PREDICTING AN UNOBSERVED DRIVER OF REGIME SHIFTS IN SOCIAL-ECOLOGICAL SYSTEMS WITH UNIVERSAL DYNAMIC EQUATIONS

A PREPRINT

Kunal J. Rathore*
College of Earth, Ocean, and Atmospheric Sciences,
Oregon State University, Corvallis, OR, USA
rathorek@oregonstate.edu

John H. Buckner
College of Earth, Ocean, and Atmospheric Sciences,
Oregon State University, Corvallis, OR, USA
bucknejo@oregonstate.edu

Zechariah D. Meunier
College of Earth, Ocean, and Atmospheric Sciences,
Oregon State University, Corvallis, OR, USA
meunierz@oregonstate.edu

Jorge Arroyo Esquivel
California Department of Fish and Wildlife,
West Sacramento, California, USA
arroyoesquiveljorge@gmail.com

James R. Watson
College of Earth, Ocean, and Atmospheric Sciences,
Oregon State University, Corvallis, OR, USA
james.watson@oregonstate.edu

June 20, 2026

ABSTRACT

Ecosystems around the world are anticipated to undergo regime shifts as temperatures rise and other climatic and anthropogenic perturbations erode the resilience of present-day states. Forecasting these nonlinear ecosystem dynamics can help stakeholders better prepare. One major challenge though is that regime shifts can be difficult to predict when they are driven by unobserved factors. In this paper, we advance Scientific Machine Learning methods, specifically Universal Dynamic Equations (UDEs), to identify changes in an unobserved bifurcation parameter that is driving an ecosystem regime shift. We demonstrate this approach using simulated data created from a dynamic model of a species population experiencing loss due to extraction or harvest that is unobserved. This could be, for example, illegal fishing from a fishery or unreported poaching in a game reserve. We show that UDEs can accurately identify changes in the slowly increasing harvest rate (the bifurcation parameter), and predict when a regime shift will occur. Compared to alternative forecasting methods, our UDE approach provides relatively accurate short-term predictions, and provides a new set of methods for ecosystem stakeholders and to manage fast-paced ecosystem change likely to happen in the coming decades.

Keywords harvest rate, nonlinear dynamics, regime shift, population dynamics, neural networks, UDE, scientific ML

1 Introduction

Human activities around the world are exerting increasing pressure on ecosystems. These impacts can be amplified by biological feedback mechanisms, leading to large and abrupt changes in the state of ecosystems called regime shifts [Scheffer et al., 2001, Rocha et al., 2018, Levin and Möllmann, 2015, Folke et al., 2004]. Ecological regime

*Corresponding author.

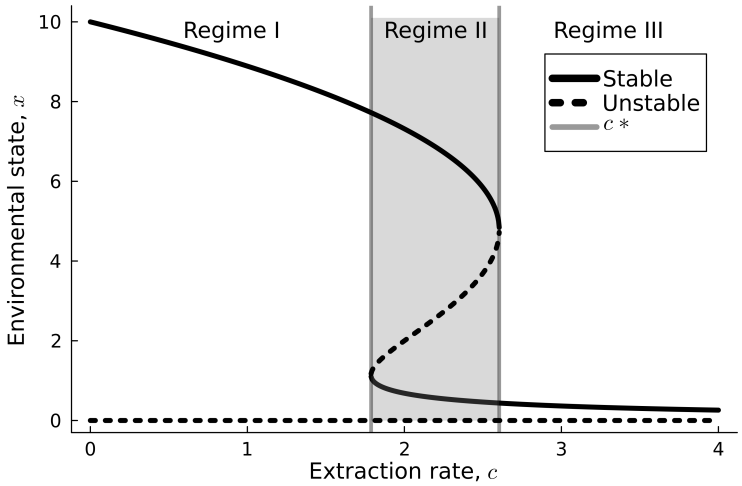

Figure 1: Bifurcation diagram illustrating the existence of three regimes in the deterministic dynamical system, where environmental state x is a function of the harvest rate c_t. Solid lines indicate stable states, and dotted lines show unstable states. Vertical gray lines mark the regime boundaries at the equilibrium harvest rates $c*$ with Regime II (shaded box) exhibiting bistability. Modified from [Tilman et al., 2024]

shifts can reduce the provision of ecosystem services, including carbon storage, biodiversity, and the maintenance of local climates, with cascading effects on food and water resources. These transitions, often driven by unobserved or nonlinear dynamics, can occur abruptly and intensify socioeconomic vulnerability by disrupting ecosystem services and livelihoods [Staver et al., 2011, Turner et al., 2020, Biggs et al., 2018]. Understanding and forecasting the state of complex human-environmental systems becomes necessary if resource managers are to maximize economic benefits while preventing irreversible shifts.

Ecological forecasting can provide ecosystem stakeholders and policymakers vital actionable information about how to adapt to changing environmental conditions [Dietze, 2017a, Dietze et al., 2024]. Predictive models play a crucial role in ecological forecasting, as they can help us to understand and prepare for often large and abrupt changes in ecosystem states [Clark et al., 2001, Kickert et al., 1999, Griffith and Fulton, 2014]. Examples include predicting species population dynamics in response to climate change [Urban et al., 2016], forecasting harmful algal blooms for effective water resource management [Anderson et al., 2012], and estimating wildfire risk based on vegetation and atmospheric conditions [Abatzoglou and Williams, 2016]. However, one major challenge in ecological forecasting is accounting for unobserved drivers of change. This is a common problem in many scientific domains such as epidemiology, where asymptomatic carriers and unreported cases alter the dynamics of disease transmission [Fraser et al., 2004]; economics, where latent variables such as consumer confidence impact macroeconomic outcomes [Stock and Watson, 2016]; and neuroscience, where unobserved neural states influence observed behavior and brain activity patterns [Paninski et al., 2010]. This challenge is particularly critical when the unobserved quantity is a bifurcation parameter that governs rapid and large shifts in the qualitative properties of ecosystem dynamics.

Here, we consider populations that exhibit regime shifts caused by a hidden source of loss, such as illegal harvest from a fishery or poaching from a game reserve. In these systems, the harvest rate can be a bifurcation parameter [May and Oster, 1976] leading to unanticipated regime shifts. In such a scenario, an ecosystem with high abundance or biomass (i.e., a stable regime, Fig. 1 Regime I) could pass through a transient regime (i.e., a flickering regime, Fig. 1 Regime II) to low abundance or biomass (i.e., the second stable regime, Fig. 1 Regime III). This presents a significant challenge for ecosystem managers: in systems where the bifurcation parameter is unobserved or poorly monitored, proximity to a critical threshold may go unrecognized until a transition is already underway [Scheffer et al., 2009, Biggs et al., 2009]. While empirical evidence for abrupt, externally-driven regime shifts remains limited to a relatively small number of well-documented cases [Scheffer et al., 2009], theory suggests that hidden drivers of loss are particularly difficult to account for under conventional monitoring frameworks. Forecasting methods that identify the unobserved parameters could therefore improve early warning and management of ecosystems as they approach critical transitions.

While early efforts to predict regime shifts relied on statistical modeling, recent advancements have increasingly integrated machine learning and, more critically, the formalisms of dynamical systems [Dakos et al., 2015, Ghadami and Epureanu, 2022, Panahi et al., 2024]. Dynamical models are key tools for forecasting regime shifts, as they offer a framework for predicting critical transitions between stable states in ecosystems based on an explicit representation of

key processes. These models use discrete-time difference and continuous-time differential equations to represent the underlying mechanisms driving the system's dynamics, allowing researchers to identify early warning signals of tipping points and to improve our understanding of why these changes occur [Barnosky et al., 2012, Kéfi et al., 2014]. By formally representing interactions among system components and their responses to external perturbations, dynamical models offer information on the conditions that precipitate abrupt changes, such as shifts in ecosystem states, climatic patterns, or disease outbreaks [Biggs et al., 2015, May, 1977a].

While early efforts to predict regime shifts relied on statistical modeling, recent advancements have increasingly integrated machine learning and, more critically, the formalisms of dynamical systems [Dakos et al., 2015, Ghadami and Epureanu, 2022, Panahi et al., 2024]. Dynamical models are key tools for forecasting regime shifts, as they offer a framework for predicting critical transitions between stable states in ecosystems based on an explicit representation of key processes. These models use discrete-time difference and continuous-time differential equations to represent the underlying mechanisms driving the system's dynamics, allowing researchers to identify early warning signals of tipping points and to improve our understanding of why these changes occur [Barnosky et al., 2012, Kéfi et al., 2014]. By formally representing interactions among system components and their responses to external perturbations, dynamical models offer information on the conditions that precipitate abrupt changes, such as shifts in ecosystem states, climatic patterns, or disease outbreaks [Biggs et al., 2015, May, 1977a].

Despite their relevance, these models have notable limitations when it comes to forecasting regime shifts. The parametrization process in dynamical models is crucial, but it can be challenging because multiple parameter sets can produce dynamics that are consistent with the data. In addition, the complexity of ecological systems characterized by numerous abiotic and biotic components, as well as direct and indirect interactions, introduces significant uncertainty in the predictions of dynamical models [Dietze, 2017b]. Furthermore, the functional forms embedded in these models, such as the choice of growth functions or predator-prey interaction terms, are often selected for mathematical tractability or historical convention rather than empirical support and may inadequately represent the nonlinear dynamics that govern regime shifts [Clark et al., 2001]. These models are also prone to overfitting (as are most models in ecology admittedly), especially with sparse or high-dimensional data, and struggle with the uncertainty and stochastic nature of ecological processes [Turchin, 2013].

In contrast, data-driven "equation-free" methods are also commonly used to make ecological forecasts. They differ from equation-based approaches in that they do not rely on predefined functional forms or assumptions about the underlying processes. One such approach, empirical dynamical modeling (EDM), uses time series data to reconstruct and infer system dynamics without assuming explicit equations [Sugihara et al., 2012, Perretti et al., 2013, Munch and Brias, 2024]. Recent extensions of EDM have further demonstrated its capacity to anticipate both the timing and type of critical transitions in nonlinear systems, as well as to detect regime shifts in chaotic dynamics and high-dimensional ecosystems [Grziwotz et al., 2023, Huang et al., 2024, Garain et al., 2025]. Despite these advances, purely equation-free methods remain limited in their ability to incorporate known mechanistic structure. An early attempt to bridge this gap was Wood's [Wood, 2001] framework of partially specified ecological models, which replaced unknown interaction terms with flexible nonparametric functions while retaining the mechanistic skeleton of a dynamic model, demonstrating that embedding data-driven flexibility within a mechanistic structure can improve both model fit and ecological interpretability [Wood, 2001].

In addition, several machine learning approaches have been developed to provide predictions of ecosystem dynamics. Advanced neural network architectures, such as the recurrent neural network and long short-term memory models, have shown great potential in overcoming the limitations of traditional equation-based methods [Hochreiter and Schmidhuber, 1997]. However, the lack of mechanistic interpretability in these methods remains a key drawback that limits their broader applicability [Lipton et al., 2015]. Similarly, by not having explicit functional forms, these equation-free approaches often overlook valuable mechanistic insights that could improve forecasts, particularly when data are limited or have observation biases [Perretti and Munch, 2015, Greff et al., 2017]. Together, these limitations motivate the need for hybrid frameworks that can embed data-driven flexibility within a mechanistic structure—a principle now formalized through universal dynamic equations (UDEs).

Recently, [Rackauckas et al., 2020, Bonnaffé et al., 2021, Arroyo-Esquivel et al., 2024, Buckner et al., 2026] demonstrated the potential of Scientific Machine Learning (SciML) methods, which combine theoretical knowledge, often represented by differential and difference equations, with data-driven neural networks to model nonlinear ecosystem dynamics. Using mathematical representations of known interactions and dynamics, in conjunction with deep neural networks, SciML bridges the gap between equation-based mechanistic approaches and equation-free data-driven approaches, offering improved forecasting skills while maintaining interpretability. Building on this foundation, we have extended the application of SciML to address two key challenges in ecosystem management: (1) identifying unobserved parameter that influence ecosystem behavior, and (2) making accurate forecasts based on this information.

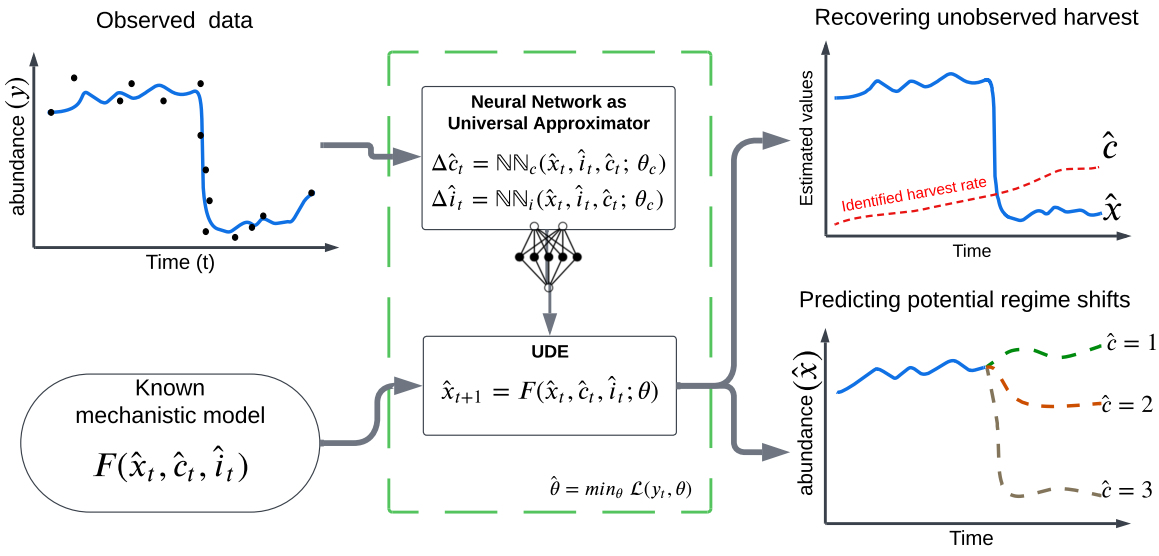

Figure 2: Workflow diagram showing the process to identify the unobserved parameter and predict regime shifts using a universal dynamic equation (UDE) model. UDE model can identify the unobserved (bifurcation) parameter and predict regime shifts in a simulated ecological population experiencing some form of harvest. We simulated 50 time-series as observed data and modeled the data using a neural network embedded in a mechanistic model of the known dynamics of the system. The UDE contained the neural network as a universal function approximator, and optimized the formulated loss function (green box). Then, the UDE model was used to estimate the unobserved harvest rate that changes over time. Finally, we generated forecasts with different harvest values to identify potential regime shifts in the ecological population.

Although we focus on social-ecological systems here, this application of SciML is generalizable to other complex systems with nonlinear dynamics driven by changes in an unobserved parameter.

In this study, we use a specific form of SciML called state-space UDEs [Buckner et al., 2026], which embed neural networks within difference equations that represent the discrete-time dynamics of an ecosystem, and the neural networks are trained on noisy time series data using a state-space modeling framework. Here we extend this approach by using state-space UDEs to estimate the unobserved bifurcation parameter of a social-ecological system undergoing a regime shift, and forecast population abundance in different regimes. We test this approach on simulated data from a social-ecological model of a harvested population that exhibits flickering between alternative stable states and a regime change as the harvest rate increases past a critical threshold. We assessed the UDE's ability to classify dynamic regimes by treating the estimated unobserved harvest parameter as an indicator of ecological regime and evaluated prediction accuracy using transition thresholds. We also compared the model's forecasting skill with various modern-day ecosystem forecasting methods.

2 Methods

2.1 Modeling social-ecological dynamics

We generated synthetic data using a discrete-time social-ecological model that combines logistic population growth with a nonlinear consumption term, following the implementation of [Tilman et al., 2024]. This class of models, rooted in the bistable harvesting frameworks of [May, 1977b] and [Scheffer et al., 2001], effectively represents the nonlinear dynamics of many social-environmental systems, accounting for environmental stochasticity, such as the collapse of fisheries due to overfishing or kelp forests affected by sea otter populations [Nicholson et al., 2024]. In the model, x represents the population abundance of the target species. Population growth is modeled as logistic growth with an intrinsic growth rate r, carrying capacity K, and multiplicative shocks to the growth rate. Animals are harvested from the population at a rate that depends on a nonlinear function of their abundance x_t, a parameter c_t that determines the harvest rate, and a parameter h, which determines the strength of nonlinearity. Population abundances update in discrete

time as follows:

$$x_{t+1} = r x_t \left(1 - \frac{x_t}{K}\right) - c_t \left(\frac{x_t^2}{x_t^2 + h^2}\right) + (1 + i_t) x_t \tag{1}$$

where i_t represents time-correlated red noise modeling environmental shocks. This type of noise characterizes scenarios where random environmental fluctuations exhibit temporal correlation rather than being fully independent over time, a pattern commonly observed in nature [Allen, 2010]. In equation (2), the parameter T controls the rate at which the autocorrelation of noise decays, with the correlation between i_t and i_{t+n} given by $(1 - 1/T)^n$, and $\eta_t \sim \mathcal{N}(0,\, \beta^2)$ is a series of independent and identically distributed normal errors.

$$i_t = \left(1 - \frac{1}{T}\right) i_{t-1} + \eta_t \tag{2}$$

Figure 1 illustrates that for low harvest rates c_t, the system settles on a single high-abundance equilibrium point. In contrast, at high values of c_t, the system moves to lower abundance stable state. For intermediate c_t values, bistability occurs, and the presence of noise induces flickering dynamics, causing the system to frequently transition between the high and low abundance basins of attraction [Tilman et al., 2024, Dakos et al., 2012]. To explore the ability of state-space UDEs to identify changes in an unobserved parameter (i.e., the harvest rate c_t) and to predict regime shifts in abundance, we generated 50 simulated time series, of length 400 timesteps, for the population abundance x_t and noise terms i_t from equations (1) and (2). Throughout the simulations, we increased the harvest rate c_t linearly from $[0, 4]$, which caused the system to pass from the high abundance regime, through the bistable flickering regime, to the low abundance regime (Fig. 3). We used the set of base parameters for the social-ecological model discussed in [Tilman et al., 2024] for each simulation (Table 2). We incorporated observation noise into the synthetic data by simulating data points y_t from a normal distribution truncated from below at zero with mean x_t and standard deviation 1.0.

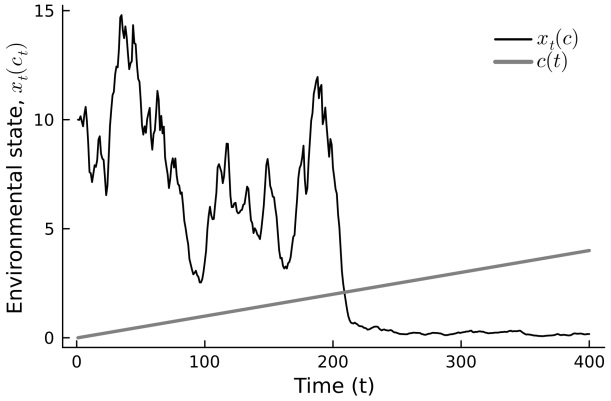

Figure 3: An example of a simulated time series of the environmental state (x_t) and harvest rate (c_t). The harvest rate is a monotonically increasing function with time, and its effects can be observed on the environmental state through the dynamical model of [Tilman et al., 2024].

2.2 Scientific machine learning of regime shifts

We developed a state-space UDE model that jointly estimates the population abundance $\hat{x}_t$, harvest rate $\hat{c}_t$, and environmental noise $\hat{i}_t$ from noisy observations of abundance, which we denote as y_t (see Fig. 2). The only input data are noisy abundance observations $y_t = x_t + \varepsilon_t$, where $\varepsilon_t \sim \mathcal{N}(0, \sigma^2)$, while the harvest rate c_t is never directly observed, and the observability of i_t depends on the availability of the data (discussed in Section 2.2.3). The system parameters r, K, and h are assumed to be known.

2.2.1 State estimation with Universal Dynamic Equations

State estimation proceeds by specifying a dynamic equation for each quantity, drawing on known ecological structure where available. Changes in $\hat{x}_t$ are governed by the known mechanistic model of [Tilman et al., 2024], which describes how logistic growth, harvesting, and environmental forcing jointly determine abundance at the next time step:

$$\hat{x}_{t+1} = \hat{x}_t + r\hat{x}_t \left(1 - \frac{\hat{x}_t}{K}\right) - \hat{c}_t \left(\frac{\hat{x}_t^2}{\hat{x}_t^2 + h^2}\right) + \hat{i}_t\, \hat{x}_t. \tag{3}$$

While the structure of this equation is fully specified, evaluating it requires knowing $\hat{c}_t$ and $\hat{i}_t$ at each time step, neither of which is directly observed. We model changes in each step using a dedicated artificial neural network, as described below. We modeled the harvesting rate $\hat{c}_t$ using a fully connected multilayer network $\mathbb{NN}_c$ with activations and weights $\boldsymbol{\theta}_c$. Rather than predicting $\hat{c}_t$ directly, $\mathbb{NN}_c$ predicts the change $\Delta\hat{c}_t = \hat{c}_{t+1} - \hat{c}_t$ at each time step, reflecting the assumption that the harvest rate varies slowly relative to the observation interval:

$$\hat{c}_{t+1} = \hat{c}_t + \mathbb{NN}_c(\hat{x}_t, \hat{i}_t, \hat{c}_t; \boldsymbol{\theta}_c). \tag{4}$$

For environmental noise, the autocorrelated structure of eq. (2), whereby $\hat{i}_t$ decays to zero over time, is preserved explicitly and $\mathbb{NN}_i$ learns the residual deviation at each step.

$$\hat{i}_{t+1} = \left(1 - \frac{1}{T}\right)\hat{i}_t + \mathbb{NN}_i(\hat{x}_t, \hat{i}_t, \hat{c}_t; \boldsymbol{\theta}_i). \tag{5}$$

Equations (3), (4), and (5) together define a one-step map F over the augmented state vector $\hat{\boldsymbol{u}}_t = (\hat{x}_t, \hat{c}_t, \hat{i}_t)^\top$:

$$\begin{pmatrix} \hat{x}_{t+1} \\ \hat{c}_{t+1} \\ \hat{i}_{t+1} \end{pmatrix} = F\Big(\hat{x}_t, \hat{c}_t, \hat{i}_t; \boldsymbol{\theta}\Big) \equiv \begin{pmatrix} \hat{x}_t + r\hat{x}_t(1 - \hat{x}_t/K) - \hat{c}_t\,\hat{x}_t^2/(\hat{x}_t^2 + h^2) + \hat{i}_t\,\hat{x}_t \\ \hat{c}_t + \mathbb{NN}_c(\hat{x}_t, \hat{i}_t, \hat{c}_t; \boldsymbol{\theta}_c) \\ (1 - 1/T)\,\hat{i}_t + \mathbb{NN}_i(\hat{x}_t, \hat{i}_t, \hat{c}_t; \boldsymbol{\theta}_i) \end{pmatrix}. \tag{6}$$

The first row of F is entirely mechanistic: the known relationship governs how the abundance changes. The second and third rows are hybrid: the known structure is preserved (incremental update for $\hat{c}_t$; learned autoregressive updates for $\hat{i}_t$), and the neural networks fill in only what the mechanistic model cannot specify. The weights of the network $\boldsymbol{\theta} = \{\boldsymbol{\theta}_c, \boldsymbol{\theta}_i\}$ and the latent trajectories $\{\hat{x}_t, \hat{c}_t, \hat{i}_t\}_{t=1}^T$ are estimated simultaneously by optimizing the composite loss described below.

2.2.2 Composite loss function

To estimate $\boldsymbol{\theta}$ and the latent trajectories, we minimize a composite loss function with three terms:

$$L_{\text{total}} = L_{\text{dyn}} + L_{\text{obs}} + L_{\text{reg}}. \tag{7}$$

The dynamic loss L_{dyn} ensures that the inferred trajectories of $\hat{x}_t$, $\hat{c}_t$, and $\hat{i}_t$ are jointly consistent with the dynamics of the model at each time step. Denoting $\Delta\hat{c}_t = \hat{c}_{t+1} - \hat{c}_t$, it penalizes the squared residual between the stored trajectory at $t+1$ and the one-step prediction of F from t:

$$L_{\text{dyn}} = \frac{1}{\sigma_{dyn}^2} \sum_{t=1}^{T-1} \left\| \hat{\boldsymbol{u}}_{t+1} - F(\hat{x}_t, \hat{c}_t, \hat{i}_t; \boldsymbol{\theta}) \right\|_2 \tag{8}$$

where σ_{dyn} is the weight and $F_x(\hat{x}_t, \hat{c}_t, \hat{i}_t)$ denotes the first row of F in equation (6), i.e., the right-hand side of equation (3).

The observational loss L_{obs} constrains $\hat{x}_t$ to match the observed abundance data y_t:

$$L_{\text{obs}} = \frac{1}{\sigma_{obs}^2} \sum_{t=1}^{T} (y_t - \hat{x}_t)^2. \tag{9}$$

where σ_{obs} controls the weight of observational loss. Finally, the regularization loss L_{reg} penalizes large network weights to reduce overfitting:

$$L_{\text{reg}} = \zeta \sum_l \boldsymbol{\theta}_l^2, \tag{10}$$

where ζ controls the relative strength of regularization. We considered $\zeta = 0.2$ an appropriate value in our case. In all experiments, we set $1/\sigma_{dyn}^2 = 1/\sigma_{obs}^2 = 0.4$, so that the dynamic and observation terms contribute equally to the composite loss. Prior work has shown that this simple equal-weighting scheme does not substantially affect the recovery of latent states in hybrid models of this kind [Treven et al.].

2.2.3 Including partially observed environmental variability

The framework above treats both c_t and i_t as unobserved, requiring their simultaneous estimation from observed abundance alone. This is a dual latent state estimation problem: since both quantities contribute to abundance dynamics through equation (3), disentangling their respective contributions relies on the structural differences in how each enters the model.

However, in many ecological monitoring programs, environmental conditions can be approximated by a measured proxy, such as a temperature or productivity index. When such a proxy $\tilde{i}_t \approx i_t$ is available, the environmental state is treated as observed and substituted directly into the abundance equation and provided as an additional input to $\mathbb{NN}_c$, while $\mathbb{NN}_i$ is removed. In our experiments, $\tilde{i}_t$ was generated from the simulation model as $\tilde{i}_t = i_t + \varepsilon_t^{(i)}$, where $\varepsilon_t^{(i)} \sim \mathcal{N}(0, \sigma_i^2)$, analogous to how noisy abundance observations y_t are obtained from x_t. The transition map F then simplifies to:

$$\hat{x}_{t+1} = \hat{x}_t + r\hat{x}_t\left(1 - \frac{\hat{x}_t}{K}\right) - \hat{c}_t\left(\frac{\hat{x}_t^2}{\hat{x}_t^2 + h^2}\right) + \tilde{i}_t\,\hat{x}_t, \qquad \hat{c}_{t+1} = \hat{c}_t + \mathbb{NN}_c(\hat{x}_t, \tilde{i}_t; \boldsymbol{\theta}). \tag{11}$$

For forecasting, where no proxy observations are available beyond the training window, $\hat{i}_t$ is propagated forward using the AR1 mean: $\hat{i}_{t+1} = (1 - 1/T)\,\hat{i}_t$. Comparing the two observability scenarios, observing only x_t versus both x_t and i_t, can indicate how much of the difficulty in recovering c_t arises from the confounding influence of unobserved environmental forcing. This comparison is central to our evaluation in Section 3.

2.3 Evaluation

We evaluated the trained UDE framework on two complementary simulation tests designed to probe distinct aspects of model performance: the ability to reconstruct the unobserved harvest rate c_t driving regime transitions and the ability to forecast future population abundances. Each test is described below in terms of its purpose, performance metrics, and comparison methods.

2.3.1 Evaluating harvest rate recovery

The recovery test evaluated how well each method reconstructs the unobserved harvest rate c_t, the bifurcation parameter that governs the regime the system occupies. Because c_t is never directly observed, this test addresses the core inferential challenge of the framework: recovering a slowly varying signal from noisy abundance observations alone. Accurate recovery of c_t helps to explain past dynamics in the system and allows the neural networks in our model to forecast future values of c_t, helping predict future transitions.

Recovery performance was assessed using two complementary metrics. Pointwise reconstruction of c_t was measured by the RMSE between the estimated $\hat{c}_t$ and the true simulated harvest. To evaluate regime-level identification, we also calculated the F-score – the harmonic mean of precision and recall, applied to regime classifications derived from $\hat{c}_t$. We considered a macro-averaged F-score metric, which treats each class performance equally, making it preferable when class frequencies differ across regimes. The critical transition values c^* derived from the mechanistic model of [Tilman et al., 2024] (Fig. 1) served as thresholds to map predicted $\hat{c}_t$ trajectories to three regime classes. This classification approach addresses a practically important question: not only how close the estimated $\hat{c}_t$ is numerically to the truth, but also whether the method correctly identifies which side of the critical threshold the system currently occupies.

To establish a benchmark, we compared the UDE modeling approach with two alternative approximation methods within the same state-space difference equation framework: 1) a hybrid model that used a Gaussian Process to learn the unknown function and 2) an algebraic solution for c_t estimated from the observation directly.

Difference Equation Gaussian Process surrogate (DE-GP) In addition to neural networks, we explored Gaussian Processes (GPs) as an alternative function approximator within the state-space dynamic equation framework [Olivier et al., 2021]. GPs have been extensively validated for differential equation problems similar to our UDEs [Medeiros et al., 2025, Raissi et al., 2019, 2017], offer superior uncertainty quantification that is critical for scientific applications [Yang et al., 2020], and allow physical constraints to be encoded through specialized kernel design [Solin and Särkkä, 2020]. The theoretical equivalence between infinitely-wide neural networks and GPs [Lee et al., 2017] makes this comparison particularly informative for understanding the relative advantages of Bayesian and neural network approximators in the scientific machine learning literature.

We incorporated GPs into the difference equation (DE) framework (referred to hereafter as DE-GP) using kernel-based representation to provide smooth, locally adaptive, and probabilistically-calibrated approximations of unobserved

dynamics while maintaining computational tractability [Turner et al., 2010]. DE-GP shares the same structural form as the UDE: the mechanistic abundance equation (eq. 3) governs the observable state, while the evolution of the unobserved harvest rate $\hat{c}_t$ is learned from data by substituting GP for the neural network $\mathbb{NN}_c$ in eqs. (8)–(9). Model parameters were optimized using a gradient-based technique and the implementation used the `AbstractGPs.jl` [Widmann et al., 2024] and `KernelFunctions.jl` [Galy-Fajou et al., 2024] packages in Julia.

Algebraic inversion reference For the algebraic inversion benchmark, we computed $\hat{c}_t$ by directly solving for c_t the mechanistic model given the observations at each time step. Rearranging equation (1) and solving for c_t given consecutive abundance observations y_t and y_{t+1}, and substituting the noisy environmental proxy $\tilde{i}_t$, yields:

$$\hat{c}_t^{\text{alg}} \approx \frac{r\, y_t(1 - y_t/K) + (1 + \tilde{i}_t)\, y_t - y_{t+1}}{y_t^2/(y_t^2 + h^2)}, \tag{12}$$

where the approximation arises because $\tilde{i}_t = i_t + \eta_t$ is a noisy proxy for the true environmental noise, and any error in $\tilde{i}_t$ propagates directly into $\hat{c}_t^{\text{alg}}$. This estimate requires no training and assumes that the environmental state is partially observable; therefore, it represents the recovery achievable by direct model inversion under realistic observational noise, establishing an upper bound on recovery skill when the model structure is known exactly.

We further evaluated the UDE and DE-GP approaches under two observability conditions; only abundance y_t was observed, and both y_t and environmental noise $\tilde{i}_t$ were observed, to isolate how much of the difficulty in recovering c_t arises from confounding with unobserved environmental forcing (Section 2.2.3).

2.3.2 Evaluating forecast skill across regimes

The forecast test assessed how accurately each method predicts future abundance $\hat{x}_t$ in dynamically distinct regimes. After training in a fixed window of simulated observations, we generated out-of-sample forecasts initialized at three regime entry points: high abundance (regime I, $t = 70$), intermediate bistability with flickering (regime II, $t = 220$), and low abundance (regime III, $t = 300$). Evaluating across these three conditions allowed us to test predictive skill in both stable regimes and near the critical transition, where small errors in state or parameter estimation compound rapidly. Forecast performance was quantified by the root mean squared error (RMSE) between predicted and observed abundances across each forecast horizon.

Three benchmark methods spanning classical statistics and modern machine learning were compared against the UDE and DE-GP frameworks. The autoregressive integrated moving average (ARIMA) model [Durbin and Koopman, 2012] serves as the canonical statistical baseline routinely applied in ecological forecasting because it captures linear temporal dependence and trends in stationary series. A random-walk null model provides the most parsimonious reference, projecting the last observed value forward without any learned structure; outperforming this baseline is the minimum bar for a useful forecasting method. Long short-term memory (LSTM) networks [Hochreiter and Schmidhuber, 1997] provide a state-of-the-art deep learning benchmark, designed to capture long-range temporal dependencies via learned gating mechanisms that address the vanishing gradient problem in conventional recurrent networks. Together, these benchmarks span the methodological space from null models to fully data-driven deep learning, allowing us to isolate the contribution of embedding mechanistic structure in the UDE framework.

2.4 SciML model training

Our state-space UDE model was trained by optimizing the composite function described in Equation (7), which serves as the objective function. In training, we used the Adam optimizer [Kingma, 2014], using adaptive estimates of lower-order moments for efficient convergence in stochastic objective functions. The optimization process was implemented using the Optimizers.jl [Ma et al., 2021] library in Julia, with a learning rate of 0.03, achieving a balanced trade-off between the optimization time and the performance of the model. The loss function gradient was calculated using automatic differentiation with the Zygote.jl [Innes, 2018] package, ensuring accurate and computationally efficient updates. The activation functions, including ReLU [Nair and Hinton, 2010], CELU [Barron, 2017], softsign [Glorot and Bengio, 2010], and tanh [LeCun et al., 2002], were evaluated on subsets of observed data to capture the best performance to predict the unobserved parameters. Among these, the CELU activation function provided the best prediction performance based on the root mean square error (RMSE) metric. During simulation, the model also predicted abundance values and environmental noise, which were further evaluated for forecast capability. To effectively manage neural network parameters and predict changes in c_t, we used the Julia packages Lux.jl [Pal, 2023] and Optimization.jl [Dixit and Rackauckas, 2023]. Lux.jl facilitated the design of a scalable neural network architecture with explicit parameter handling, improving both flexibility and efficiency in model development. We performed training and evaluation 50 times randomly to ensure robustness and assess the stability of model performance across

different simulated datasets and $\boldsymbol{\theta}$ initializations. This helps quantify uncertainty in parameter estimation and forecasting skills under stochastic conditions.

2.5 Sensitivity analysis and generalization tests

To assess how general our findings our, beyond the primary training setting, we conducted three additional robustness experiments. First, we evaluated how forecast and recovery performance degrade under reduced data availability, comparing a fully observed time series ($N = 380$) against two sparse sampling scenarios ($N = 126$ and $N = 63$; Appendix B). Second, we tested sensitivity to the assumed functional form of the consumption term by comparing UDE and DE-GP across four specifications – Holling type III, constant, linear, and quadratic, to assess robustness to model misspecification (Appendix C). Third, we examine the sensitivity to values of ecological parameters, evaluating the harvest rate recovery and forecast accuracy under parameter perturbations from the baseline (Appendix D).

3 Results

3.1 Harvest rate recovery across methods

In an experimental sample run, our state-space UDE model successfully recovered the linearly increasing trend in the unobserved harvest rate (Fig. 4). In this example, the UDE closely tracked the true trajectory c_t across all three regimes, correctly crossing both critical thresholds c_1^* and c_2^* at approximately the right times.

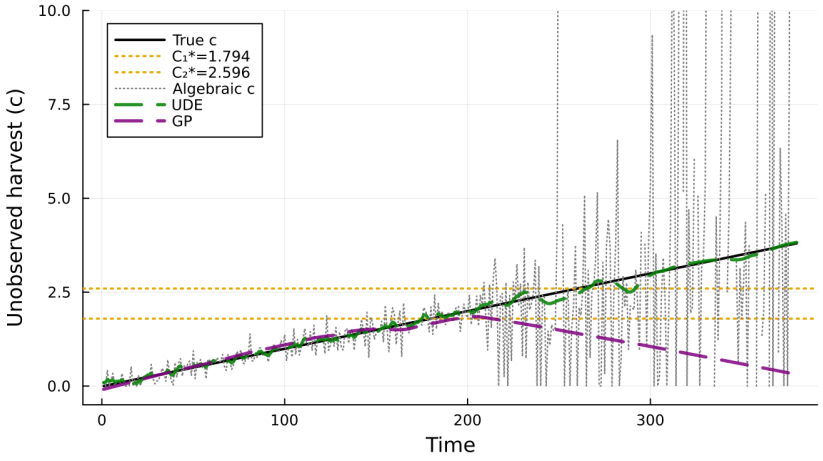

Figure 4: **Recovery of the unobserved parameter c_t from one of the random experiments.** The actual parameter values (solid black line) are compared with estimates from three methods: DE-Gaussian Process (GP; dashed purple line), and UDE (dashed green line) and algebraically-calculated harvest rate (dotted gray line). The two values of $c*$ separating regimes I, II, and III are given as horizontal intercepts (dashed yellow lines).

The DE-GP estimate followed the true trajectory reasonably well in the high abundance regime but diverged after c_1^*, failing to capture the continued rise in harvest rate and drifting toward zero late in the time series – a pattern we attribute to the kernel's limited capacity to track a non-stationary latent trend over long horizons. The algebraic inversion served as a closed-form reference for harvest rate, but because it propagates observation noise directly through the nonlinear denominator, its pointwise estimates are highly variable and unsuitable for regime classification. Taken together, this example suggests that jointly optimizing the latent trajectory and neural network weights, as in the UDE formulation, produces a more stable and ecologically-interpretable recovery of the unobserved harvest than either DE-GP or direct algebraic inversion. These observations were further examined in repeated experiments (Section 2.3), with recovery performance c_t varying systematically with the size of the training set and the dynamic regime.

In regimes I and II (Fig. 5, N=70, N=220), DE-GP outperformed UDE, both methods achieving substantially lower RMSE than the algebraic method. This indicates that a smooth functional approximation underlying DE-GP is well suited to the lower-variance, quasi-stationary conditions where the latent trend is gradual. However, in regime III (N=300), this advantage reversed: UDE outperformed both DE-GP and the algebraic method, with the flexibility of the neural network capturing the accelerating rise in c_t through bifurcation in a way that the GP could not sustain (similar

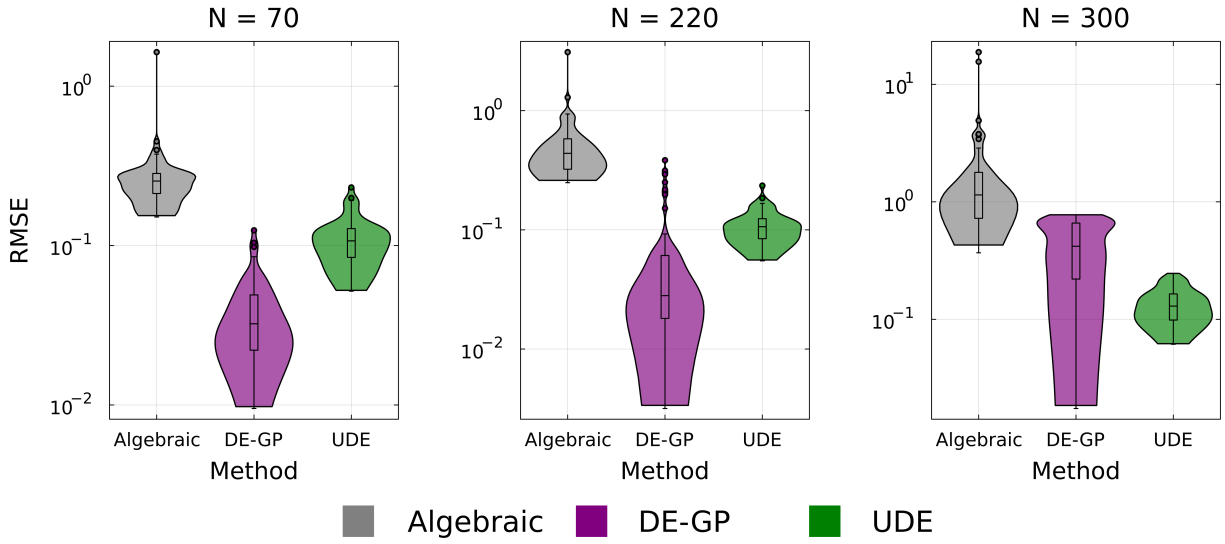

Figure 5: **RMSE distributions for predicted parameter $\hat{c}_t$ across three abundance regimes**. Results are shown for 50 random experiments across high abundance (Regime I, $N = 70$), intermediate flickering (Regime II, $N = 220$), and low abundance (Regime III, $N = 300$). Violin-boxplots display the median, quartiles, and outliers (values beyond $1.5\times$ IQR). Model performance is compared between the universal dynamic equation with a neural network (UDE, green), DE-Gaussian process (DE-GP, purple), and algebraically-calculated harvest rates (gray).

to Fig. 4). This reversal highlights a key asymmetry in model behavior across regimes and underscores the practical importance of model selection under non-stationarity.

Regime classification evaluations using F-scores corroborate this finding, further confirming that the advantage of UDE in regime III reflects structural learning of the bifurcation dynamics rather than improved point estimates of $\hat{c}_t$ alone (Fig. 6). The distributions of the F-score varied substantially between regimes: regime I produced tightly clustered high values (first quartile exceeding 0.8), regime II showed moderate dispersion, and regime III revealed a marked drop in DE-GP performance while UDE maintained the comparatively highest median F-score of $\sim$0.8. The algebraic method achieved a median F-score of $\sim$0.8 in regime I, comparable to UDE and DE-GP, but degraded substantially in regimes II and III, where it plateaued near ~ 0.6, significantly below UDE. This regime-specific pattern underscores that UDE is more robust to the classification challenge posed by nonlinear transitions and ecological degradation.

3.2 Near-term forecasts of abundance

Methods leveraging known ecological dynamics (UDE and DE-GP) produced lower forecast errors than data-driven benchmarks (ARIMA, Random Walk, LSTM) across regimes I and II, with the size of this gap depending on both the dynamical regime and the forecast horizon (Fig. 7). In regimes I and II (N=70, N=220), UDE and DE-GP achieved consistently lower RMSE values, where DE-GP showed a competitively similar median performance in the forecast of one-step and five-step horizons, while UDE showed fewer extreme errors across horizons, as reflected in the spread of the box plot (Fig. 7). The Mann-Whitney U-tests confirm that UDE significantly outperforms ARIMA, Random Walk, and LSTM at horizons 1 and 5 in regimes I and II ($p < 0.02$), indicating that process-informed methods provide a consistent advantage in forecasting in the short-to-medium range under stable and flickering conditions. Forecast errors increased with the horizon length for all methods, consistent with the high stochasticity of the system.

In regime III (N=300), the performance ordering shifted. UDE achieved the lowest RMSE at the one-step horizon (0.195 ± 0.207) and retained this advantage at the 15 step horizon (0.260 ± 1.043), while DE-GP's errors increased sharply with the length of the horizon (RMSE 0.542 ± 2.935 at 15 steps), falling below even ARIMA and Random Walk. The Mann-Whitney U-tests confirm that UDE significantly outperforms DE-GP at horizons 5, 10, and 15 in regime III ($p < 0.02$), suggesting that DE-GP overfits the training dynamics in a way that becomes unstable once the population has settled into the low-abundance state, whereas UDE maintains more reliable forecasts throughout. Overall, UDE provided the most consistent forecast accuracy across all three regimes, with particular strength at short-to-medium horizons.

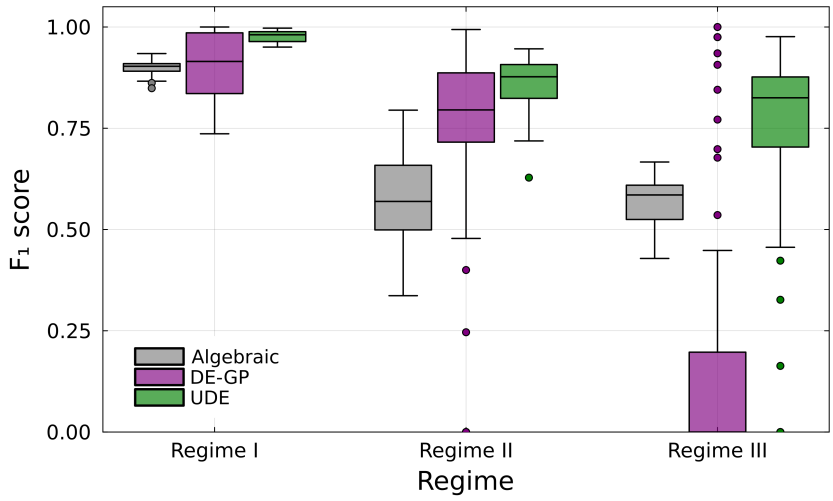

Figure 6: **Regime classification performance based on predictions of the unobserved harvesting parameter** $\hat{c}_t$. Each predicted $\hat{c}_t$ value is mapped to one of three dynamical regimes (I, II, III) defined by the critical thresholds c_1^* and c_2^*, and F-scores (higher is better) summarize classification accuracy across seeds. Box colors indicate methods: algebraic (yellow), DE-GP (purple), and UDE (green).

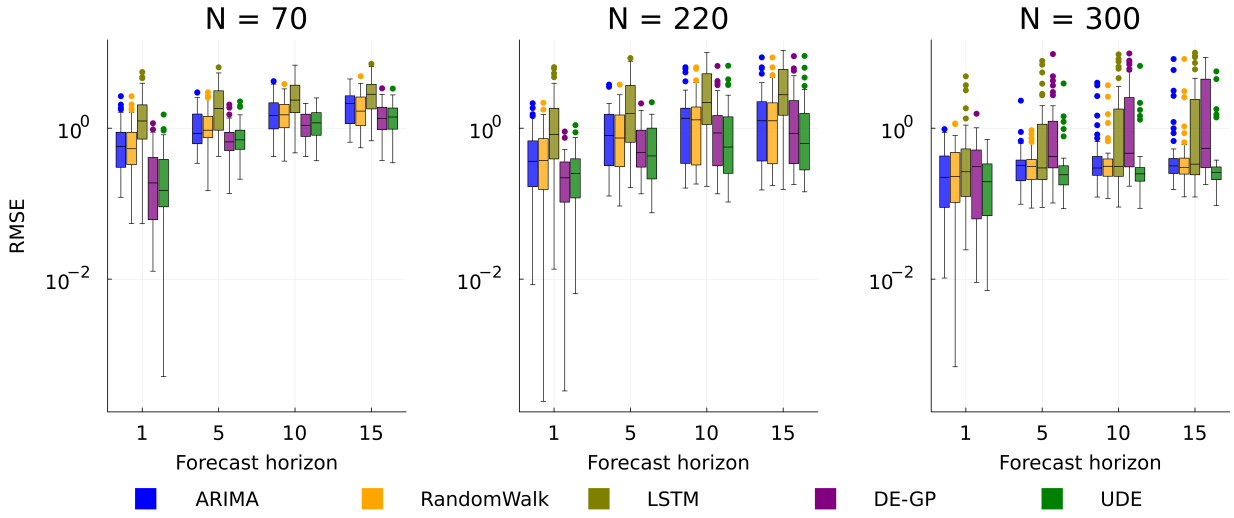

Figure 7: Comparative forecast performance using the root mean square error (RMSE) metric across different regimes and forecast horizons. Blue, yellow, olive, purple, and green boxes represent autoregressive integrated moving average (ARIMA), RandomWalk, long short-term memory (LSTM), DE-Gaussian process (DE-GP), and universal dynamic equation with a neural network (UDE) methods, respectively. Panels separate performance metric results after training on 70 data points (Regime I), 220 data points (Regime II), and 300 data points (Regime III).

3.3 Performance based on system observability

Table 1 summarizes latent parameter recovery and forecasting performance for UDE and DE-GP across the two observability scenarios introduced in the Methods 2.2.3 — fully observed (x, i) and partially observed (x only). In both observation scenarios, UDE substantially outperformed DE-GP in recovery of the latent harvest rate: RMSE-c dropped from 1.017 to 0.228 in the fully observed case and from 1.305 to 1.110 in the partially observed case, with a corresponding improvement in the F-score from 0.501 to 0.894 and from 0.225 to 0.613, respectively. UDE reduced the error in harvest rate recovery by a factor of roughly $35\times$ relative to the algebraic reference in the fully observed setting (RMSE-$c = 0.228$ vs. RMSE-c-alg$= 8.111$), and by nearly $9\times$ in the partially observed setting (RMSE-$c = 1.110$ vs. RMSE-c-alg$= 9.779$), demonstrating that jointly optimizing the latent trajectory and network weights yields substantially more reliable recovery than algebraic inversion alone. The forecast accuracy, measured by RMSE for x at horizons 1 through 15, remained comparable between the two UDE variants, suggesting that the additional difficulty of the partially observed setting penalizes latent recovery more than the prediction of short-horizon abundance. In contrast, the DE-GP forecast error grew substantially with the horizon in both settings, indicating that its latent c_t estimates do not generalize well beyond the training window under the increasing harvest rate. Finally, UDE was consistently faster to train than DE-GP by a factor of roughly 7–$8\times$, a practical advantage in high-volume simulation studies.

Table 1: Performance comparison using root mean square error (RMSE) for the harvest rate (c) and environmental state (x) at different forecast horizons (1, 5, 10, 15 time steps) and under different levels of system knowledge, i.e., fully or partially observed system. The two dynamic equation methods are compared with the better performance in bold for each metric.

Observability	Method	F-score	RMSE-c	RMSE-1 (x)	RMSE-5 (x)	RMSE-10 (x)	RMSE-15 (x)	Training Time (s)
Fully Observed (x, i)	UDE	**0.894 ± 0.08**	**0.228 ± 0.10**	**0.202 ± 0.15**	**0.237 ± 0.10**	**0.241 ± 0.08**	**0.244 ± 0.08**	**260.1 ± 15.8**
	DE-GP	0.501 ± 0.16	1.017 ± 0.28	0.365 ± 0.31	1.271 ± 1.35	3.287 ± 2.78	4.469 ± 3.11	1872.5 ± 43.2
	Algebraic		8.111 ± 17.39					
Partially Observed (x)	UDE	**0.613 ± 0.14**	**1.110 ± 0.65**	**0.219 ± 0.15**	**0.242 ± 0.10**	**0.245 ± 0.08**	**0.248 ± 0.08**	**331.6 ± 11.6**
	DE-GP	0.225 ± 0.04	1.305 ± 0.14	0.362 ± 0.30	0.970 ± 0.79	2.657 ± 2.07	3.678 ± 2.60	2761.2 ± 62.8
	Algebraic		9.779 ± 21.67					

4 Discussion

Predicting regime shifts in ecosystems where key drivers of change are unobserved remains a significant challenge across ecology and natural resource management [Dakos et al., 2015, Hamilton, 1989, Akbal, 2024]. This study demonstrates that UDEs (based on both neural networks and Gaussian Processes) can effectively recover unobserved harvest rate and identify ecological regimes from population abundance observations alone. Across training sizes and random seeds, UDE outperformed DE-GP and the purely statistical alternatives in recovering the hidden harvest rate with lower prediction errors and more reliable regime classification. By embedding known population dynamics within the model, the framework can distinguish between changes in abundance driven by harvest and those driven by environmental fluctuations, something a purely data-driven approach has no ability to do. We also evaluated how this performance holds under reduced data availability and under different assumptions about the functional form of the consumption term, finding that UDE performed well under moderate data sparsity and was more accurate when the assumed model structure matched the true system (Appendix Tables 3 and 4).

When it comes to forecasting population abundance over short horizons, UDE outperformed competing methods consistently in the flickering and low abundance stable regimes, while DE-GP outperformed in the high abundance stable regime, where smooth dynamics are easier to track without mechanistic grounding. This result is worth interpreting carefully rather than treating as a shortcoming. [Boettiger, 2022] has shown that choosing models based solely on the forecast evaluation metric can lead to a situation where the statistically superior model produces worse management outcomes in practice, because the forecast evaluation captures how well a model tracks the observable past, not how faithfully it represents what is actually driving the system. A method that closely follows recent abundance trends may still be entirely blind to whether harvest has been quietly increasing toward a critical threshold. The UDE framework, by recovering the estimated harvest along with the abundance forecast, provides information about what drives the system that no purely statistical forecast can offer. When an environmental signal was also available as an observed input, separating harvest from environmental forcing became considerably easier, improving both recovery accuracy and forecast performance (Table 1).

Taken together, our results suggest that state-space UDEs offer the greatest advantage over purely data-driven methods under three conditions: (1) when an estimated parameter (e.g., harvest rate) is the quantity of primary scientific or

management interest, rather than the state variable (e.g., abundance) alone; (2) when the system is near a regime boundary; and (3) when observations are moderately sparse. Conversely, when the goal is purely short-term abundance forecasting with dense observations, simpler data-driven approaches can match or exceed UDE performance at a lower computational cost. This pattern connects to a broader point about model selection in ecology. A random walk can achieve low prediction error on abundance simply by following recent trends, but it tells a manager nothing about whether harvest pressure has been increasing or how close the system is to a point of no return. UDEs, by explicitly recovering the estimated harvest trajectory within a known ecological framework, provide the kind of mechanistic insight that forecast accuracy alone cannot deliver.

The ability to recover unobserved harvest rates from abundance data alone has direct relevance to conservation and resource management contexts where harvest would be unreported or deliberately concealed. Unobserved drivers of population change can emerge from multiple sources: technological and measurement limitations [Rocke et al., 2003, Decorte et al., 2024], microclimate heterogeneity [Mislan and Helmuth, 2008], temporal scale mismatches between ecological processes and monitoring programs [Winkler et al., 2021], and economic constraints that limit monitoring coverage [Sparrow et al., 2020]. Such drivers can also emerge from illegal activities including clandestine fishing, poaching, logging, or other forms of extraction deliberately concealed from monitoring systems. For example, illegal, unreported, and unregulated fishing constitutes one of the most significant sources of bias in global fisheries stock assessments [Agnew et al., 2009, Pauly et al., 2002, Widjaja et al., 2023]. In all of these settings, the UDE framework could provide a principled basis for inferring the (hidden) harvest rates necessary to produce the observed abundances. For managers of social-ecological systems, the estimated harvest rates are directly interpretable in terms of known ecological thresholds, giving decision-makers actionable information about proximity to critical transitions that forecasts of abundances alone cannot provide.

Several extensions would substantially broaden the applicability of this framework. Incorporating uncertainty quantification, through ensemble methods or Bayesian approaches, would translate point estimates of harvest into probabilistic early warning signals more suitable for formal decision-making. Developing computationally efficient implementations would enable application to higher-dimensional systems with multiple interacting variables, which is currently a practical barrier. The observability experiments conducted here suggest that even a noisy environmental proxy substantially improves harvest recovery; future work could explore how proxy quality trades off against recovery accuracy in real monitoring contexts. Extensions to real fisheries data, including additional stocks beyond Northern Cod [Cook et al., 1997], would test generalizability to systems where the true harvest trajectory is never observed, requiring validation strategies based on retrospective stock assessments. Finally, robustness to model misspecification could be improved by treating selected system parameters as learnable within the UDE framework, which would reduce the degree of ecological expertise required to apply the method in new systems.

This study demonstrates that using neural network components within a known mechanistic population model can reliably recover unobserved harvest and identify ecological regimes from abundance observations alone. Applied to a simulated harvesting system, UDE substantially outperformed competing methods in estimating the hidden harvest and classifying the type of regime, while remaining competitive in short-horizon abundance forecasting. The broader significance of this result is that the UDE framework recovers not only what the system will do next, but also what the drivers are, information that is more directly useful for management intervention than forecast accuracy alone. As unobserved and unreported pressures on natural systems become increasingly consequential, methods that can infer hidden drivers within an ecologically-grounded framework offer a practical path toward more informed and timely governance of social-ecological systems.

Appendix

A Mechanistic model parameters and data pre-processing

We simulated the [Tilman et al., 2024] model to generate time series data with different random seeds, by solving the discrete state-space equation (1). We considered the harvest rate as an unobserved parameter that monotonically increases between the suggested range [0,4] in [Tilman et al., 2024]. This process provides time-series observations $\hat{x}, \hat{i}$ corresponding to the environmental state and the noise term in the system. The parameter values are listed in Table 2.

Table 2: [Tilman et al., 2024] mechanistic model parameters

Variable	Value	Description
x_t	0-20	Current environmental state (x_0 = 10.0)
i_t		Auto-correlated red noise
T	30	Timescale over which noise becomes uncorrelated
η_t	0	i.i.d. normal error term
β	0.07	The standard deviation of η's
r	1	Resource growth rate
K	10	Resource carrying capacity
c_t	0.0-4.0	Harvest rate
h	1	Harvest half-saturation constant

B Performance evaluation based on data availability

We also extended our experiments to evaluate how performance degrades under reduced data availability, summarized in Table 3, covering three sampling scenarios: fully observed ($N = 380$), and two sparse observation scenarios sub-sampling every third ($N = 126$) and every sixth ($N = 63$) time step.

UDE maintained strong harvest recovery under moderate sparsity, with RMSE-c increasing only marginally from 0.228 at full density to 0.314 at $N = 126$, and F-score remaining high at 0.843 significantly higher than algebraic method (0.201 at $N = 126$), suggesting that the joint trajectory-network optimization is robust to a threefold reduction in observations. At the more aggressive sampling scenario ($N = 63$), RMSE-c rose to 0.715 and F-score dropped to 0.620 (whereas algebraic F-score ~ 0.174), indicating that reliable regime classification becomes more challenging when fewer than one-sixth of time steps are observed, though UDE still outperformed DE-GP at this level (0.715 vs. 0.905). DE-GP performance degraded more gradually in RMSE-c across sampling scenarios ($1.017 \rightarrow 1.070 \rightarrow 0.905$), but its forecast error in longer horizons remained consistently high, with RMSE-15 near 4.5 in both the full and $N = 126$ settings, converging toward UDE only at the most aggressive sparsity, where both methods face recovery limits.

Table 3: UDE performance based on root mean square error (RMSE) for varying amounts of observed training data (N) under different sampling scenarios. Values in bold indicate higher performance between methods within sampling scenario.

	Fully observed data ($N = 380$)		**Sampled every 3 steps** ($N = 126$)		**Sampled every 6 steps** ($N = 63$)	
Method	**UDE**	**DE-GP**	**UDE**	**DE-GP**	**UDE**	**DE-GP**
F-score	$\mathbf{0.894 \pm 0.08}$	0.501 ± 0.16	$\mathbf{0.843 \pm 0.12}$	0.460 ± 0.19	$\mathbf{0.620 \pm 0.19}$	0.505 ± 0.22
RMSE-c-alg	8.111 ± 17.39	8.111 ± 17.39	3.610 ± 5.15	3.610 ± 5.15	3.411 ± 3.87	3.411 ± 3.87
RMSE-c	$\mathbf{0.228 \pm 0.10}$	1.017 ± 0.28	$\mathbf{0.314 \pm 0.17}$	1.070 ± 0.30	$\mathbf{0.715 \pm 0.34}$	0.905 ± 0.34
RMSE-1	$\mathbf{0.202 \pm 0.15}$	0.365 ± 0.31	$\mathbf{0.244 \pm 0.41}$	0.295 ± 0.27	$\mathbf{0.270 \pm 0.37}$	0.259 ± 0.23
RMSE-5	$\mathbf{0.237 \pm 0.10}$	1.271 ± 1.35	$\mathbf{0.398 \pm 1.04}$	1.425 ± 1.41	$\mathbf{0.843 \pm 1.29}$	0.876 ± 0.96
RMSE-10	$\mathbf{0.241 \pm 0.08}$	3.287 ± 2.78	$\mathbf{0.437 \pm 1.31}$	3.456 ± 2.99	$\mathbf{1.718 \pm 2.83}$	2.273 ± 2.54
RMSE-15	$\mathbf{0.244 \pm 0.08}$	4.469 ± 3.11	$\mathbf{0.449 \pm 1.37}$	4.486 ± 3.30	$\mathbf{2.096 \pm 3.29}$	3.219 ± 3.04

C Performance evaluation based on system knowledge

We further evaluated robustness to the assumed functional form of the consumption term by comparing UDE and DE-GP performance across four harvesting specifications: Holling type-III ($x^2/(x^2 + h^2)$), constant (1), linear (x), and

quadratic (x^2), with results summarized in Table 4. Although we note that in alternative harvesting term specification, the recovered term may include more than pure harvest rate, this experiment focuses on comparative assessment of recovered unobserved factors.

UDE achieved its strongest latent recovery under the Holling type-III term (RMSE-$c = 0.228$, F-score $= 0.894$), which corresponds to the true generating model, confirming that correct functional specification yields the most accurate regime identification. Under misspecified consumption terms, UDE performance degraded but remained competitive: RMSE-c ranged from 0.978 (quadratic) to 1.573 (constant), with F-score highest under the quadratic term (0.548) and lowest under the linear term (0.213), where both methods effectively collapsed to near-random regime classification. DE-GP showed a similar pattern of degradation but with consistently higher RMSE-c and lower F-score than UDE across all four terms, with the exception of the constant specification where both methods performed comparably (1.573 vs. 1.521). Forecast accuracy followed a different ordering: the linear consumption term yielded the lowest forecast error for both methods (RMSE-15 of 0.068 and 1.251 respectively), likely because the simpler functional form produces smoother abundance trajectories that are easier to extrapolate, even when latent c_t recovery is poor. Training time was broadly consistent across consumption terms for UDE, while DE-GP showed higher variance under constant specification, suggesting the sensitivity of kernel optimization to the shape of the latent trajectory.

Table 4: Comparison of UDE and DE-GP performance across different consumption terms (known dynamics) for the full time series with a fully observed system where abundance and noise (x, i) are both known. Table shows average values over experiments.

Consumption Type	Method	F-score	RMSE-c-alg	RMSE-c	RMSE-1	RMSE-5	RMSE-10	RMSE-15
Holling-III ($x^2/(x^2+h^2)$)	ARIMA	-	-	-	0.257	0.294	0.311	0.312
	RandomWalk	-	-	-	0.273	0.286	0.302	0.304
	LSTM	-	-	-	0.354	0.461	0.622	0.799
	UDE	0.894	8.111	0.228	0.202	0.237	0.241	0.244
	GP	0.501	8.111	1.017	0.365	1.271	3.287	4.469
Constant (1)	ARIMA	-	-	-	0.257	0.292	0.308	0.310
	RandomWalk	-	-	-	0.273	0.284	0.300	0.302
	LSTM	-	-	-	0.348	0.453	0.613	0.787
	UDE	0.440	8.043	1.573	0.339	0.951	1.730	2.081
	GP	0.349	8.043	1.521	0.462	2.543	4.756	5.521
Linear (x)	ARIMA	-	-	-	0.047	0.059	0.065	0.067
	RandomWalk	-	-	-	0.042	0.049	0.052	0.052
	LSTM	-	-	-	0.032	0.038	0.040	0.040
	UDE	0.213	1.089	1.262	0.039	0.052	0.061	0.068
	GP	0.212	1.089	1.750	0.056	0.163	0.662	1.251
Quadratic (x^2)	ARIMA	-	-	-	0.059	0.074	0.097	0.122
	RandomWalk	-	-	-	0.062	0.065	0.071	0.073
	LSTM	-	-	-	0.057	0.062	0.069	0.075
	UDE	0.548	0.617	0.978	0.169	0.415	0.492	0.516
	GP	0.213	0.617	1.692	0.214	0.959	1.214	1.267

D Performance evaluation based on parameter sensitivity

Latent harvest rate recovery under the same parameter misspecification scenarios is reported in Table 5, complementing the forecast results in Table 6.

Under the baseline configuration, UDE achieved strong latent recovery (RMSE-$c = 0.228$, F-score $= 0.894$), substantially outperforming DE-GP (RMSE-$c = 1.017$, F-score $= 0.501$). Misspecifying $h = 2.0$ caused a sharp deterioration in UDE latent recovery (RMSE-$c = 3.349$), while DE-GP degraded more modestly (RMSE-$c = 1.332$). This is because a larger half-saturation constant flattens the consumption response at low abundance, reducing the signal available to distinguish harvesting from growth, and making the joint trajectory optimization more sensitive to initialization. Treating h as a learnable parameter within the UDE framework would be a natural direction to reduce this sensitivity. Under $K = 20.0$, both methods struggled with regime classification (UDE F1-score $= 0.321$, DE-GP $= 0.102$), as the wider dynamic range of x_t stretches the latent space in which c_t must be recovered, effectively diluting the regime-relevant signal in the flickering regime. Incorporating K-normalized state representations or abundance-conditioned priors on c_t could help constrained recovery in high-capacity systems. At high growth rate ($r = 2.0$, $K = 1.0$), UDE and DE-GP showed poor performance (F1-score $= 0.268$ vs. 0.272), suggesting that when population dynamics are fast and occur over a narrow abundance range, the signatures of regime transitions become difficult to detect from noisy observations alone; denser sampling or longer training windows may help in such settings.

Table 5: UDE model performance comparison for recovery of harvest rate under the parameter sensitivity analysis.

UDE-Parameters	Method	RMSE-c-alg	RMSE-c	F1-score
$(r = 1.0, K = 10.0, h = 1.0)$	UDE	8.111 ± 17.39	0.228 ± 0.10	0.894 ± 0.08
	DE-GP	8.111 ± 17.39	1.017 ± 0.28	0.501 ± 0.16
$(r = 1.0, K = 10.0, h = 2.0)$	UDE	33.421 ± 69.47	3.349 ± 0.99	0.765 ± 0.13
	DE-GP	33.421 ± 69.47	1.332 ± 0.30	0.487 ± 0.19
$(r = 1.0, K = 20.0, h = 1.0)$	UDE	10.178 ± 16.83	4.057 ± 1.23	0.321 ± 0.14
	DE-GP	10.178 ± 16.83	3.538 ± 0.72	0.102 ± 0.07
$(r = 2.0, K = 1.0, h = 1.0)$	UDE	13.605 ± 28.10	2.258 ± 0.48	0.268 ± 0.12
	DE-GP	13.605 ± 28.10	1.683 ± 0.23	0.272 ± 0.15

Table 6: UDE model performance comparison for forecast accuracy at different time horizons under the parameter sensitivity analysis.

UDE-Parameters	Method	RMSE-1	RMSE-5	RMSE-10	RMSE-15
	ARIMA	0.257 ± 0.23	0.294 ± 0.15	0.311 ± 0.13	0.312 ± 0.13
	RW	0.273 ± 0.21	0.286 ± 0.14	0.302 ± 0.12	0.304 ± 0.12
	LSTM	0.354 ± 0.57	0.461 ± 0.85	0.622 ± 1.28	0.799 ± 1.66
$(r = 1.0, K = 10.0, h = 1.0)$	UDE	0.202 ± 0.15	0.237 ± 0.10	0.241 ± 0.08	0.244 ± 0.08
$(r = 1.0, K = 10.0, h = 1.0)$	DE-GP	0.365 ± 0.31	1.271 ± 1.35	3.287 ± 2.78	4.469 ± 3.11
	ARIMA	0.256 ± 0.23	0.295 ± 0.16	0.313 ± 0.13	0.313 ± 0.13
	RW	0.272 ± 0.22	0.287 ± 0.15	0.303 ± 0.12	0.304 ± 0.12
	LSTM	0.358 ± 0.58	0.467 ± 0.86	0.632 ± 1.30	0.812 ± 1.68
$(r = 1.0, K = 10.0, h = 2.0)$	UDE	0.202 ± 0.17	0.244 ± 0.11	0.254 ± 0.09	0.255 ± 0.09
$(r = 1.0, K = 10.0, h = 2.0)$	DE-GP	0.459 ± 0.43	2.765 ± 2.28	5.524 ± 3.17	6.412 ± 3.38
	ARIMA	0.257 ± 0.23	0.294 ± 0.15	0.311 ± 0.13	0.312 ± 0.13
	RW	0.273 ± 0.21	0.286 ± 0.14	0.302 ± 0.12	0.304 ± 0.12
	LSTM	0.352 ± 0.57	0.460 ± 0.85	0.622 ± 1.28	0.800 ± 1.66
$(r = 1.0, K = 20.0, h = 1.0)$	UDE	0.235 ± 0.20	0.411 ± 0.77	0.992 ± 3.39	1.510 ± 5.26
$(r = 1.0, K = 20.0, h = 1.0)$	DE-GP	0.453 ± 0.44	3.064 ± 3.55	8.895 ± 6.40	11.848 ± 7.23
	ARIMA	0.257 ± 0.23	0.294 ± 0.15	0.311 ± 0.13	0.312 ± 0.13
	RW	0.273 ± 0.21	0.286 ± 0.14	0.302 ± 0.12	0.304 ± 0.12
	LSTM	0.352 ± 0.57	0.460 ± 0.85	0.622 ± 1.28	0.800 ± 1.66
$(r = 2.0, K = 1.0, h = 1.0)$	UDE	0.499 ± 0.48	3.348 ± 2.97	4.700 ± 3.69	5.075 ± 3.92
$(r = 2.0, K = 1.0, h = 1.0)$	DE-GP	0.716 ± 0.67	5.505 ± 2.94	6.871 ± 3.33	7.200 ± 3.46

Declaration

Data Availability The code used to generate synthetic data, models, and their analysis is available in the Github repository `https://github.com/kjrathore/C_Star` .

Author contributions Conceptualization: KJR, JHB, JRW, JAE, ZDM; Software: KJR, JHB, JAE, ZDM; Formal Analysis: KJR, JHB; Original Writing Draft: KJR, JHB; Writing Review and Editing: JHB, ZDM, JAE, JRW; Funding Acquisition: JAE, JRW Supervision: JRW; Project administration: JRW

Funding This research was supported by the National Science Foundation awards #2233982 and #2233983 to JRW and LCM on Model-Enabled Machine Learning to Predict Ecosystem Regime Shifts.

Acknowledgments This paper is a product of the model-enabled machine learning for ecology working group, which includes the authors of the paper, Lisa McManus, Ariel Greiner, Nathan Fitzpatrick, Cheyenne Jarman, and Emerson Arehart, all of whom provided valuable contributions to the intellectual environment that led to this paper. We also thank Chris Rackauckas for help working with Julia Scientific Machine Learning tools and the Hawai'i Institute of Marine Biology for hosting a workshop where the ideas for this project were developed.

References

Marten Scheffer, Steve Carpenter, Jonathan A. Foley, Carl Folke, and Brian Walker. Catastrophic shifts in ecosystems. *Nature*, 413(6856):591–596, October 2001. ISSN 1476-4687. doi:10.1038/35098000.

Juan C Rocha, Garry Peterson, Örjan Bodin, and Simon Levin. Cascading regime shifts within and across scales. *Science*, 362(6421):1379–1383, 2018.

Phillip S. Levin and Christian Möllmann. Marine ecosystem regime shifts: Challenges and opportunities for ecosystem-based management. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1659):20130275, January 2015. ISSN 0962-8436. doi:10.1098/rstb.2013.0275.

Carl Folke, Steve Carpenter, Brian Walker, Marten Scheffer, Thomas Elmqvist, Lance Gunderson, and C. S. Holling. Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. *Annual Review of Ecology, Evolution, and Systematics*, 35(Volume 35, 2004):557–581, December 2004. ISSN 1543-592X, 1545-2069. doi:10.1146/annurev.ecolsys.35.021103.105711.

A. Carla Staver, Sally Archibald, and Simon A. Levin. The global extent and determinants of savanna and forest as alternative biome states. *Science*, 334(6053):230–232, 2011. doi:10.1126/science.1210465. URL `https://www.science.org/doi/abs/10.1126/science.1210465`.

Monica G Turner, W John Calder, Graeme S Cumming, Terry P Hughes, Anke Jentsch, Shannon L LaDeau, Timothy M Lenton, Bryan N Shuman, Merritt R Turetsky, Zak Ratajczak, et al. Climate change, ecosystems and abrupt change: science priorities. *Philosophical Transactions of the Royal Society B*, 375(1794):20190105, 2020.

Reinette Biggs, Garry D. Peterson, and Juan C. Rocha. The regime shifts database: a framework for analyzing regime shifts in social-ecological systems. *Ecology and Society*, 23(3), 2018. ISSN 17083087. URL `https://www.jstor.org/stable/26799133`.

Michael Dietze. Ecological forecasting. In *Ecological Forecasting*. Princeton University Press, 2017a.

Michael Dietze, Ethan P White, Antoinette Abeyta, Carl Boettiger, Nievita Bueno Watts, Cayelan C Carey, Rebecca Chaplin-Kramer, Ryan E Emanuel, SK Morgan Ernest, Renato J Figueiredo, et al. Near-term ecological forecasting for climate change action. *Nature Climate Change*, pages 1–9, 2024.

James S Clark, Steven R Carpenter, Mary Barber, Scott Collins, Andy Dobson, Jonathan A Foley, David M Lodge, Mercedes Pascual, Roger Pielke Jr, William Pizer, et al. Ecological forecasts: an emerging imperative. *science*, 293 (5530):657–660, 2001.

Ronald N Kickert, Giorgio Tonella, Alexander Simonov, and Sagar V Krupa. Predictive modeling of effects under global change. *Environmental Pollution*, 100(1-3):87–132, 1999.

Gary P Griffith and Elizabeth A Fulton. New approaches to simulating the complex interaction effects of multiple human impacts on the marine environment. *ICES Journal of Marine Science*, 71(4):764–774, 2014.

Mark C Urban, Greta Bocedi, Andrew P Hendry, J-B Mihoub, Guy Pe'er, Alex Singer, JR Bridle, LG Crozier, Luc De Meester, William Godsoe, et al. Improving the forecast for biodiversity under climate change. *Science*, 353(6304): aad8466, 2016.

Donald M Anderson, Allan D Cembella, and Gustaaf M Hallegraeff. Progress in understanding harmful algal blooms: paradigm shifts and new technologies for research, monitoring, and management. *Annual review of marine science*, 4:143–176, 2012.

John T Abatzoglou and A Park Williams. Impact of anthropogenic climate change on wildfire across western us forests. *Proceedings of the National Academy of sciences*, 113(42):11770–11775, 2016.

Christophe Fraser, Steven Riley, Roy M Anderson, and Neil M Ferguson. Factors that make an infectious disease outbreak controllable. *Proceedings of the National Academy of Sciences*, 101(16):6146–6151, 2004.

James H Stock and Mark W Watson. Dynamic factor models, factor-augmented vector autoregressions, and structural vector autoregressions in macroeconomics. *Handbook of Macroeconomics*, 2:415–525, 2016.

Liam Paninski, Yashar Ahmadian, Daniel Gil Ferreira, Shinsuke Koyama, Kamiar Rahnama Rad, Michael Vidne, Joshua Vogelstein, and Wei Wu. A new look at state-space models for neural data. *Journal of computational neuroscience*, 29(1):107–126, 2010.

Andrew R. Tilman, Elisabeth H. Krueger, Lisa C. McManus, and James R. Watson. Maintaining human wellbeing as socio-environmental systems undergo regime shifts. *Ecological Economics*, 221:108194, 7 2024. ISSN 0921-8009. doi:10.1016/J.ECOLECON.2024.108194.

Robert M May and George F Oster. Bifurcations and dynamic complexity in simple ecological models. *The American Naturalist*, 110(974):573–599, 1976.

Marten Scheffer, Jordi Bascompte, William A Brock, Victor Brovkin, Stephen R Carpenter, Vasilis Dakos, Hermann Held, Egbert H Van Nes, Max Rietkerk, and George Sugihara. Early-warning signals for critical transitions. *Nature*, 461(7260):53–59, 2009.

Reinette Biggs, Stephen R Carpenter, and William A Brock. Turning back from the brink: Detecting an impending regime shift in time to avert it. *Proceedings of the National Academy of Sciences*, 106(3):826–831, 2009.

Vasilis Dakos, Stephen R Carpenter, Egbert H van Nes, and Marten Scheffer. Resilience indicators: prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1659):20130263, 2015.

Amin Ghadami and Bogdan I Epureanu. Data-driven prediction in dynamical systems: recent developments. *Philosophical Transactions of the Royal Society A*, 380(2229):20210213, 2022.

Shirin Panahi, Ling-Wei Kong, Mohammadamin Moradi, Zheng-Meng Zhai, Bryan Glaz, Mulugeta Haile, and Ying-Cheng Lai. Machine learning prediction of tipping in complex dynamical systems. *Phys. Rev. Res.*, 6: 043194, November 2024. doi:10.1103/PhysRevResearch.6.043194. URL `https://link.aps.org/doi/10.1103/PhysRevResearch.6.043194`.

Anthony D Barnosky, Elizabeth A Hadly, Jordi Bascompte, Eric L Berlow, James H Brown, Mikael Fortelius, Wayne M Getz, John Harte, Alan Hastings, Pablo A Marquet, et al. Approaching a state shift in earth's biosphere. *Nature*, 486 (7401):52–58, 2012.

Sonia Kéfi, Vishwesha Guttal, William A Brock, Stephen R Carpenter, Aaron M Ellison, Valerie N Livina, David A Seekell, Marten Scheffer, Egbert H Van Nes, and Vasilis Dakos. Early warning signals of ecological transitions: methods for spatial patterns. *PloS one*, 9(3):e92097, 2014.

Reinette Biggs, Maja Schlüter, and Michael L Schoon. Principles for building resilience: sustaining ecosystem services in social-ecological systems. 2015.

Robert M May. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, 269(5628): 471–477, 1977a.

Michael C. Dietze. *Ecological Forecasting*. Princeton University Press, 2017b. ISBN 9780691160573. URL `http://www.jstor.org/stable/j.ctvc7796h`.

Peter Turchin. *Complex population dynamics: a theoretical/empirical synthesis (MPB-35)*. Princeton university press, 2013.

George Sugihara, Robert May, Hao Ye, Chih hao Hsieh, Ethan Deyle, Michael Fogarty, and Stephan Munch. Detecting causality in complex ecosystems. *Science*, 338(6106):496–500, 2012. doi:10.1126/science.1227079. URL `https://www.science.org/doi/abs/10.1126/science.1227079`.

Charles T Perretti, George Sugihara, and Stephan B Munch. Nonparametric forecasting outperforms parametric methods for a simulated multispecies system. *Ecology*, 94(4):794–800, 2013.

Stephan B. Munch and Antoine Brias. Empirical dynamic programming for model-free ecosystem-based management. *Methods in Ecology and Evolution*, 15(4):769–778, February 2024. ISSN 2041-210X. doi:10.1111/2041-210x.14302. URL `http://dx.doi.org/10.1111/2041-210X.14302`.

Florian Grziwotz, Chun-Wei Chang, Vasilis Dakos, Egbert H. van Nes, Markus Schwarzländer, Oliver Kamps, Martin Heßler, Isao T. Tokuda, Arndt Telschow, and Chih-Hao Hsieh. Anticipating the occurrence and type of critical transitions. *Science Advances*, 9(1):eabq4558, January 2023. ISSN 2375-2548. doi:10.1126/sciadv.abq4558.

Yong-Jin Huang, Chun-Wei Chang, and Chih-hao Hsieh. Detecting shifts in nonlinear dynamics using empirical dynamic modeling with nested-library analysis. *PLOS Computational Biology*, 20(1):1–15, January 2024. doi:10.1371/journal.pcbi.1011759.

Koushik Garain, Chao-Jui Chang, Jer-Horng Wu, Hsiao-Pei Lu, Arndt Telschow, Chun-Wei Chang, and Chih-hao Hsieh. Early warning signals to anticipate critical transitions in high-dimensional systems, November 2025. ISSN 2693-5015.

Simon N. Wood. Partially Specified Ecological Models. *Ecological Monographs*, 71(1):1–25, 2001. ISSN 1557-7015. doi:10.1890/0012-9615(2001)071[0001:PSEM]2.0.CO;2.

Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural Computation*, 9:1735–1780, 11 1997. doi:10.1162/neco.1997.9.8.1735.

Zachary C. Lipton, John Berkowitz, and Charles Elkan. A critical review of recurrent neural networks for sequence learning, 2015. URL `https://arxiv.org/abs/1506.00019`.

Charles T Perretti and Stephan B Munch. On estimating the reliability of ecological forecasts. *Journal of Theoretical Biology*, 372:211–216, 2015.

Klaus Greff, Rupesh K. Srivastava, Jan Koutnik, Bas R. Steunebrink, and Jurgen Schmidhuber. Lstm: A search space odyssey. *IEEE Transactions on Neural Networks and Learning Systems*, 28(10):2222–2232, 10 2017. ISSN 2162-2388. doi:10.1109/tnnls.2016.2582924. URL http://dx.doi.org/10.1109/TNNLS.2016.2582924.

Christopher Rackauckas, Yingbo Ma, Julius Martensen, Collin Warner, Kirill Zubov, Rohit Supekar, Dominic Skinner, Ali Ramadhan, and Alan Edelman. Universal differential equations for scientific machine learning. 1 2020. URL http://arxiv.org/abs/2001.04385.

Willem Bonnaffé, Ben C Sheldon, and Tim Coulson. Neural ordinary differential equations for ecological and evolutionary time-series analysis. *Methods in Ecology and Evolution*, 12(7):1301–1315, 2021.

Jorge Arroyo-Esquivel, Christopher A Klausmeier, and Elena Litchman. Using neural ordinary differential equations to predict complex ecological dynamics from population density data. *Journal of the Royal Society Interface*, 21(214): 20230604, 2024.

Jack H Buckner, Zechariah D Meunier, Jorge Arroyo-Esquivel, Nathan Fitzpatrick, Ariel Greiner, Lisa C McManus, and James R Watson. Recovering complex ecological dynamics from time series using state-space universal dynamic equations. *Communications Earth & Environment*, 7(1):112, 2026.

Robert M. May. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. *Nature*, 269(5628): 471–477, October 1977b. ISSN 1476-4687. doi:10.1038/269471a0.

Teri E. Nicholson, Loren McClenachan, Kisei R. Tanaka, and Kyle S. Van Houtan. Sea otter recovery buffers century-scale declines in California kelp forests. *PLOS Climate*, 3(1):1–18, January 2024. doi:10.1371/journal.pclm.0000290. URL https://doi.org/10.1371/journal.pclm.0000290. Publisher: Public Library of Science.

Linda JS Allen. *An introduction to stochastic processes with applications to biology*. CRC press, 2010.

Vasilis Dakos, Stephen R. Carpenter, William A. Brock, Aaron M. Ellison, Vishwesha Guttal, Anthony R. Ives, Sonia Kéfi, Valerie Livina, David A. Seekell, Egbert H. van Nes, and Marten Scheffer. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS ONE*, 7, 7 2012. ISSN 19326203. doi:10.1371/journal.pone.0041010.

Lenart Treven, Philippe Wenk, Florian Dörfler, and Andreas Krause. Distributional gradient matching for learning uncertain neural dynamics models. In *Proceedings of the 35th International Conference on Neural Information Processing Systems*, NIPS '21, pages 29780–29793. Curran Associates Inc. ISBN 978-1-7138-4539-3.

Audrey Olivier, Michael D. Shields, and Lori Graham-Brady. Bayesian neural networks for uncertainty quantification in data-driven materials modeling. *Computer Methods in Applied Mechanics and Engineering*, 386:114079, December 2021. ISSN 0045-7825. doi:10.1016/j.cma.2021.114079. URL https://www.sciencedirect.com/science/article/pii/S0045782521004102.

Lucas P. Medeiros, Darian K. Sorenson, Bethany J. Johnson, Eric P. Palkovacs, and Stephan B. Munch. Revealing unseen dynamical regimes of ecosystems from population time-series data. *Proceedings of the National Academy of Sciences*, 122(24):e2416637122, 2025. doi:10.1073/pnas.2416637122. URL https://www.pnas.org/doi/abs/10.1073/pnas.2416637122.

M. Raissi, P. Perdikaris, and G.E. Karniadakis. Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations. *Journal of Computational Physics*, 378:686–707, 2019. ISSN 0021-9991. doi:https://doi.org/10.1016/j.jcp.2018.10.045. URL https://www.sciencedirect.com/science/article/pii/S0021999118307125.

Maziar Raissi, Paris Perdikaris, and George Em Karniadakis. Machine learning of linear differential equations using gaussian processes. *J. Comput. Phys.*, 348:683–693, November 2017. ISSN 0021-9991. doi:10.1016/j.jcp.2017.07.050. URL https://doi.org/10.1016/j.jcp.2017.07.050.

Liu Yang, Dongkun Zhang, and George Em Karniadakis. Physics-informed generative adversarial networks for stochastic differential equations. *SIAM Journal on Scientific Computing*, 42(1):A292–A317, 2020. doi:10.1137/18M1225409. URL https://doi.org/10.1137/18M1225409.

Arno Solin and Simo Särkkä. Hilbert space methods for reduced-rank gaussian process regression. *Statistics and Computing*, 30(2):419–446, 2020.

Jaehoon Lee, Yasaman Bahri, Roman Novak, Samuel S. Schoenholz, Jeffrey Pennington, and Jascha Narain Sohl-Dickstein. Deep neural networks as gaussian processes. *ArXiv*, abs/1711.00165, 2017. URL https://api.semanticscholar.org/CorpusID:3708505.

Ryan Turner, Marc Deisenroth, and Carl Rasmussen. State-space inference and learning with gaussian processes. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 868–875. JMLR Workshop and Conference Proceedings, 2010.

David Widmann, Will Tebbutt, st, Théo Galy-Fajou, Sharan Yalburgi, Hong Ge, Simone Carlo Surace, david vicente, Jacob Vaverka, John Skovbekk, Nathanael Bosch, Niklas Schmitz, Ross Viljoen, Tom Wright, Vikram, and andreaskoher. Juliagaussianprocesses/abstractgps.jl: v0.5.21, March 2024. URL https://doi.org/10.5281/zenodo.10801764.

Théo Galy-Fajou, David Widmann, Sharan Yalburgi, Will Tebbutt, st, Isak Falk, Simone Carlo Surace, Steffen Ridderbusch, Tom Wright, Hong Ge, Sebastian Khan, Pietro Monticone, Letif Mones, david vicente, Johannes Giersdorf, Julia TagBot, Ross Viljoen, Simon Schölly, Tor Erlend Fjelde, and Kaan Öcal. JuliaGaussianProcesses/KernelFunctions.jl: v0.10.64, July 2024. URL https://doi.org/10.5281/zenodo.12793559.

James Durbin and Siem Jan Koopman. *Time series analysis by state space methods*, volume 38. OUP Oxford, 2012.

Diederik P Kingma. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.

Yingbo Ma, Vaibhav Dixit, Michael J Innes, Xingjian Guo, and Chris Rackauckas. A comparison of automatic differentiation and continuous sensitivity analysis for derivatives of differential equation solutions. In *2021 IEEE High Performance Extreme Computing Conference (HPEC)*, pages 1–9. IEEE, 2021.

Mike Innes. Don't unroll adjoint: Differentiating ssa-form programs. *arXiv preprint arXiv:1810.07951*, 2018.

Vinod Nair and Geoffrey E. Hinton. Rectified linear units improve restricted boltzmann machines. In *Proceedings of the 27th International Conference on International Conference on Machine Learning*, ICML'10, pages 807–814, Madison, WI, USA, June 2010. Omnipress. ISBN 978-1-60558-907-7.

Jonathan T. Barron. Continuously Differentiable Exponential Linear Units, 4 2017. URL http://arxiv.org/abs/1704.07483. arXiv:1704.07483.

Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the thirteenth international conference on artificial intelligence and statistics*, pages 249–256. JMLR Workshop and Conference Proceedings, 2010.

Yann LeCun, Léon Bottou, Genevieve B Orr, and Klaus-Robert Müller. Efficient backprop. In *Neural networks: Tricks of the trade*, pages 9–50. Springer, 2002.

Avik Pal. Lux: Explicit Parameterization of Deep Neural Networks in Julia, April 2023. URL https://doi.org/10.5281/zenodo.7808904. If you use this software, please cite it as below.

Vaibhav Kumar Dixit and Christopher Rackauckas. Optimization.jl: A unified optimization package, March 2023. URL https://doi.org/10.5281/zenodo.7738525.

James D Hamilton. A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica: Journal of the econometric society*, pages 357–384, 1989.

Omer Faruk Akbal. Regime-switching factor models and nowcasting with big data. *IMF Working Papers*, 2024(190): A001, 2024. doi:10.5089/9798400286407.001.A001. URL https://www.elibrary.imf.org/view/journals/001/2024/190/article-A001-en.xml.

Carl Boettiger. The forecast trap. *Ecology Letters*, 25(7):1655–1664, 2022.

David M. Rocke, Blythe Durbin, Machelle Wilson, and Henry D. Kahn. Modeling uncertainty in the measurement of low-level analytes in environmental analysis. *Ecotoxicology and Environmental Safety*, 56(1):78–92, 2003. ISSN 0147-6513. doi:https://doi.org/10.1016/S0147-6513(03)00052-6. URL https://www.sciencedirect.com/science/article/pii/S0147651303000526.

Thomas Decorte, Steven Mortier, Jonas J Lembrechts, Filip JR Meysman, Steven Latré, Erik Mannens, and Tim Verdonck. Missing value imputation of wireless sensor data for environmental monitoring. *Sensors*, 24(8):2416, 2024.

K. A. S. Mislan and B. Helmuth. Microclimate. In Sven Erik Jørgensen and Brian D. Fath, editors, *Encyclopedia of Ecology*, pages 2389–2393. Academic Press, Oxford, January 2008. ISBN 978-0-08-045405-4. doi:10.1016/B978-008045405-4.00520-6. URL https://www.sciencedirect.com/science/article/pii/B9780080454054005206.

Klara J. Winkler, Marie C. Dade, and Jesse T. Rieb. Mismatches in the Ecosystem Services Literature—a Review of Spatial, Temporal, and Functional-Conceptual Mismatches. *Current Landscape Ecology Reports*, 6(2):23–34, June 2021. ISSN 2364-494X. doi:10.1007/s40823-021-00063-2. URL https://doi.org/10.1007/s40823-021-00063-2.

Ben D. Sparrow, Will Edwards, Samantha E.M. Munroe, Glenda M. Wardle, Greg R. Guerin, Jean-Francois Bastin, Beryl Morris, Rebekah Christensen, Stuart Phinn, and Andrew J. Lowe. Effective ecosystem monitoring requires a multi-scaled approach. *Biological Reviews of the Cambridge Philosophical Society*, 95(6):1706–1719, December 2020. ISSN 1464-7931. doi:10.1111/brv.12636. URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7689690/.

David J. Agnew, John Pearce, Ganapathiraju Pramod, Tom Peatman, Reg Watson, John R. Beddington, and Tony J. Pitcher. Estimating the Worldwide Extent of Illegal Fishing. *PLOS ONE*, 4(2):e4570, February 2009. ISSN 1932-6203. doi:10.1371/journal.pone.0004570.

Daniel Pauly, Villy Christensen, Sylvie Guénette, Tony J. Pitcher, U. Rashid Sumaila, Carl J. Walters, R. Watson, and Dirk Zeller. Towards sustainability in world fisheries. *Nature*, 418(6898):689–695, August 2002. ISSN 1476-4687. doi:10.1038/nature01017.

Sjarief Widjaja, Tony Long, Hassan Wirajuda, Hennie Van As, Per Erik Bergh, Annie Brett, Duncan Copeland, Miriam Fernandez, Ahmad Gusman, Stephanie Juwana, Toni Ruchimat, Steve Trent, and Chris Wilcox. Illegal, Unreported and Unregulated Fishing and Associated Drivers. In Jane Lubchenco and Peter M. Haugan, editors, *The Blue Compendium: From Knowledge to Action for a Sustainable Ocean Economy*, pages 553–591. Springer International Publishing, Cham, 2023. ISBN 978-3-031-16277-0. doi:10.1007/978-3-031-16277-0_15.

R. M. Cook, A. Sinclair, and G. Stefánsson. Potential collapse of North Sea cod stocks. *Nature*, 385(6616):521–522, February 1997. ISSN 1476-4687. doi:10.1038/385521a0.