, making it an extremely versatile method. These other applications are described in the discussion.

2.2 | Defining NODEs

For the sake of simplicity, we consider extending simple ODE population dynamic models into NODE population dynamic models. To start with, we define several key elements of simple ODE population dynamic models: let X denote the density of species X , let r_x denote the per capita growth rate of species X . The dynamics of X in continuous time can then be written as a simple differential equation $dX/dt = r_x X$. The dynamics of any other species Y , for instance a predator of X , can be written as $dY/dt = r_y Y$. Traditionally, we would define per capita growth rates as parametric equations, based on our biological knowledge of the drivers of the changes in the density of these populations. For instance, the Lotka–Volterra model assumes a linear decrease in the growth rate of the prey with predator density, $r_x = a - bY$, and a linear increasing function of prey density for the per capita growth rate of the predator, namely $r_y = cX - d$. This parametric construction captures the direction of the ecological interaction between the two species, which in this case is a trophic interaction benefiting the lynx. The actual direction and strength of the interaction can be obtained by estimating the parameters a , b , c and d from time-series data by fitting the ODE system to the time series of population densities (e.g. Demyanov et al., 2006; Fussmann et al., 2017; Kendall et al., 2005).

The benefit of this approach is that it embeds the ODE population dynamic model in the real world by providing parameters that are consistent with the dynamics of the real system. The issue with this approach is that the values of the parameters, and therefore of the direction and strength of the ecological interaction, are contingent on the choice of mathematical function for the growth rates. In practice, due to the complexity of the processes that underlie the per capita growth rates of species, we have a low chance of choosing an appropriate mathematical form. Furthermore, because all variables are coupled dynamically, misspecifying the per capita growth rate in a single species will impact the dynamics of all others, and potentially bias the inference of all species per capita growth rates. While it is feasible to test different mathematical functions through model selection in a low dimensional system, it becomes rapidly infeasible in larger systems. For instance, a three-species population dynamic model features nine possible pairwise interactions (i.e. density dependencies in per capita growth rates), so that even simply testing for the absence or presence of each interaction requires model selection in a model space of size 81.

This is why we propose to use a NODE system instead (or in addition to the parametric approach). A NODE is a class of ODE in which some functions are inferred nonparametrically with ANNs, rather than parametrically. ANNs should not be viewed as anything more than flexible functions of biological variables (e.g. species density) that can produce any continuous shape, and which can describe any biological quantity (e.g. per capita growth rates). Therefore, while fitting a parametric ODE model provides estimates of parameters, fitting a NODE model provides estimates of entire functions. And instead of looking at specific parameter values, biological processes are studied by looking at the shape of the nonparametric functions, and dynamical behaviour they enforce. The benefit of this approach is that the inference of biological functions that underlie the dynamics of the system, such as per capita growth rates, as well as any inference derived from them are not contingent on our assumptions or on the mathematical functions we might have chosen to portray them.

More particularly in the case of population dynamics, we can turn the simple ODE population dynamic model that we have developed so far into a NODE system by defining the per capita growth rate of the populations as nonparametric functions of the species densities, instead of a parametric one, more formally, $r_x = s(X, Y)$, where s is a nonparametric function of the density of species X and Y . The dynamical equations for species X and Y thus become

$$\begin{aligned} \frac{dX}{dt} &= s1(X, Y)X \\ \frac{dY}{dt} &= s2(X, Y)Y \end{aligned} \quad (1)$$

where $s1$ and $s2$ denote the fact that the per capita growth rates of X and Y are determined by two independent nonparametric functions s . The only biological assumption that needs to be made upon defining a NODE system to infer per capita growth rates is the dependences of the growth rates of the species on the other states of the system.

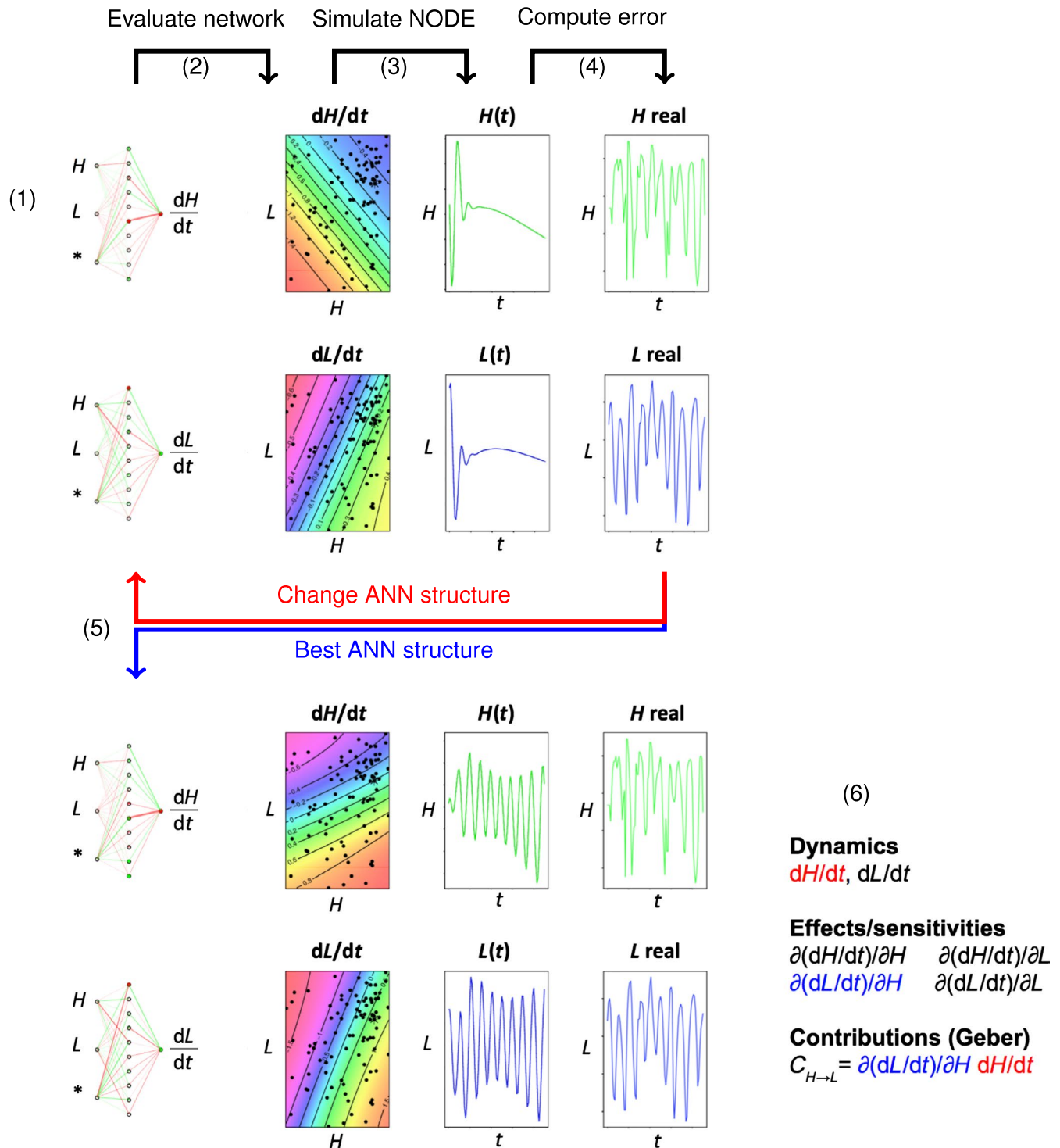


FIGURE 1 Overview of training and analysis of neural ordinary differential equation (NODE) systems. (1) The dynamics of the state variables are defined as neural networks with inputs being the state of the system (e.g. hare H , lynx L , other variables $*$). (2) This imposes a shape for the dynamics of each population as a function of the state of the system. (3) The model can be simulated like any regular ODE to generate predicted time series for the system states. (4) The predictions are compared to observed states assuming an error model (in the present study a Bayesian model). (5) The error is minimised by repeating steps (1) to (4) with different network parameters, at which point the networks will have converged to the shapes that best describe the observed dynamics. (6) Finally, the best NODE system can be analysed to explore state dependence in dynamics (effects), dynamical coupling (contributions) and more

For instance, the construction $r_x = s(X)$ would imply that the per capita growth rate r_x may only be subjected to intraspecific density dependence, and may not be affected by changes in the density of

species Y . In practice, this can be done simply by specifying manually which variables should be considered as inputs in the nonparametric functions s .

The second step is to define the properties of the nonparametric functions s , that is the specifications of the ANNs. In this paper, we consider a specific class of ANNs, called single layer perceptrons (SLPs), which is a widespread and mathematically convenient choice due to their simplicity (e.g. Aarts & Van Der Veer, 2001; Funahashi & Nakamura, 1993; Mai et al., 2016; Sinchev et al., 2018; Wu et al., 2005). They are the most simple form of ANNs as they correspond to a weighted sum of basis functions f , more formally

$$s(X, Y, \dots) = \beta_0 + \beta_1 f(\beta_{10} + \beta_{11}X + \beta_{12}Y + \dots) + \beta_2 f(\beta_{20} + \beta_{21}X + \beta_{22}Y + \dots) + \dots + \beta_i f(\beta_{i0} + \beta_{i1}X + \beta_{i2}Y + \dots), \quad (2)$$

where β_{ij} are the parameters of the network.

The first key element of the networks that we need to define is the basis functions f . There is strong incentive for using sigmoid basis function given that the capacity of SLPs with sigmoid basis function to reproduce the shape of any continuous function has been established theoretically (Chen & Chen, 1993). This property has been extended to NODEs relying on SLPs with sigmoid basis functions which have been shown to be able to reproduce the behaviour of any dynamical system with continuously defined differential operators (Funahashi & Nakamura, 1993). In practice, however, certain basis functions are better at producing specific shapes. For instance, Gaussian functions are suitable for producing Gaussian-like shapes, with multiple peaks and valleys, which may be easier to train if the phase space is thought to be bell shaped. The identity basis functions simplify the SLPs back to a linear model. In a population dynamic context, this implies that there is no context dependence (or nonlinearities) in the pairwise interactions between species. It is hence possible to test the robustness of the results to the implementation of context dependence in the effects of variables by switching the basis function from sigmoid to identity.

The second and final network element that needs to be specified is the number of intermediate nodes, namely the number of times the basis functions are summed in Equation (2). There is a strong theoretical grounding for choosing a number of nodes that is at least twice as large as the number of input variables in the network to ensure that any continuous shape can be produced by the neural network (Chen & Chen, 1993). For instance, if we modelled the per capita growth rate of the lynx r_y as a function of hare and lynx density X and Y , namely $r_y = s(X, Y)$, we would need at least $n = 4$ nodes in the network. In practice, this number should also be as low as possible as it would otherwise complicate the training of the neural network (e.g. Wu et al., 2005). It is common practice to use 10 nodes, which is suitable for small ODE systems (e.g. Wu et al., 2005). In practice, there is no need to manipulate the number of nodes to change the complexity of the shape of the network as this is handled by the parameters of the networks.

2.3 | NODE fitting

Fitting NODE systems to time series involves the same steps that are required for ODE systems, namely defining and optimising an error function. However, they require a couple of extra steps to account for the fact that NODEs are prone to overfitting and getting trapped in local maxima, due to being larger than parametric ODEs by at least an order of magnitude.

The first step is to define a simple Bayesian model to compute the posterior probability density of the population density predicted by the NODE system given the population densities in the time series, which can be written as $p(X, Y | x, y) \propto p(x, y | X, Y) p(\beta)$, where x and y are the observed population densities, X and Y the population densities predicted by the NODE system, obtained by simulating the NODE system with an ODE solver and β are the parameters of the neural networks. For the sake of simplicity, we assume normal distributions for the residuals of model predictions with a dispersion parameter σ_1 that we use to control the precision of the fit, more formally, $e_t \sim \mathcal{N}(0, \sigma_1^2)$, iid. Given the flexibility of these models, we treat the dispersion σ_1 as known to dictate how closely the model should fit the data and therefore reduce the chance of overfitting. In addition, we define normal prior distributions for the parameters of the network with a mean of 0 and dispersion σ_2 that we use to control the complexity of the network, $\beta_{ij} \sim \mathcal{N}(0, \sigma_2)$, iid. The idea is that low values of dispersion σ_2 constrain the parameters of the networks to be close to 0, which also constrains the shape of the neural network to be simple and linear. In contrast, increasing the dispersion of network parameters allows for more complex, nonlinear shapes for the network. This provides a way of limiting the risk of overfitting, while assessing the robustness of results and dynamics to increasing levels of complexity/nonlinearity in the shape of the population growth rates.

The second step consists in fitting the NODE system, which can be done by finding the network parameters (and thereby the shape of the functions) that maximise the posterior distribution. In this manuscript, we rely on the Broyden–Fletcher–Goldfarb–Shanno (BFGS) gradient descent method to optimise the posterior distribution. There exist other techniques to fit these models which we discuss in the code implementation section.

The third and final step, which is not necessary if we are only interested in finding the best fit, is to quantify the uncertainty around the NODE predictions given the choice of the dispersion parameter for the likelihood and priors. In this manuscript, we use a rejection sampling scheme of the posterior distribution to recover approximated posterior distributions of the network parameters (Gelman et al., 2013), and thereby of the shapes of the per capita growth rate and biological quantities that we derive from them. Samples from the posterior distribution are generated by randomly drawing parameters according to a known probability distribution, centred around the best identified parameter vector and discarded with a probability proportional to their posterior density (Gelman et al., 2013).

2.4 | NODE analysis

Once the fitting of the NODE system to the time series is completed, we can analyse the model. The general NODE model that we have developed in the context of population dynamic returns nonparametric approximations of the per capita growth rates of each species, as a function of the density of other species. Analysing the model comes down to analysing the per capita growth rate functions of each species, and studying the dynamical behaviour that they impose, such as oscillations and equilibrium points.

2.4.1 | Inferring direction and strength of ecological interactions

Per capita growth rate functions contain essential information regarding the ecological interactions between the species of the system. First, the direction of the interaction between the species can be approached by studying how its per capita growth rate changes with the density of other species. This information is captured by the slope, or sensitivity, of the per capita growth rate with respect to the density of each species, noted $\partial r_Y / \partial X$ where r_Y is the per capita growth rate of species Y , and X is the density of species X . For instance, considering the hare-lynx system, $\partial r_l / \partial H = 0.5$ would mean that the per capita growth rate of the lynx (r_l) increases by 0.5 per unit of hare density (H). In biological terms, this means that hare density tends to increase the per capita growth rate of the lynx, and therefore that a higher prey density improves the predator population growth rate.

2.4.2 | Dynamical coupling between variables

Computing effects alone is necessary but not sufficient to establish whether two variables are coupled dynamically, that is whether changes observed in one variable were caused by changes in the other. For example, the growth rate of a prey may be affected negatively by predator density, but the actual contribution of the predator population to the dynamics of the prey can either be positive, zero or negative depending on whether predator density is decreasing, remaining constant or increasing. In the first case, the prey population would grow as a result of the alleviation of predation pressure. It is thus important to consider contributions in addition to effects/sensitivities. Contributions of the species X to the dynamics of any species Y can be computed following the Geber method (Hairston et al., 2005), which amounts to multiplying the dynamics of X by the effect of X on the growth rate of Y , more formally $C: X \rightarrow Y = dX/dt \times \partial r_Y / \partial X$. The total contribution of a variable over a time period can be computed by summing the absolute value of its contribution at each time step across all time steps. This value gives an idea of how important a given variable is for the dynamics of the focal population.

2.4.3 | Finding equilibrium points

In addition to effects and contributions, it is also possible to derive equilibrium points from the analysis of the nonparametric approximations of the population growth rates. Equilibrium points correspond to points in state space (i.e. particular combinations of population densities) at which the growth rate of the populations are null. Equilibrium points can be found numerically by evaluating the dynamics of the system for different combinations of population densities, just like any regular non-analytically tractable ODE systems. This last application of NODEs can prove particularly useful for managing natural populations, as equilibrium points found are not subjected to potential misspecifications of the mathematical structure of the model.

2.5 | Code implementation

There are two well-established and tested repositories for NODEs in Python (Chen et al., 2019) and Julia (Rackauckas et al., 2019). In both cases, the repositories provide state of the art fitting algorithms, such as the adjoint sensitivity method, and methods for analysing NODEs. However, these repositories do not provide options for performing Bayesian inference, and there are no implementations directly available to recover the biological quantities of interest discussed in the previous section. We hence provide an R implementation of the NODE population dynamic models presented in the previous sections. It is our view that an R implementation would be more useful to many ecologists given the predominance of the R language in the field. Furthermore, we coded the repository with a focus on simplicity and clarity rather than computational efficiency to facilitate and encourage customisations of the method to pursue other biological aims beyond the ones presented here. The code is modular as well and can be applied to any time series, be it formatted as a table with time steps in the first column, and handles automatically the specifications of the networks and of the fitting.

The steps to implement a NODE for a given time series in the R language are as follows: (a) format and standardise the time series (TS.format function), (b) create a NODE object list containing the dependence of each variable on the variables of the system (NODE.init function), (c) define the dispersion parameters of the likelihood and prior functions σ_1 and σ_2 (function error.normalBayesian), (d) optimise the posterior distribution for a specified number of iterations and dispersion parameters, based on the *ode* function in the *odeSolve* package for simulation and *optim* function that performs BFGS gradient-descent optimisation, which is not as efficient as the fitting methods in the Python and Julia repositories, but is still robust and commonly used for fitting ODEs and ANNs, (e) quantify uncertainty around model predictions through rejection sampling of the posterior distribution, which is an approximate Bayesian computational (ABC) method that provides approximations of the error around model estimates (function .ABC) and (f) extract and visualise

effects, contributions, equilibrium points with the various functions implemented in the repository.

The code is provided in full for the case study on a GitHub repository referenced in the data accessibility section. The implementation we provide also allows for the use a simple LASSO/ridge least square regression of the NODE system, in case the Bayesian implementation is deemed unnecessary. In addition to that, we provide supplementary files on the GitHub repository containing (a) a minimalistic implementation of a two-variable NODE system, geared towards computational efficiency and fine customisation, as well as (b) an implementation of a two-variable NODE system with a time dependence embedded to model time-varying interactions between variables.

2.6 | Case study: di-trophic prey–predator dynamics

2.6.1 | System

In this section, we provide an example of NODE implementation to study density dependence in prey–predator cycles, using the well-known hare–lynx system as a model (Krebs et al., 2006), thus facilitating the assessment of the consistency of our results with the literature. Hare and lynx display clear 10-year long population cycles, which are driven by a combination of phase dependence in predation and negative density dependence in lynx (Krebs et al., 2006; Slough & Mowat, 1996; Stenseth et al., 2002). Our aim is to use a NODE system to learn the per capita growth rate of the prey and predator from time series of lynx and hare density. Then, by analysing the system, we quantify intra- and interspecific density dependence in growth rates, and through the Geber method (Hairston et al., 2005), establish the respective contribution of intra- and interspecific interactions to the demography of each population. The code, and additional explanations, can be found on a GitHub repository referenced in the data accessibility section.

We analyse a 90-year long time series of hare and lynx pelts (in tens of thousands), collected by trappers in the Hudson Bay area, from 1845 to 1935 (Figure 4 first row, a and b), which displays characteristic oscillations, each lasting 10 years on average (Odum & Barrett, 1972). We chose this time series as it has been studied extensively, allowing for the most sensible comparisons between the results of our case study and the literature (Krebs et al., 2006; Stenseth et al., 2002).

2.7 | Model specifications

We define a NODE system in which changes in hare and lynx densities are governed by a differential equation, and their per capita growth rate is approximated by an ANN.

$$\begin{aligned} \frac{dH}{dt} &= r_H(H, L, \Omega_H)H \\ \frac{dL}{dt} &= r_L(H, L, \Omega_L)L, \end{aligned} \quad (3)$$

where r_H and r_L denote the per capita population growth rate of the hare and lynx population, as a function of hare and lynx density H and L , respectively, and Ω_H and Ω_L are the network parameters governing the shape of the growth rates. In practice, it is recommended to model the log of the density H and L for this allows to: (a) write the model in a simpler form, namely $d(\log(H))/dt = r_H$, and $d(\log(L))/dt = r_L$, (b) model the posterior distribution of the log density as a simple Bayesian model with a Gaussian likelihood and (c) normalise the distribution of observed densities which facilitates the training of the SLPs.

The per capita growth rates are approximated with SLP functions (Equation 2).

$$\begin{aligned} r_H(H, L, \Omega_H) &= \beta_0 + \beta_1 f(\beta_{10} + \beta_{11}H + \beta_{12}L) \\ &+ \beta_2 f(\beta_{20} + \beta_{21}H + \beta_{22}L) + \dots \\ &+ \beta_l f(\beta_{l0} + \beta_{l1}H + \beta_{l2}L), \quad l = 10, \end{aligned} \quad (4)$$

of the hare and lynx densities H and L , weight vector Ω_H and Ω_L containing the network parameters β , and where the number of intermediate neurons is $l = 10$. The activation function is chosen to be sigmoid with form $f(x) = 1/(1 + \exp(-x))$. The value of 10 hidden neurons was chosen by trial and error; it should be well over the dimension of the input (Chen & Chen, 1993), but as low as possible (Wu et al., 2005).

2.8 | Model fitting

The model was fitted to the time series following the approach described in the model fitting section of the methods (Section 2.3). The fitting consisted in finding the parameters (and thereby the functional shape) for the nonparametric per capita growth rates that maximise the posterior density of the population densities predicted by the dynamical system given the population densities observed in the time series and the prior distribution of the network parameters, more formally $p(H, L | h, l) \propto p(h, l | H, L)p(\Omega)$. We also compared model fits with different prior specifications, between $\sigma_2 = 0.1$ and $\sigma_2 = 0.3$, the former enforcing more simple linear functional shapes and the latter allowing for more complex nonlinear relationships between the species.

2.9 | Model analysis

Once this process was complete, the shape of nonparametric per capita growth rates was analysed to determine the direction and strength of interactions between the two species, respective contribution of each species to the dynamics of the other, and candidate equilibrium points where both populations growth rates are 0. In particular, pairwise interactions between species were recovered by computing the sensitivity of the per capita growth rates with respect to the other species density, for instance the effect of a change in hare density on the lynx growth rate was

computed as Effect: $H \rightarrow L = \partial r_L / \partial H$. In turn, the contribution of the prey population dynamics to the dynamics of the predator population was computed by multiplying the dynamics of the prey population by its effect on the predator per capita growth rate, Contribution: $H \rightarrow L = \partial r_L / \partial H \times dH/dt$ (Hairston et al., 2005). We computed the average effect of the prey population on predator dynamics by averaging the sensitivity of the predator per capita

growth rate across the entire time series. The mean strength of the coupling between the dynamics of the prey and predator can be obtained by averaging the absolute value (or the square) of each pairwise contribution across the time series. Finally, we identified equilibrium points in the state space of the system by numerically solving the NODE system, that is finding H^* and L^* for which $dH/dt = dL/dt = 0$.

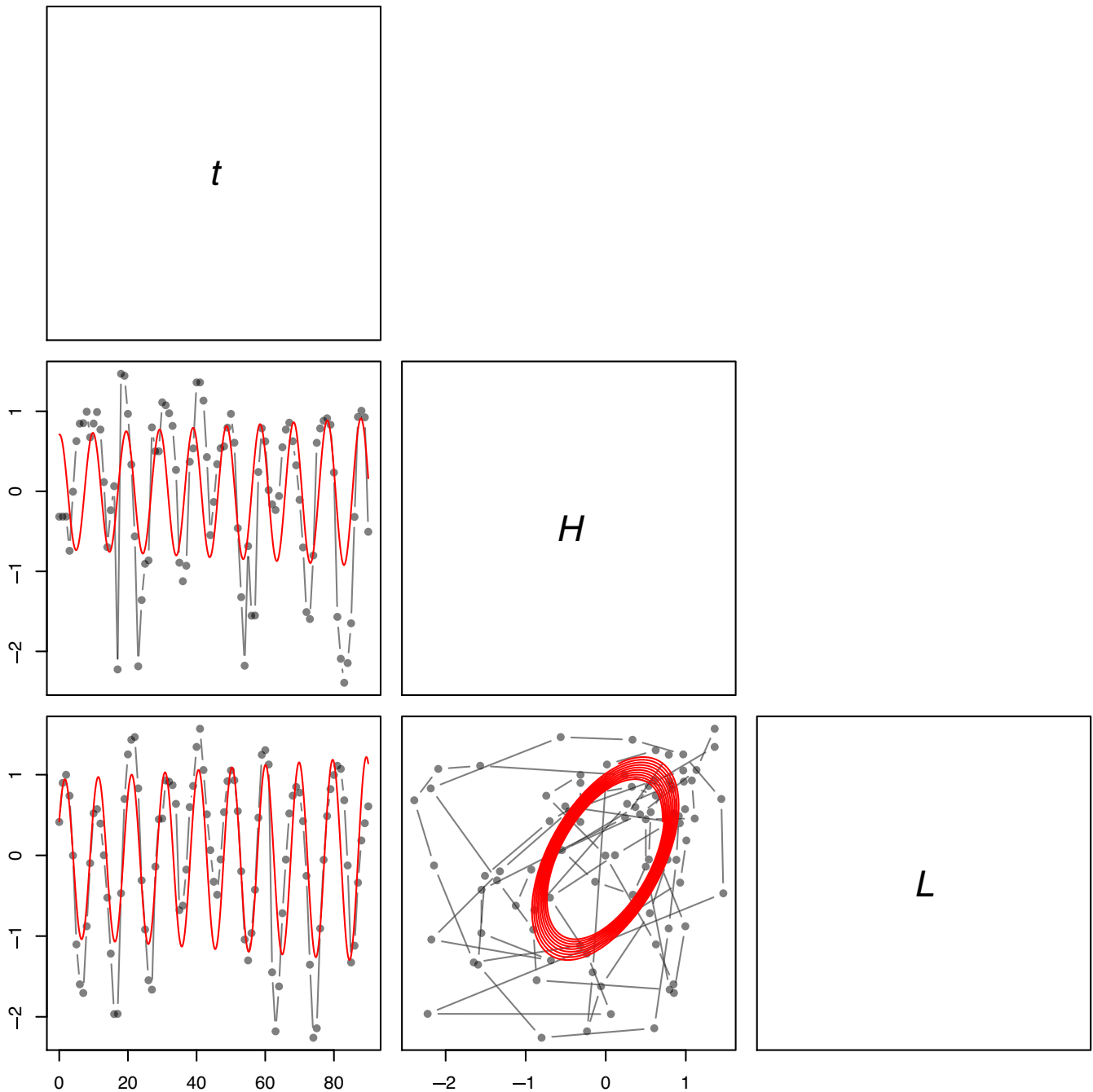


FIGURE 2 Fit of the hare-lynx neural ordinary differential equation (NODE) system to hare-lynx pelt count time series (Odum & Barrett, 1972). The three state variables, that is time t , hare and lynx density H and L , are presented on the diagonal. The grey dots correspond to the number of either hare or lynx pelt counts reported for a given year. The red line corresponds to the number of pelt counts predicted by the best NODE system identified. The hare and lynx density were log transformed and standardised to prevent non-real numbers for densities and improve fitting performance

3 | RESULTS

3.1 | Model fit

The fit of the best NODE system to the hare-lynx pelt count time series is presented in Figure 2. We were not able to obtain a fit for values of parameter prior dispersion below $\sigma_2 = 0.15$. The NODE system does well at capturing the main component of the 10-year oscillations and shows a slight trend for an increase in the amplitude of the oscillations towards the end of the time series. The fit yields a value of r -squared around 0.42, and the analysis of the residual autocorrelation demonstrates that the model does not capture faster oscillatory components of the dynamics (<10 years). The increase in the prior dispersion parameter (up to 0.3), which enables greater nonlinearity in the growth rates, did not improve the fit of the model. In addition, we find very narrow error bars around model predictions, estimated from rejection sampling, which means that estimated dynamics, effects and contributions are highly identifiable for a given fit.

3.2 | Dynamical interaction network and phase space

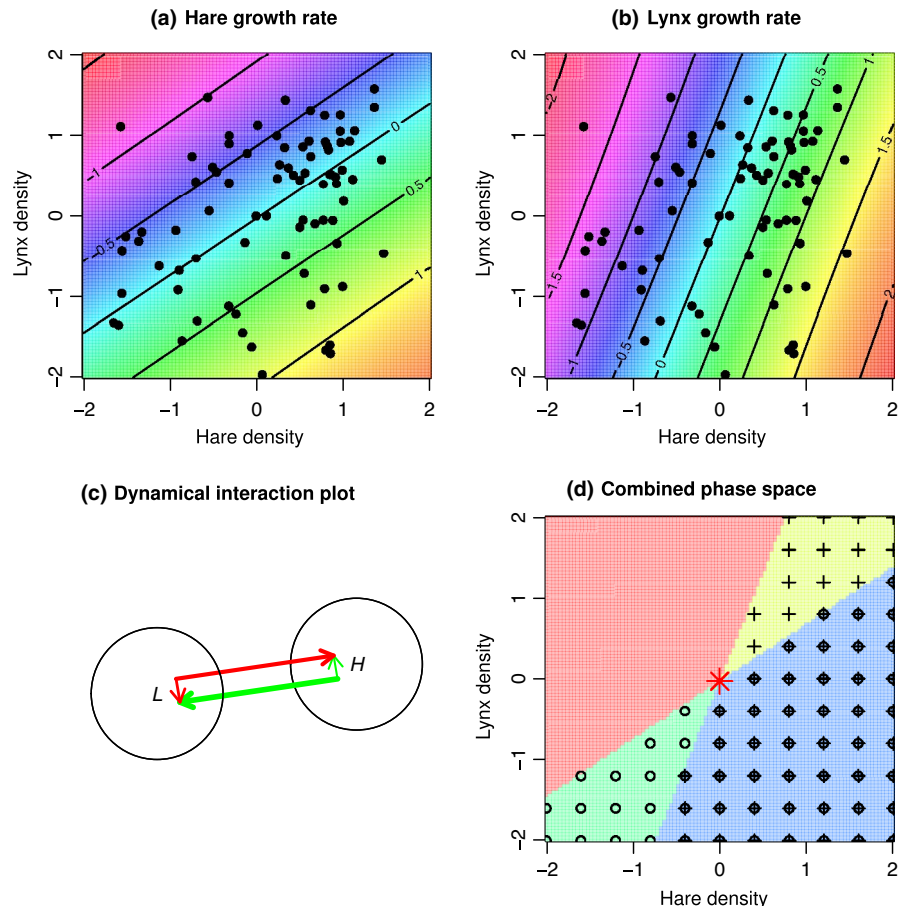
Figure 3 displays the shape of the hare and lynx growth rate (a and b) learned by the neural networks. These reveal a decrease in the

hare growth rate with lynx density (Figure 3a), and an increase with hare density, indicating positive intraspecific density dependence. The lynx growth rate (Figure 3b) increases with hare density and decreases with lynx density, indicating negative intraspecific density dependence. Overall, there is no strong evidence of nonlinearity in the shape of the growth rates. The combination of the two growth rate landscapes gives the combined phase space in Figure 3d, which reveals that the oscillations can be divided in four stages, where the two dominating stages correspond to areas where both populations are either growing or declining (Figure 3d, red and blue area), which indicate tightly coupled oscillations. We found a unique equilibrium point in that system centred around the mean population densities (0 corresponds to the mean of the log densities). Overall, these results are consistent with a prey-predator interaction with negative density dependence in the predator and positive density dependence in the prey (Figure 3c).

3.3 | Effects and contributions

These results are further confirmed when dissecting the effects and contributions underlying the time series (Figure 4). We find that the prey growth rate is affected positively by hare density and negatively by lynx density (Figure 4a middle row). As for the lynx, we find a negative effect of lynx density on lynx growth, and a positive effect

FIGURE 3 Neural ordinary differential equations (NODE) approximation of the population growth rates. Graphs a and b correspond to the per capita growth rates of the hare and lynx populations, respectively, as a function of hare and lynx density H and L . The black dots correspond to the data points in the time series. Cold colours (purple to blue) and contour lines below 0 indicate areas where the populations are declining. Graph c. corresponds to the dynamical interaction network obtained by computing the sensitivity of each population growth rate to hare and lynx density, red and green arrows indicate negative and positive effects respectively. Graph d. shows areas where both populations decline (red area), both grow (blue area, crossed circles), only the hare population grows (green area with empty dots) or only the lynx population grows (yellow area with crosses). The red star indicates the equilibrium point in the system (found by numerically solving the NODE system)



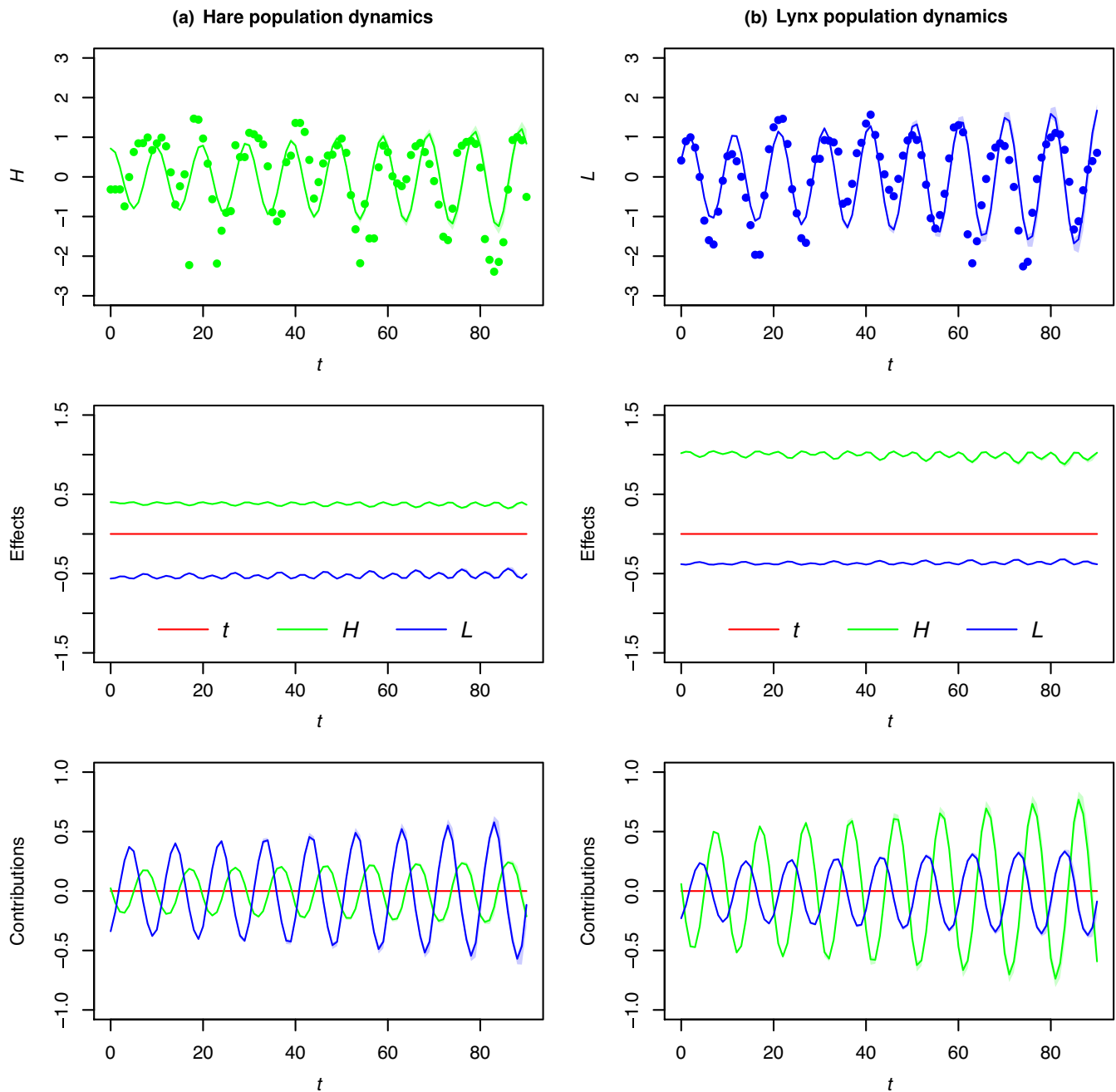


FIGURE 4 Effects and contributions derived from the neural ordinary differential equation (NODE) system fitted to the hare–lynx time series. Graphs on the first row correspond to hare (a) and lynx (b) population densities either observed (dots) or predicted by the best NODE system (solid line). Graphs on the second row display the effects/sensitivity of population growth rates to changes in hare and lynx density (i.e. the Jacobian elements of the networks in the NODE system). The graphs in the last row correspond to the contribution of these effects to hare and lynx population dynamics, respectively, computed by multiplying the dynamics of state variables (first row) with their effect (second row). In all graphs, the red lines indicate 0, and the shaded areas (very narrow) give the error around estimates derived from the model (light shading: 5%–95%, dark shading: 25%–75%, obtained from rejection sampling)

of hare density (Figure 4b middle row). Overall, we find very limited variation of the effects throughout the time series (Figure 4a,b middle row). The contributions reveal that the main contribution to the dynamics of the prey is that of the predator (Figure 4a last row), similarly the prey dynamics is the main contributor to the growth rate of the predator (Figure 4b last row). This indicates that the interspecific interaction is the main driver of the dynamics of both populations.

4 | DISCUSSION

The aim of this paper is to introduce the mathematical framework of NODEs to ecologists and evolutionary biologists and to illustrate how it can be used to learn ecological and evolutionary dynamical processes from time-series data. We present the general framework and use it to identify the drivers of prey–predator

oscillations. In particular, we infer per capita growth rates in a time series of hare and lynx pelt counts from the Hudson Bay Company records (Odum & Barrett, 1972). We were able to quantify precisely the intra- and interspecific drivers of the dynamics of each population, revealing that the interspecific interaction is the main driver of the oscillations, with no clear evidence for nonlinearity in the relationship. These results demonstrate that the approach can provide valuable biological insights into the drivers of population dynamics, in a real setting. We discuss below the biological relevance of our results in light of the literature. We conclude by comparing this novel approach to existing methods for time-series analysis, and highlight further applications to study ecological and evolutionary time series and connections with existing empirical and theoretical frameworks.

4.1 | Hare-lynx case study

Overall, we find that the di-trophic NODE system considered in the present work fits reasonably well the hare-lynx pelt count time series, in spite of residual oscillatory components. We find that increasing the nonlinearity of the hare and lynx growth did not improve the fit of the model, contrary to the general belief that neural networks can easily overfit. This strongly suggests that the residual variation is attributable to state variables that are not modelled in the present system, such as climatic variables (Holyoak & Heath, 2016), fluctuations in hunting pressures (Deng, 2018) or age structure (Stenseth et al., 2002), and that increasing the complexity of the functional relationship between the hare and lynx population alone is insufficient to explain those residual dynamics. If nothing else, our study demonstrates that there seems to be a limit to the amount of variation that can be explained with di-trophic interactions as it accounts for up to 40% of the variation in the dynamics of the system, no matter how complicated the driving functions are.

We found that the interspecific interaction was the main driver of the dynamics in that system, but that intraspecific density dependence, positive in hare and negative in lynx, also contributed substantially. This is consistent with empirical and experimental studies that ruled out the influence of adverse interactions with conspecifics on hare population growth, partly in line with low food constraints (Krebs et al., 2006; Stenseth et al., 2002). Negative density dependence was also found in boreal lynx populations, as individuals tend to be more territorial before the population peak (Slough & Mowat, 1996). Overall, our results demonstrate that NODEs can provide sensible biological insights in light of the biology of the system.

4.2 | Comparison with related approaches

Traditionally, ecological dynamical systems such as the hare-lynx system are analysed with parametric dynamical models, such as ODE models in continuous time (e.g. Deng, 2018) and auto-regressive

models (or difference equation) in discrete time (e.g. Stenseth et al., 2002), which rely on specific mathematical constructions. The NODE approach provides nonparametric evidence for intra- and interspecific effects in this system, which are thereby robust to assumptions about the nature of the ecological interaction between the hare and lynx population. More generally, parametric models require (a) a strong knowledge to formulate assumptions regarding the dynamical processes, (b) advanced mathematical skills to translate these assumptions into parametric equations and (c) a considerable amount of model selection to find a suitable model, which becomes impractical with increasing process complexity. NODEs address each of these points, as ANNs (a) require a limited amount of assumptions, which merely consists in the specification of inputs, (b) emulate infinitely many different process shapes and (c) browse through them continuously, regardless of the complexity of the process. The main benefit of NODEs is hence that they permit the user to infer dynamical processes where complexity and lack of prior knowledge makes a classical parametric approach impractical, and otherwise provides first-hand information to guide the design and selection of parametric equations. In particular, NODEs provide a powerful alternative to parametric models when it comes to the identification and categorisation of equilibrium points and steady-state behaviours empirically informed by time series.

In the context of the NODE literature, our study provides the first successful assessment of the suitability of NODEs to analyse real ecological time series and derive meaningful biological insights. We acknowledge that NODE models have been applied successfully in the past to model prey-predator dynamics, especially in the field of artificial intelligence (e.g. Mai et al., 2016; Massaroli et al., 2020; Rackauckas et al., 2019). This work was instrumental in establishing the workings and validity of the approach, and we also provide one such proof of validity in the Supporting Information of the present work by analysing a simulated time series with known parameters and find an agreement between the true functions and the functions learned by the networks. However, none of the studies above considered real biological systems, and proposed steps to extract biological quantities from those, so that much work remains to be done to find real-world applications for NODEs. We view our study as a step in this direction as we consider an iconic ecological real time series, and provide steps to derive biologically sensitive quantities such as ecological interactions, dynamical coupling and equilibrium points. We further argue that the range of applications of NODEs exceeds largely than those considered in this manuscript.

We view NODEs as beneficial when compared to pure non-parametric dynamical system approaches, such as RNNs (e.g. Chon et al., 2001; Matsunaga et al., 2013; Obach et al., 2001), which rely on single ANNs to approximate the dynamics of the entire system. In doing so, RNNs are often criticised for their black box nature which greatly reduces interpretability of the model, and its applicability as an explanatory tool, making them more suitable for forecasting. NODEs circumvent this problem by embedding ANNs in a consistent theoretical framework outlined by the parametric ODE equations, making ANNs more interpretable and meaningful,

which is an idea that preceded the birth of NODEs (Aarts & Van Der Veer, 2001). For instance, if we can assume that the life cycle of a given species is well known, we can model it in the NODE with specific life stage variables, use parametric functions for known transitions, and neural networks to learn unknown transition functions. By doing so, we progressively dissect ANNs into possible mechanisms, thus opening the black box (Aarts & Van Der Veer, 2001).

We view that our work also improves over closely related non-parametric dynamical methods in ecology, such as semi-specified models (Wood, 2001), gradient matching (Ellner et al., 2002) and residual minimisation training (Wu et al., 2005). These studies have in common the idea of embedding nonparametric functions, such as splines and neural networks, into dynamical systems. However, the work remained limited to inferring specific functions (e.g. density dependence in recruitment rates, Ellner et al., 2002), and did not consider system-level properties, such as ecological interaction networks or equilibrium points, as the aim of the inference. Furthermore, each study provided their own mathematical constructions and way of fitting these models, which hindered uptake of the methodology. Here we propose to rely on the recently discovered framework of NODEs to widen the scope considered by these pioneering studies, and cast these models in an integrated Bayesian framework.

4.3 | Further applications to ecological and evolutionary dynamics

The NODE system in our case study can be used to identify the drivers of the demography of populations from time series (Figure 5, 1a,b), and is readily applicable to a range of other ecological interactions, whether it is to quantify competitive interactions (Gamelon et al., 2019) or trophic interactions (Hiltunen et al., 2013). This can be achieved by modelling changes in population densities with differential equations (Figure 5, 1c), and by approximating demographic processes, such as per capita growth rates, birth or death rates, consumption rates, with ANNs (Figure 5, 1d). Then by analysing the system, we can establish the networks of ecological interactions and their contribution to the dynamics of each species in the community. In doing so, NODEs also provide a bridge between empirical systems and theoretical ecological interactions frameworks, such as that of Press Perturbation (Montoya et al., 2009) or classic trophic ODE models (Hiltunen et al., 2013).

Similarly, NODEs can be applied to identify the drivers of evolution in natural systems (Figure 5, 2a) from time series of the traits, demography and environmental variables (Figure 5, 2b), such as those available for the Darwin's Finches system (Grant & Grant, 1993). This can be done by defining a set of differential equations describing changes in phenotypic, demographic and environmental states (Figure 5, 2c), for instance by merging demographic models with the Price equation (Ellner et al., 2011; Lion, 2018), and by approximating the fitness landscape and fitness gradient with ANNs (Figure 5, 2d). Following an application of the Geber method to both the fitness

landscape and fitness gradient, we could derive the contribution of ecology and evolution to the dynamics of the phenotype and populations (Hairston et al., 2005). In this way, NODEs provide an 'equation-free' (i.e. nonparametric) way of applying the Geber method, and of testing theoretical frameworks for eco-evolutionary feedbacks in real time series, which would benefit many existing studies in this area (e.g. Hairston et al., 2005; Lion, 2018; Patel et al., 2018; Pigeon et al., 2017; Van Velzen & Gaedke, 2017).

NODEs can also be used to learn the drivers of spatial dynamics of individuals and populations (Figure 5, 3a) from records of paths and trajectories (e.g. foraging, dispersal, migrations; Figure 5, 3b). This can be achieved by modelling the changes in the spatial position of individuals, or populations with differential equations (Figure 5, 3c), and approximating the movement kernels, that is the probability of a given move depending on a set of cues, with ANNs (Figure 5, 3d). We view this last application of NODEs as particularly useful, given that we often lack a consistent mathematical framework to explain animal movement (Dalziel et al., 2008), and would lead to novel insights if combined with theoretical frameworks for shoaling and flocking behaviours, or flying predators (e.g. Brighton et al., 2017; Corcoran & Conner, 2016; Olberg, 2012).

4.4 | Limits

The main limitation of the NODE approach is that they are hard to fit, due to a high number of network parameters and hence high computational costs. NODE and ODE systems in general need to be simulated for each set of parameters in order to compute the error with respect to the observations, contrary to linear models, or ANNs, which can be evaluated at all points in the time series simultaneously. Simulating the model through an ODE solver prevents the computation of gradients of the error with respect to each model parameter, which further prevents the use of standard gradient descent methods. That, in addition to the high number of network parameters, overall makes the training of these models challenging and computationally intensive (Chen et al., 2019). There have been steps taken to solve this issue, for instance by relying on highly sophisticated techniques such as automatic differentiation (Chen et al., 2019) and the adjoint sensitivity method (Rackauckas et al., 2019), but it is our view that the training of these models remains nonetheless intensive.

5 | CONCLUSIONS

NODEs combine the explanatory power of dynamical mathematical models and the learning capacity of ANNs, giving them the potential to approximate the dynamics of any biological system based on simple time-series data. This framework has not been applied yet, its potential to further our understanding of system dynamics anywhere theory is missing is unexplored. We show that NODEs can be applied successfully to analyse time series in ecology, using

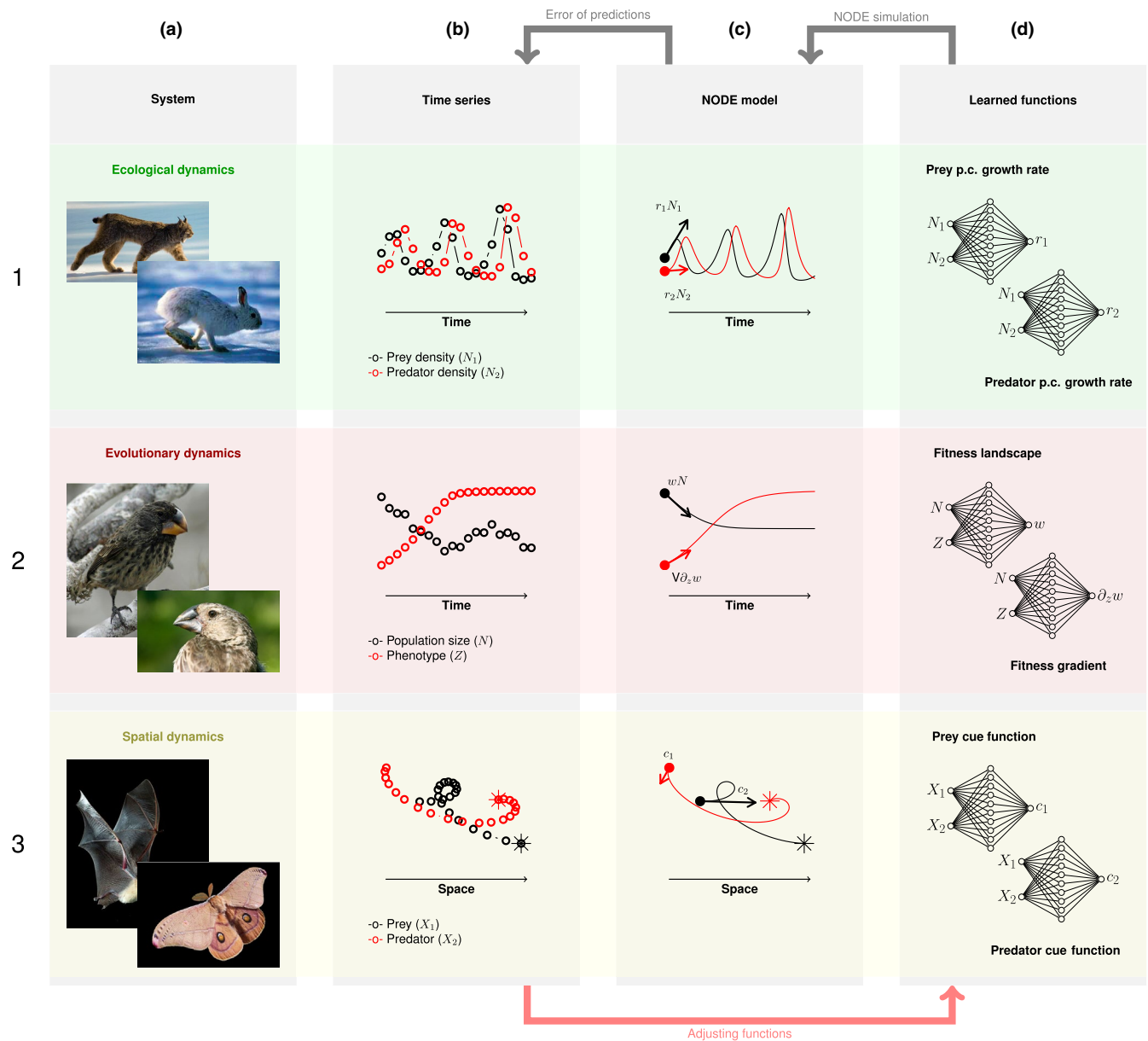


FIGURE 5 Possible applications of neural ordinary differential equations (NODEs) to ecological, evolutionary and spatial dynamics. NODEs can be used to model a wide range of biological functions and systems, including (1) population dynamics, to learn ecological interaction networks beyond the hare–lynx system considered here, (2) evolutionary dynamics, where they could be used to learn fitness landscapes and gradients from time series of phenotypic change and population density, for example and (3) in a spatial dynamic context, where they can be used to learn individual cue functions by analysing continuous time spatial trajectories, such as those of flying predators like bats, birds of prey, dragonflies

the hare–lynx prey–predator system as a case study. We further provided several directions for addressing key questions in ecological, spatial and evolutionary dynamics, with NODEs, highlighting its generalisability and its potential to generate new insights in ecology and evolution. Successfully applying and developing this framework will also result in concrete applications, as NODE systems could be used to control biological systems in real time, whether it is to maintain them at a maximum sustainable yield (e.g. in fisheries, agro-ecological systems), or to restore and protect them more efficiently against perturbations (e.g. invasive species, climate change).

ACKNOWLEDGEMENTS

The authors thank warmly the Ecological and Evolutionary Dynamics Lab and Sheldon Lab Group at the department of Zoology for their feedback and support, Mike Bonsall and José Lourenco for their feedback on early versions of the methods. They thank the reviewers for their enthusiasm and insightful comments. The work was supported by the NERC DTP and Oxford Oxitec Scholarship.

AUTHORS' CONTRIBUTIONS

W.B. designed the method, performed the analysis and wrote the manuscript; B.C.S. provided input for the manuscript and commented

on the manuscript; T.C. led investigations, provided input for the manuscript and commented on the manuscript.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.13606>.

DATA AVAILABILITY STATEMENT

All data and code are fully available at <https://github.com/Willembonnafe/NODEs/tree/master/NODER> and on Zenodo at <https://doi.org/10.5281/zenodo.4630706> (Bonnafe et al., 2021).

ORCID

Willem Bonnafe  <https://orcid.org/0000-0002-5053-8891>

Ben C. Sheldon  <https://orcid.org/0000-0002-5240-7828>

Tim Coulson  <https://orcid.org/0000-0001-9371-9003>

REFERENCES

- Aarts, L. P., & Van Der Veer, P. (2001). Neural network method for solving partial differential equations. *Neural Processing Letters*, 14(3), 261–271. <https://doi.org/10.1023/A:1012784129883>
- Berg, J., & Nyström, K. (2019). Data-driven discovery of PDEs in complex datasets. *Journal of Computational Physics*, 384, 239–252. <https://doi.org/10.1016/j.jcp.2019.01.036>
- Bonnafe, W., Sheldon, B. C., & Coulson, T. (2021). Neural ordinary differential equations for ecological and evolutionary time series analysis. *Zenodo*, <https://doi.org/10.5281/zenodo.4630706>
- Brighton, C. H., Thomas, A. L. R., & Taylor, G. K. (2017). Terminal attack trajectories of peregrine falcons are described by the proportional navigation guidance law of missiles. *Proceedings of the National Academy of Sciences of the United States of America*, 114(51), 13495–13500. <https://doi.org/10.1073/pnas.1714532114>
- Chen, R. T. Q., Rubanova, Y., Bettencourt, J., & Duvenaud, D. (2019). Neural ordinary differential equations. *arXiv*, 1–19.
- Chen, T., & Chen, H. (1993). Approximations of continuous functionals by neural networks with application to dynamic systems. *IEEE Transactions on Neural Networks*, 4(6), 910–918. <https://doi.org/10.1109/72.286886>
- Chon, T.-S., Kwak, I.-S., Park, Y.-S., Kim, T.-H., & Kim, Y. S. (2001). Patterning and short-term predictions of benthic macroinvertebrate community dynamics by using a recurrent artificial neural network. *Ecological Modelling*, 146(1–3), 181–193. [https://doi.org/10.1016/S0304-3800\(01\)00305-2](https://doi.org/10.1016/S0304-3800(01)00305-2)
- Chon, T.-S., Park, Y.-S., Kim, J.-M., Lee, B.-Y., Chung, Y.-J., & Kim, Y. S. (2000). Use of an artificial neural network to predict population dynamics of the forest-pest pine needle Gall Midge (Diptera: Cecidomyiidae). *Environmental Entomology*, 29(6), 1208–1215. <https://doi.org/10.1603/0046-225x.29.6.1208>
- Corcoran, A. J., & Conner, W. E. (2016). How moths escape bats: Predicting outcomes of predator-prey interactions. *Journal of Experimental Biology*, 219(17), 2704–2715. <https://doi.org/10.1242/jeb.137638>
- Cortez, M. H., & Patel, S. (2017). The effects of predator evolution and genetic variation on predator-prey population-level dynamics. *Bulletin of Mathematical Biology*, 79(7), 1510–1538. <https://doi.org/10.1007/s11538-017-0297-y>
- Cotto, O., Wessely, J., Georges, D., Klonner, G., Schmid, M., Dullinger, S., Thuiller, W., & Guillaume, F. (2017). A dynamic eco-evolutionary model predicts slow response of alpine plants to climate warming. *Nature Communications*, 8, 1–9. <https://doi.org/10.1038/ncomm.s15399>
- Dalziel, B. D., Morales, J. M., & Fryxell, J. M. (2008). Fitting probability distributions to animal movement trajectories: Using artificial neural networks to link distance, resources, and memory. *The American Naturalist*, 172(2), 248–258. <https://doi.org/10.1086/589448>
- Demyanov, V., Wood, S. N., & Kedwards, T. J. (2006). Improving ecological impact assessment by statistical data synthesis using process-based models. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 55(1), 41–62. <https://doi.org/10.1111/j.1467-9876.2005.00527.x>
- Deng, B. (2018). An inverse problem: Trappers drove hares to eat lynx. *Acta Biotheoretica*, 66(3), 213–242. <https://doi.org/10.1007/s10441-018-9333-z>
- Ellner, S. P., Geber, M. A., & Hairston, N. G. (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. <https://doi.org/10.2307/3072057>
- Ferguson, N. M., Laydon, D., Nedjati-Gilani, G., Imai, N., Ainslie, K., Baguelin, M., Bhatia, S., Boonyasiri, A., Cucunubá, Z., Cuomo-Dannenburg, G., Dighe, A., Dorigatti, I., Fu, H., Gaythorpe, K., Green, W., Hamlet, A., Hinsley, W., Okell, L. C., Van Elsland, S., ... Ghani, A. C. (2020). Impact of non-pharmaceutical interventions (NPIs) to reduce COVID-19 mortality and healthcare demand. Prepr. Spiral March, 3–20. <https://doi.org/10.25561/77482>
- Funahashi, K., & Nakamura, Y. (1993). Approximation of dynamical systems by continuous time recurrent neural networks. *Neural Networks*, 6(6), 801–806. [https://doi.org/10.1016/S0893-6080\(05\)80125-X](https://doi.org/10.1016/S0893-6080(05)80125-X)
- Fussmann, K. E., Rosenbaum, B., Brose, U., & Rall, B. C. (2017). Interactive effects of shifting body size and feeding adaptation drive interaction strengths of protist predators under Warming. *bioRxiv*, 1–34. <https://doi.org/10.1101/101675>
- Gamelon, M., Vriend, S. J. G., Engen, S., Adriaensen, F., Dhondt, A. A., Evans, S. R., Matthysen, E., Sheldon, B. C., & Sæther, B.-E. (2019). Accounting for interspecific competition and age structure in demographic analyses of density dependence improves predictions of fluctuations in population size. *Ecology Letters*, 22(5), 797–806. <https://doi.org/10.1111/ele.13237>
- Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Chapman & Hall/CRC Texts in Statistical Science.
- Grant, B. R., & Grant, P. R. (1993). Evolution of Darwin's finches caused by a rare climatic event. *Proceedings of the Royal Society B: Biological Sciences*, 251(1331), 111–117. <https://doi.org/10.1098/rspb.1993.0016>
- Hairston, N. G., 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. <https://doi.org/10.1111/j.1461-0248.2005.00812.x>
- Hiltunen, T., Jones, L. E., Ellner, S. P., & Hairston, N. G. (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>
- Holyoak, M., & Heath, S. K. (2016). The integration of climate change, spatial dynamics, and habitat fragmentation: A conceptual overview. *Integrative Zoology*, 11(1), 40–59. <https://doi.org/10.1111/1749-4877.12167>
- 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. <https://doi.org/10.1890/03-4056>
- Krebs, C. J., Boonstra, R., Boutin, S., & Sinclair, A. R. E. (2006). What drives the 10-year cycle of snowshoe hares? *BioScience*, 51(1), 1–35. [https://doi.org/10.1641/0006-3568\(2001\)051\[0025:wdtyco\]2.0.co;2](https://doi.org/10.1641/0006-3568(2001)051[0025:wdtyco]2.0.co;2)

- Lek, S., & Gue, J. F. (1999). Artificial neural networks as a tool in ecological modelling, an introduction. *Ecological Modelling*, 120, 1–9. [https://doi.org/10.1016/S0304-3800\(99\)00092-7](https://doi.org/10.1016/S0304-3800(99)00092-7)
- Lion, S. (2018). Theoretical approaches in evolutionary ecology: Environmental feedback as a unifying perspective. *The American Naturalist*, 191(1), 21–44. <https://doi.org/10.1086/694865>
- Mai, M., Shattuck, M. D., & O'Hern, C. S. (2016). Reconstruction of ordinary differential equations from time series data. *arXiv*, 1–15. 1605.05420v1.
- Massaroli, S., Poli, M., Park, J., Yamashita, A., & Asama, H. (2020). Dissecting neural ODEs. *arXiv*, 1–19.
- Matsunaga, F., Rakocevic, M., & Brancher, J. D. (2013). Artificial neural networks in modeling of environmental time series for yerbamate growth dynamics. *Proceedings of the 7th International Conference on Functional-Structural Plant Models*, 2016, 2005–2007.
- Montoya, J. M., Woodward, G., Emmerson, M. C., & Solé, R. V. (2009). Press perturbations and indirect effects in real food webs. *Ecology*, 90(9), 2426–2433. <https://doi.org/10.1890/08-0657.1>
- Norberg, J., Urban, M. C., Vellend, M., Klausmeier, C. A., & Loeuille, N. (2012). Eco-evolutionary responses of biodiversity to climate change. *Nature Climate Change*, 2(10), 747–751. <https://doi.org/10.1038/Nclimate1588>
- Obach, M., Wagner, R., Werner, H., & Schmidt, H.-H. (2001). Modelling population dynamics of aquatic insects with artificial neural networks. *Ecological Modelling*, 146(1–3), 207–217. [https://doi.org/10.1016/S0304-3800\(01\)00307-6](https://doi.org/10.1016/S0304-3800(01)00307-6)
- Odum, E. P., & Barrett, G. W. (1972). Fundamentals of ecology. *Journal of Wildlife Management*, 36(4), 1372. <https://doi.org/10.2307/3799291>
- Olberg, R. M. (2012). Visual control of prey-capture flight in dragonflies. *Current Opinion in Neurobiology*, 22(2), 267–271. <https://doi.org/10.1016/j.conb.2011.11.015>
- Patel, S., Cortez, M. H., & Schreiber, S. J. (2018). Partitioning the effects of eco-evolutionary feedbacks on community stability. *The American Naturalist*, 191(3), 381–394. <https://doi.org/10.1086/695834>
- Pierson, J. C., Beissinger, S. R., Bragg, J. G., Coates, D. J., Oostermeijer, J. G. B., Sunnucks, P., Schumaker, N. H., Trotter, M. V., & Young, A. G. (2015). Incorporating evolutionary processes into population viability models. *Conservation Biology*, 29(3), 755–764. <https://doi.org/10.1111/cobi.12431>
- Pigeon, G., Ezard, T. H. G., Festa-Bianchet, M., Coltman, D. W., & Pelletier, F. (2017). Fluctuating effects of genetic and plastic changes in body mass on population dynamics in a large herbivore. *Ecology*, 98(9), 2456–2467. <https://doi.org/10.1002/ecy.1940>
- 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.
- Rubanov, Y., Chen, R. T. Q., & Duvenaud, D. (2019). Latent ODEs for irregularly-sampled time series, *arXiv*, 1–11.
- Sinchev, B., Sibanbayeva, S. E., Mukhanova, A. M., Nurgulzhanova, A. N., Zaurbekov, N. S., Imanbayev, K. S., Gagarina, N. L., & Baibolova, L. K. (2018). Some methods of training radial basis neural networks in solving the Navier-Stokes equations. *International Journal for Numerical Methods in Fluids*, 86(10), 625–636. <https://doi.org/10.1002/fld.4470>
- Slough, B. G., & Mowat, G. (1996). Lynx population dynamics in an untrapped refugium. *Journal of Wildlife Management*, 60(4), 946–961. <https://doi.org/10.2307/3802397>
- Stenseth, N. C., Falck, W., Chan, K.-S., Bjornstad, O. N., O'Donoghue, M., Tong, H., Boonstra, R., Boutin, S., Krebs, C. J., & Yoccoz, N. G. (2002). From patterns to processes: Phase and density dependencies in the Canadian lynx cycle. *Proceedings of the National Academy of Sciences of the United States of America*, 95(26), 15430–15435. <https://doi.org/10.1073/pnas.95.26.15430>
- 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. <https://doi.org/10.1126/science.1227079>
- Sussillo, D., Churchland, M. M., Kaufman, M. T., & Shenoy, K. V. (2015). A neural network that finds a naturalistic solution for the production of muscle activity. *Nature Neuroscience*, 18(7), 1025–1033. <https://doi.org/10.1038/nn.4042>
- Van Velzen, E., & Gaedke, U. (2017). Disentangling eco-evolutionary dynamics of predator-prey coevolution: The case of antiphase cycles. *Scientific Reports*, 7(1), 1–11. <https://doi.org/10.1038/s41598-017-17019-4>
- Wood, S. N. (2001). Partially specified ecological models. *Ecological Monographs*, 71(1), 1–25. [https://doi.org/10.1890/0012-9615\(2001\)071\[0001:PSEM\]2.0.CO;2](https://doi.org/10.1890/0012-9615(2001)071[0001:PSEM]2.0.CO;2)
- 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. <https://doi.org/10.1016/j.ecolmodel.2005.04.013>

How to cite this article: Bonnaffé W, Sheldon BC, Coulson T. Neural ordinary differential equations for ecological and evolutionary time-series analysis. *Methods Ecol Evol*. 2021;12:1301–1315. <https://doi.org/10.1111/2041-210X.13606>