



Revealing unseen dynamical regimes of ecosystems from population time-series data

Lucas P. Medeiros^{a,b,1} , Darian K. Sorenson^c , Bethany J. Johnson^{d,e} , Eric P. Palkovacs^{a,f} , and Stephan B. Munch^{d,g,1}

Affiliations are included on p. 9.

Edited by Nils Stenseth, Universitetet i Oslo, Oslo, Norway; received August 15, 2024; accepted May 11, 2025

Many dynamical systems can exist in alternative regimes for which small changes in an environmental driver can cause sudden jumps between regimes. In ecology, predicting the regime of population fluctuations under unobserved levels of an environmental driver has remained an unsolved challenge with important implications for conservation and management. Here, we show that integrating time-series data and information on a putative driver into a Gaussian Process regression model for the system's dynamics allows us to predict dynamical regimes without the need to specify the equations of motion of the system. As a proof of concept, we demonstrate that we can accurately predict fixed-point, cyclic, or chaotic dynamics under unseen levels of a control parameter for a range of simulated population dynamics models. For a model with an abrupt population collapse, we show that our approach goes beyond an early warning signal by characterizing the regime that follows the tipping point. We then apply our approach to data from an experimental microbial food web and from a lake planktonic food web. We find that we can reconstruct transitions away from chaos in the microbial food web and anticipate the dynamics of the oligotrophic regime in the planktonic food web. These results lay the groundwork for making rational decisions about preventing, or preparing for, regime shifts in natural ecosystems and other dynamical systems.

regime shifts | population dynamics | empirical dynamic modeling | early warning signal | tipping points

Dynamical systems in physics, climate, and ecology can exist in several distinct alternative regimes, and rapid shifts from one regime to another may result from gradual changes in some environmental driver or “control parameter” (1–3). In ecology, well-documented examples of regime shifts include clear macrophyte-dominated lakes transitioning into turbid phytoplankton-dominated lakes due to nutrient input (4), kelp forests becoming sea urchin barrens due to predator loss (5), coral reefs giving way to macroalgae reefs as a result of warming (6), and forests shifting to savannas after changes in precipitation (7). Because ecosystem services depend greatly on the state of the system, there has been considerable interest in understanding and predicting the dynamical regimes that may emerge under changing environmental conditions.

Predicting regime shifts in time to avert them has been an important avenue of research and many Early Warning Signals (EWS) have been developed to do so (1). These signals can be computed directly from time-series data and include variance, skewness, the AR(1) coefficient, and the power spectrum (8). The unifying idea behind EWS is that changes in the control parameter cause the system to cross a tipping point (i.e., a bifurcation) and most are based on “critical slowing down”—in which the stability of a fixed point relaxes as the bifurcation is approached (9). Empirical evidence for critical slowing down has been found in plankton in lakes (10), photoinhibition in cyanobacterial populations (11), declining food sources for *Daphnia magna* (12), and increased mortality in yeast (13). A key feature of EWS is that they are generic, that is, the dynamical behavior close to a bifurcation is well described by a handful of canonical “normal forms” (14) such that we do not need detailed knowledge of the underlying dynamics to anticipate a shift. However, EWS pertain to a relatively narrow range of bifurcations—namely the fold, transcritical, and Hopf bifurcations (3, 15)—and approaches that apply to a wider range of cases are currently needed. Most importantly, EWS only tell us that the system might be approaching a regime shift but give no information about what the upcoming regime will be.

Predicting the dynamics of ecosystems under unobserved values of an environmental driver, such as beyond a proposed tipping point, has been a major challenge in ecology. Although recent progress has been made in anticipating the composition and structure of certain ecosystems under unseen conditions (16, 17), taking the next step and predicting

Significance

Small changes in conditions can cause dynamical systems to suddenly shift to new, potentially undesirable, regimes. However, characterizing previously unseen dynamical regimes—a critical bit of information in evaluating the utility of costly interventions—remains an open challenge. Here, we focus on ecosystems and introduce a data-driven modeling approach that can extrapolate the effect of environmental conditions on population growth, allowing us to predict how populations will fluctuate under unobserved conditions. We test our approach with simulations and apply it to empirical data from microbial and planktonic food webs. By anticipating the dynamics that follows a tipping point (e.g., plankton fluctuations after a lake transition), we illustrate how our approach can facilitate decision-making in ecosystems with multiple regimes.

Author contributions: L.P.M., D.K.S., B.J.J., E.P.P., and S.B.M. designed research; S.B.M. developed the method; L.P.M. and D.K.S. performed research; L.P.M. and D.K.S. analyzed data; D.K.S., B.J.J., and E.P.P. revised the manuscript; and L.P.M. and S.B.M. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission.

Copyright © 2025 the Author(s). Published by PNAS. This open access article is distributed under Creative Commons Attribution-NonCommercial-NoDerivatives License 4.0 (CC BY-NC-ND).

¹To whom correspondence may be addressed. Email: lucas.medeiros@whoi.edu or smunch@ucsc.edu.

This article contains supporting information online at <https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2416637122/-DCSupplemental>.

Published June 13, 2025.

population fluctuations remains notoriously difficult. The difficulty stems from accurately extrapolating the effect of the driver on species abundances using data on few species and over a limited range of the driver. Unfortunately, most EWS are unable to perform this extrapolation (18, 19) and the few approaches that do [e.g., approaches leveraging deep learning (20)] require prior knowledge of the class of bifurcations. Although model-based approaches provide more specific predictions (3, 21), these require detailed knowledge about the underlying dynamics and are unlikely to be transferable across ecosystems. Therefore, an alternative approach, with generality similar to that of existing EWS, but that also predicts the dynamics beyond a tipping point is highly desirable.

Empirical Dynamic Modeling (EDM) is a framework based on nonparametric function approximation and time-delay embedding that has proved useful to understand and forecast a wide range of dynamical systems. Examples include forecasting recruitment in fisheries (22), understanding the environmental drivers of diseases (23), inferring causal interaction networks (24), and quantifying population responses to perturbations (25). By learning about system dynamics directly from time-series data (through nonparametric function approximation) and by accounting for unobserved state variables (through time-delay embedding), EDM can give insights about the dynamics without the need to specify a model, as with EWS. Several studies have used EDM to develop new EWS, which include approaches based on prediction accuracy (26), nonlinearity (27), and eigenvalues of the Jacobian matrix (28). Other recent studies have used EDM to detect whether and when a regime shift has happened in the past based on prediction accuracy tests (29, 30). What is currently lacking from both EWS and EDM approaches is the capacity to predict dynamics at novel driver levels. Yet, being able to leverage information on putative drivers (e.g., temperature, nutrients, rainfall, fishing pressure) could allow us to anticipate unseen regimes and at the same time advance our understanding about a given system.

Here we introduce an approach based on EDM that allows us to predict the dynamical behavior of an ecosystem at unobserved levels of an environmental driver. Our approach works by first integrating population time-series data and information on the putative driver into a Gaussian Process regression with time-delay embedding [GP-EDM, (31)] to learn how the driver affects population dynamics. We then predict the dynamics at unobserved levels of the driver by interpolating or extrapolating the GP-EDM model. Using several population dynamics models, we demonstrate that we can accurately predict unseen fixed-point, cyclic, or chaotic regimes, in addition to predicting the location of multiple tipping points in a bifurcation diagram. We also show that, for a model with an abrupt population collapse, our approach can anticipate the regime that follows the tipping point. Unsurprisingly, the accuracy of these predictions relies on informative training data and declines with long-range extrapolation. Finally, we apply our approach to empirical time series from both an experimental microbial and a lake planktonic food web. We show that we can reconstruct transitions away from chaos in the experimental food web and anticipate the dynamics of the oligotrophic regime in the lake food web.

Results

Testing the Approach on Model-Generated Data. We first tested the ability of our approach to predict unseen dynamical regimes and the location of tipping points using four discrete-time population dynamics models (*SI Appendix*, section S1). Here,

we focus on a two-species predator–prey model and detailed results for other models can be found in *SI Appendix*. For each model, we trained a GP-EDM model [*SI Appendix*, section S3; (31)] using time series from a single species (x_i) generated at four different levels of an environmental driver (i.e., a control parameter p ; Fig. 1A and *SI Appendix*, Fig. S1). The trained GP-EDM models produced highly accurate leave-one-out predictions under the observed data (R^2 : single-species model, 0.992; two-species competition model with harvesting, 0.998; two-species predator–prey model, 0.979; three-species competition model, 0.954; *SI Appendix*, Fig. S2). Note that we use R^2 here to indicate the mean squared error of leave-one-out predictions divided by the total time-series variance (Eq. 4) and that model selection based on out-of-sample predictions is standard in EDM (32–34). The chosen embedding dimension ($E + 1$) for each model is shown in *SI Appendix*, Fig. S2 and plots of model residuals in *SI Appendix*, Fig. S3. Then, using the delay-embedding map G_i (Eq. 3) approximated via GP-EDM (Fig. 1B), we performed predictions of x_i over unobserved values of p to reconstruct the bifurcation diagram (Fig. 1 C and D and *SI Appendix*, Fig. S1). Although G_i clearly has a different shape than the true map (F_i , Eq. 2), Fig. 1B shows that these functions change in analogous ways as p is increased. In *SI Appendix*, Fig. S1, we illustrate the case where G_i and F_i have the same inputs and, therefore, are more directly comparable to each other.

We were able to accurately predict the dynamics of x_i across many unobserved levels of p for all four models (Fig. 2 and *SI Appendix*, Fig. S4). For instance, having observed the prey time series at four regimes of the two-species predator–prey model (fixed points at $p = 2.4$ and $p = 2.6$, cycle at $p = 3.15$, and chaos at $p = 3.4$), we accurately predicted unobserved fixed-point, cyclic, or chaotic regimes (Fig. 2A). We assessed the reconstructed bifurcation diagram using the Jensen–Shannon divergence (JSD) and found that the distribution of predicted abundances of species i ($\{\hat{x}_i(t)\}$, $t = 1, \dots, T$) for each p significantly matched the distribution of true abundances ($\{x_i(t)\}$, $t = 1, \dots, T$; Fig. 2B and *SI Appendix*, Fig. S4). Note that an accurate reconstruction of the dynamics of species i does not imply an accurate reconstruction of the dynamics of other unobserved species. Nevertheless, we obtained similar results when changing the species used to perform time-delay embedding (*SI Appendix*, Fig. S5). As expected, our approach was most accurate when we used all species in the GP-EDM model (i.e., native coordinates; *SI Appendix*, Fig. S6) instead of using only a single species with time-delay embedding (i.e., delay coordinates; *SI Appendix*, Fig. S4). In addition to low JSD, predicted regimes showed dynamics with a similar time scale (*SI Appendix*, Fig. S7), periodicity (*SI Appendix*, Fig. S8), and stability (*SI Appendix*, section S4 and Fig. S8) as the true regimes. Although with slightly less accuracy, we also detected the location of tipping points (i.e., bifurcations) that separate different regimes. This includes the location of a Hopf bifurcation (Fig. 2), period doubling bifurcations (*SI Appendix*, Fig. S4 A and D), and a fold bifurcation associated with population collapse (*SI Appendix*, Fig. S4B).

We evaluated the sensitivity of our results by performing the above analysis for different dynamical regimes in our training data (*SI Appendix*, Fig. S9). We found that we can most accurately predict unseen regimes and detect the location of tipping points when the training data includes a chaotic regime, which provides information for the GP-EDM model across the entire range of x_i . We also performed these analyses for several different training datasets with different amounts of process noise. Although high noise can make our approach less precise, we found that, on average, the JSD between true and predicted

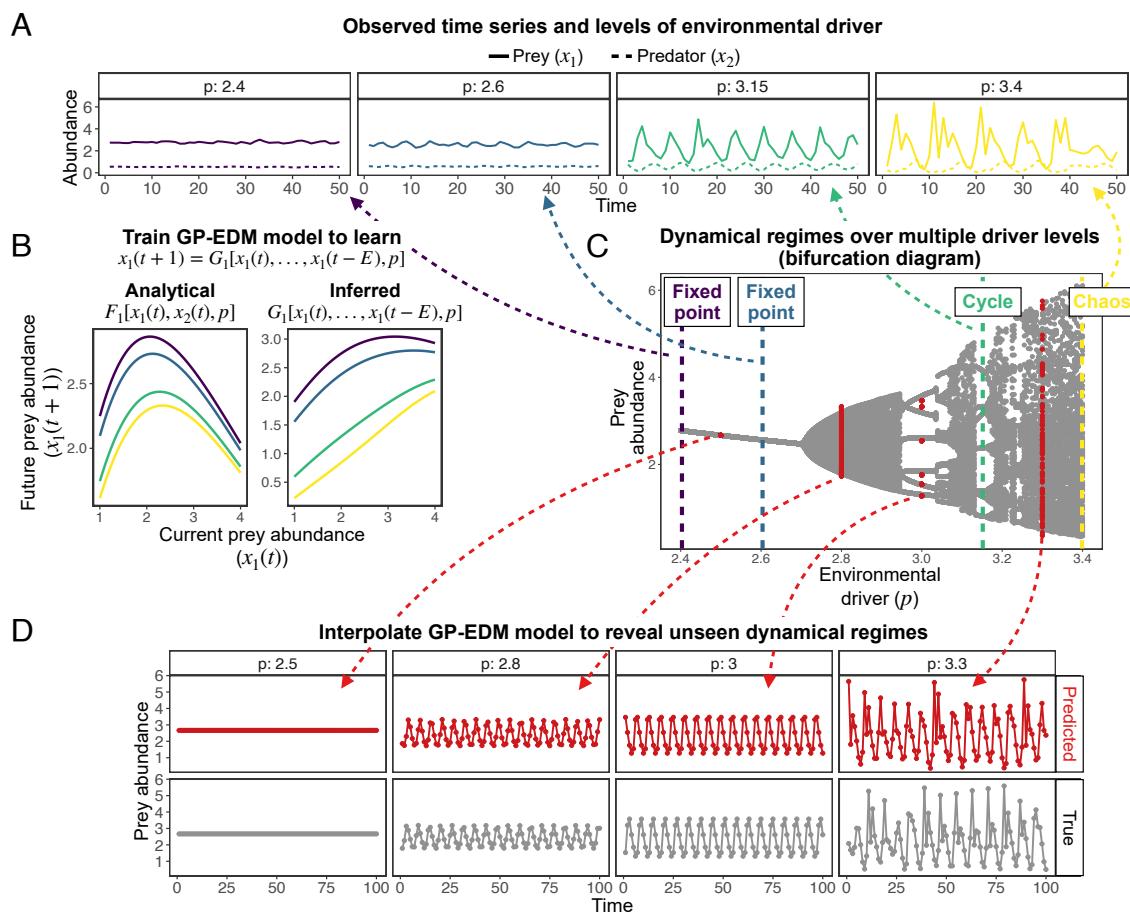


Fig. 1. GP-EDM approach to reveal unseen dynamical regimes. (A) Population time series generated from a predator-prey model (*SI Appendix, section S1*) at four different levels of the attack rate parameter (i.e., the environmental driver p). Low values of p generate fixed points ($p = 2.4$ in purple and $p = 2.6$ in blue) whereas high values of p generate cycles ($p = 3.15$ in green) or chaos ($p = 3.4$ in yellow). (B) Left panel depicts the true relationship between current ($x_1(t)$) and future ($x_1(t+1)$) prey abundance at different levels of p (different colors). Right panel depicts the same relationship but obtained from the trained GP-EDM model. Note that $x_2(t)$ is also an input of function F_1 and $x_1(t-1), x_1(t-2)$, and $x_1(t-3)$ are also inputs of function G_1 , but these inputs are held constant here for simplicity. (C) Bifurcation diagram of the predator-prey model showing the different dynamical regimes that emerge as we vary p (gray points). The four regimes shown in (A) are represented as vertical dashed lines. We can then interpolate the trained GP-EDM model to unseen levels of p , which are shown in red. (D) Predicted dynamical regimes at unseen levels of p (in red) match closely the true dynamics (in gray). These regimes are fixed point ($p = 2.5$), low-amplitude cycle ($p = 2.8$), high-amplitude cycle ($p = 3.0$), and chaos ($p = 3.3$). Note that the correct comparison is between the red and gray dynamical regimes and not exact trajectories, given that initial conditions are arbitrary.

dynamics was low (*SI Appendix, Fig. S10*). Finally, we obtained an accurate reconstruction of the bifurcation diagram when using a continuous-time model (*SI Appendix, section S2* and *Fig. S11*).

We then tested whether our approach can be used as an EWS of an upcoming population collapse using a two-species competition model with harvesting (*SI Appendix, section S1*). To do so, we generated a time series for which the harvest rate (i.e., the control parameter p) changed linearly from 0.42 to 0.48 and a tipping point occurred at $p = 0.464$ according to a change-point analysis (Fig. 3*A*). We then sequentially updated our GP-EDM model as the harvest rate increased toward the tipping point and, for each update, we extrapolated the entire bifurcation diagram (i.e., all dynamical regimes from $p = 0.42$ to $p = 0.48$). Again, the best trained GP-EDM model produced highly accurate leave-one-out predictions ($R^2 = 0.986$).

Our approach was able to anticipate population collapse when the training data ended 4 time steps before the tipping point (Fig. 3*B*). Although the predicted abundance at the unseen regime was higher than the true abundance, our approach clearly anticipated an abrupt abundance shift. The harvest rate of population collapse was predicted to be $p = 0.4608$, which is slightly below the tipping point location (i.e., $p = 0.464$). When

the training data ended exactly at the tipping point, our approach predicted a more accurate tipping point location at $p = 0.465$. Finally, when the training data contained the collapsed regime, our approach provided a very accurate reconstruction of the entire bifurcation diagram, as expected. Note that the tipping point location ($p = 0.464$) will be different for every simulation due to process noise. We obtained similar although slightly less accurate results when using training data with tipping points at different locations (*SI Appendix, Fig. S12*). Finally, the rolling variance of abundance only started to increase after the tipping point (Fig. 3*A*), suggesting a limited use of this EWS under this particular model.

Applying the Approach to an Experimental Microbial Ecosystem. Next, we applied our approach to population time-series data from a seminal chemostat experiment that discovered transitions from fixed point to chaos, then to cycles (p ; i.e., mortality) was reduced (35). This dataset is analogous to our simulation set up from Fig. 2 and, therefore, we used the same analytical approach (*Materials and Methods*).

For each of the three species in this experimental food web, we obtained leave-one-out predictions that are 49 to 67% more

Two-species predator-prey model

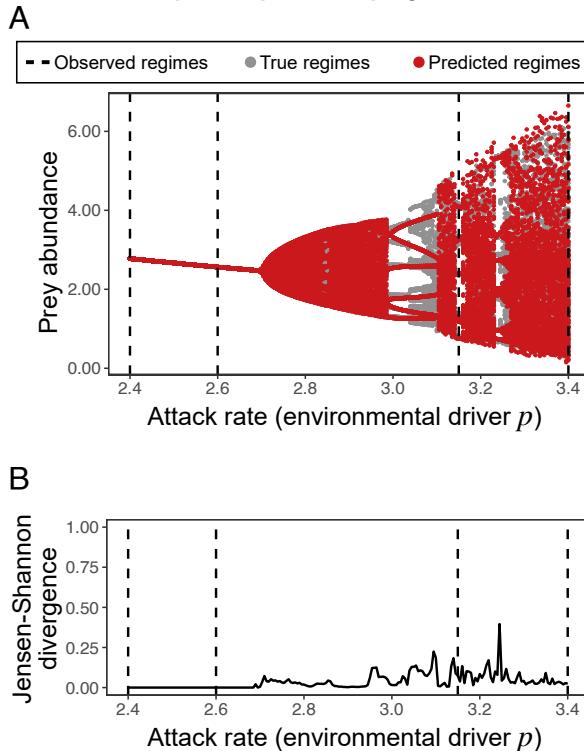


Fig. 2. GP-EDM approach accurately predicts a range of unseen dynamical regimes. (A) True (in gray) and predicted (in red) bifurcation diagram for a two-species predator-prey population dynamics model (*SI Appendix*, section S1). The bifurcation diagram depicts the abundance of the prey species after the transient period (y-axis) for a range of control parameter values (p , x-axis). (B) Jensen-Shannon divergence (JSD) computed at each value of p . The JSD ranges from 0 to 1, where 0 represents a close match between true and predicted dynamics. The average JSD across all values of p is 0.037. In (A and B) the vertical dashed lines depict the values of p for which we observed population time series to train the GP-EDM model (as in Fig. 1C).

accurate than simply using their average abundance as the prediction (R^2 : predator, 0.490; preferred prey, 0.508; less-preferred prey, 0.670; *SI Appendix*, Fig. S13). The chosen embedding dimension ($E + 1$) for each species is shown in *SI Appendix*, Fig. S13 and plots of model residuals in *SI Appendix*, Fig. S14. This suggests that GP-EDM is reasonably approximating the function G_i (Eq. 3) at the observed dilution rates. More importantly, the bifurcation diagrams reconstructed independently from each species exhibit a large region of chaos around $p = 0.5/\text{day}$, cycles around $p = 0.45/\text{day}$ and slightly above $p = 0.6/\text{day}$, and a fixed point for $p > 0.7/\text{day}$ (Fig. 4). Our results also reveal the dilution rate levels at which we should observe unseen bifurcations. For example, the appearance of a fixed point is predicted to occur somewhere between $p = 0.62/\text{day}$ and $p = 0.68/\text{day}$ (Fig. 4). Although we cannot verify this experimentally, the fact that similar results were obtained independently for each species increases our confidence that the approach accurately illuminates previously invisible regimes and bifurcations.

We evaluated whether these results were robust to how we scaled species abundances and assigned prior information to p as a model input. Results for predator and preferred prey species were robust to these changes, whereas results for the less-preferred prey species were not (*SI Appendix*, Fig. S15). We also excluded a given dilution rate from our training data and repeated our analyses to verify whether we could predict the regime under the

excluded dilution rate. We could accurately predict the fixed-point and cyclic regimes of all species (i.e., $p = 0.45$, $p = 0.75$, and $p = 0.9$), but not the chaotic regime (i.e., $p = 0.5$; *SI Appendix*, Fig. S16). Hence, similar to our model results, we found that regimes that explore the entire state space (whether through chaos or process noise) are the most informative pieces of the training data.

Applying the Approach to a Lake Planktonic Ecosystem. For our last set of analyses, we applied the EWS version of our approach to the well-known case of lakes that shift from an eutrophic to an oligotrophic regime. We used monthly time-series data from Lake Zurich of 13 plankton functional groups, phosphate concentration (i.e., the control parameter p), and water temperature from January, 1978 to December, 2019 [*SI Appendix*, Fig. S17; $T = 504$ data points; (16)]. Because this planktonic food web is thought to have shifted from an eutrophic to an oligotrophic regime via a gradual reduction of phosphate concentration (16), it is analogous to our simulations from Fig. 3.

We first performed a change-point analysis to determine the location of a potential tipping point for each functional group. We found that large green algae had the most well-resolved shift followed by small cryptophytes and omnivores (*SI Appendix*, Table S1). This potential tipping point for large green algae occurred in November, 1987 ($t_{\text{split}} = 119$) at a phosphate concentration of 0.0472 mg/l (Fig. 5A). The potential tipping points for small cryptophytes (October, 1988; $t_{\text{split}} = 130$) and omnivores (June, 1986; $t_{\text{split}} = 102$) occurred around the same time, suggesting an ecosystem-wide shift.

To separate the effects of the control parameter from seasonality, we first applied a centered moving average smoother to the phosphate time series, removing seasonal oscillations while retaining the trend. We included seasonality in our GP-EDM model by adding water temperature as an input. We then used data from January, 1978 to October, 1986 (106 data points) to train GP-EDM models and verify which functional groups were most predictable. Small green algae showed the best ($R^2 = 0.801$), whereas large green algae showed the second-best ($R^2 = 0.770$; *SI Appendix*, Fig. S18) leave-one-out prediction accuracy (*SI Appendix*, Table S2). Due to strong evidence of a tipping point for large green algae from the change-point analysis, we focused subsequent analyses on this functional group. The chosen embedding dimension and plots of model residuals for large green algae are shown in *SI Appendix*, Table S2 and Fig. S18. The GP-EDM model suggested that higher phosphate concentrations increase the nonlinearity in density dependence of large green algae (*SI Appendix*, Fig. S19).

We then reconstructed bifurcation diagrams for large green algae using successively larger training datasets (Fig. 5A). We found evidence of a previously unseen low abundance regime when the training data ended 1 y before the potential tipping point (i.e., training data up to October, 1986; Fig. 5B). We then clearly predicted an abrupt shift in the dynamics of large green algae when the training data ended 1 mo before the potential tipping point (i.e., training data up to October, 1987; Fig. 5B). This abrupt shift was predicted to occur at a phosphate concentration of 0.0470 mg/l, very close to the phosphate concentration at the potential tipping point detected via change-point analysis. Interestingly, the bifurcation diagram reconstructed with training data up to October, 1987 was very similar to the bifurcation diagram reconstructed with

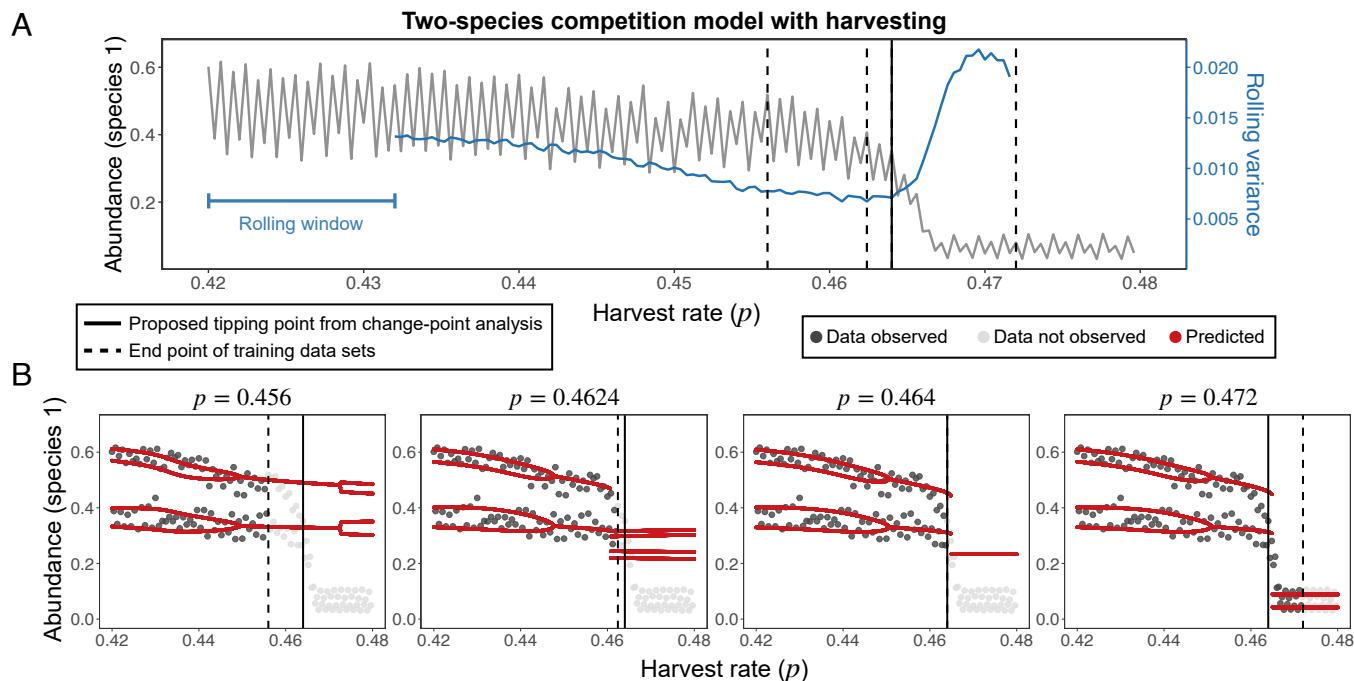


Fig. 3. Illustration of GP-EDM approach as an Early Warning Signal. (A) Population time series (in gray) of the harvested species in a two-species competition model (*SI Appendix*, section S1). Harvest rate (p) is depicted in the x-axis and is proportional to time. The population shows a tipping point at $p = 0.464$ (vertical solid line) according to a change-point analysis. The vertical dashed lines show the end points of the training datasets used with the GP-EDM model. The rolling variance (a typical EWS) computed using a window with 30 points is shown in blue. (B) Reconstructed bifurcation diagrams (in red) using the GP-EDM model trained up to each of the four dashed lines. Bifurcation diagrams depict the population abundance after the transient period (y-axis) for a range of harvest rate values (x-axis). The value of p at the end of each training dataset is shown above each plot. Gray points denote population abundance values used to train the GP-EDM model (dark gray) or not yet observed (light gray).

training data up to October, 1996 (i.e., 9 y after the tipping point; Fig. 5B), reinforcing that the former predictions were accurate. In contrast to these results, the rolling variance of large green algae abundance—the EWS that we used in Fig. 3—did not steadily increase prior to the tipping point (*SI Appendix*, Fig. S20).

Our results were qualitatively the same when removing temperature as a model input, assigning prior information to p , and changing the phosphate moving average smoother (*SI Appendix*, Fig. S21). Last, we mimicked our analysis with the two-species predator-prey model (*SI Appendix*, Fig. S8) and computed the power spectrum and largest Lyapunov exponent of the predicted regimes of large green algae. We found evidence that the dynamics of large green algae transitions from chaos to stable cycles as phosphate concentration decreases during lake oligotrophication (*SI Appendix*, Fig. S22).

Discussion

Foreseeing the dynamical regimes of ecosystems under unobserved environmental conditions is an important ecological challenge, with applications in both conservation and management (2, 19, 36, 37). Although a large body of research has developed EWS that foretell an impending regime shift (1, 8), anticipating what type of dynamics we will observe after the shift has remained elusive. The GP-EDM approach introduced here integrates time-series data and information on a putative driver to enable interpolation and extrapolation to previously unseen dynamical regimes and serve as an EWS. In both simulated and empirical data, we found that GP-EDM can provide useful inference on when and how these systems change to a different regime.

Many EWS have been developed and employed to detect upcoming regime shifts, which include variance, skewness, the

AR(1) coefficient, and the power spectrum (8). Although these indicators cannot give information about unseen regimes, other recent EWS based on deep learning or reservoir computing can perform this task (20, 38, 39). However, these approaches require large amounts of training data, information on the nonlinearities present in the underlying dynamics, or knowledge of the class of bifurcations. In contrast, GP-EDM can be applied with more modest amounts of data and is completely agnostic about the underlying dynamics. Therefore, our approach is able to identify bifurcations that are not among those typically considered in ecological studies. Most importantly, our approach leverages information on the putative driver of regime shifts and, by doing so, allows us to better understand the range of conditions under which an ecosystem exhibits a given type of dynamics.

Although our results are promising, several limitations of GP-EDM are important to note. First, to be used as an EWS, the GP-EDM model must extrapolate to new regions of the control parameter (p) space. Since the prediction accuracy of any Gaussian Process regression declines with long-range extrapolation (40, 41), this limits our ability to anticipate distant regime changes. However, when the inverse-length scale parameter ϕ_i associated with p is small we can extrapolate much farther than when ϕ_i is large (*SI Appendix*, section S3). A small ϕ_i indicates that the map given by function F_i (Eq. 2) is more linear in p . Second, we found that when the training data includes a limited range of abundances (e.g., when it contains mostly fixed points), GP-EDM is likely to determine that the control parameter has little to no effect and the resulting bifurcation diagram is inaccurate (*SI Appendix*, Fig. S9). Therefore, our approach is more accurate when the training data spans a large portion of the state space, either because of process noise or large deterministic fluctuations (e.g., chaos). Note that this is

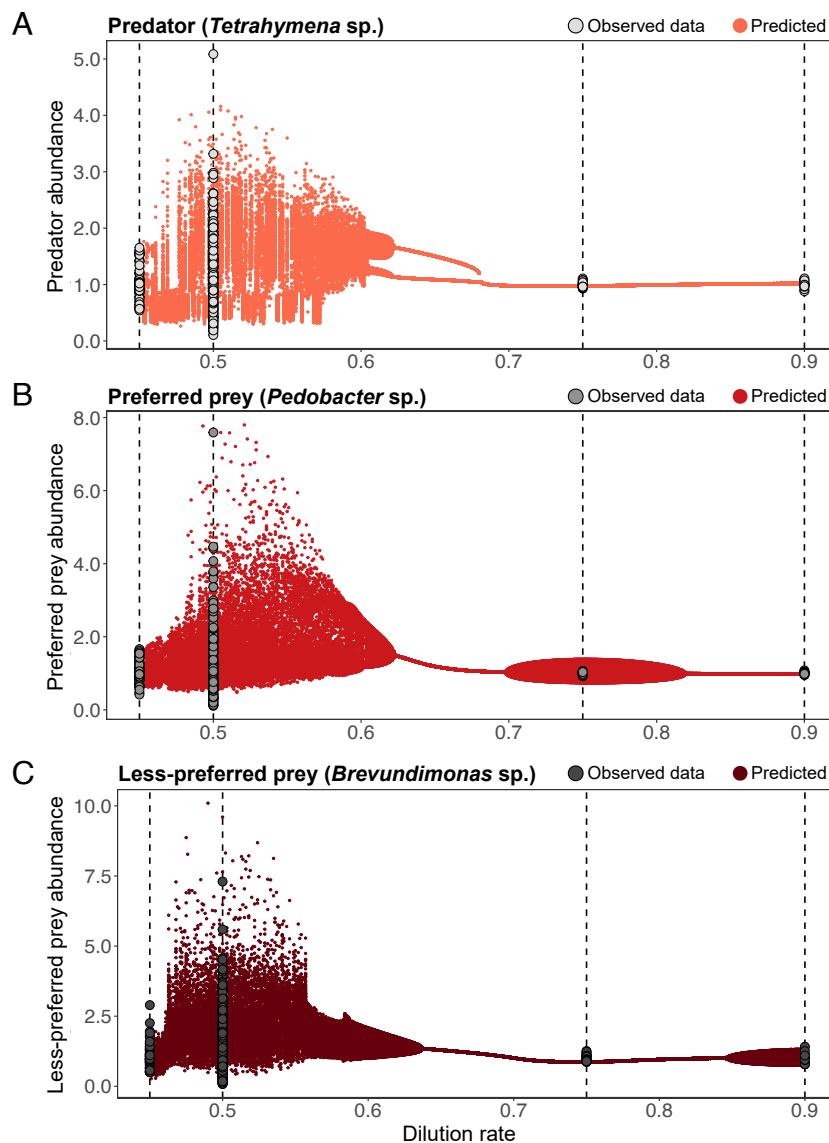


Fig. 4. Predicting unseen dynamical regimes in an experimental microbial food web. Bifurcation diagrams reconstructed with our GP-EDM approach (red points) for the predator species [*Tetrahymena* sp., light red, (A)], preferred prey species [*Pedobacter* sp., moderate red, (B)], and less-preferred prey species [*Brevundimonas* sp., dark red, (C)]. Each bifurcation diagram depicts the abundance of a given species after the transient period (y-axis) for a range of dilution rate values (x-axis). Vertical dashed lines represent the dilution rates of the experimental treatments (0.45/day, 0.5/day, 0.75/day, and 0.9/day). Gray points show the abundance values at these dilution rates used to train a GP-EDM model separately for each species. Abundances are in arbitrary units due to our data scaling process (*Materials and Methods*).

a typical requirement of statistical methods that aim to infer dynamical properties of a system from time-series data (38, 42). In situations where we know that the control parameter is important but GP-EDM determines that it has no effect, we suggest fixing ϕ_i associated with p to 1 (*SI Appendix, section S3*). A related point concerns the required time-series length. Based on previous EDM studies, we recommend using time series that covers several multiples of the organism's generation time for each level of the control parameter (22, 34). Third, ecosystems are typically under the effect of multiple environmental drivers. Our study, like most work in EWS, was restricted to a single control parameter that leads to codimension-one bifurcations. Future studies could include additional control parameters in the GP-EDM model to evaluate whether it is possible to anticipate more complicated regime shifts. Finally, in contrast to other EWS, GP-EDM requires data on the putative control parameter,

or a suitable proxy. In the absence of a known driver, slow feature analysis (43) may be useful to construct a suitable proxy and future studies could explore this idea.

In addition to testing our approach with model-generated data (Figs. 2 and 3), we applied our approach to two empirical datasets: an experimental microbial ecosystem [Fig. 4; (35)] and a lake planktonic ecosystem [Fig. 5; (16)]. Both applications revealed alternative dynamical regimes and shifts between regimes in these ecosystems. Regarding the experimental microbial ecosystem, previous work had determined the existence of three qualitative distinct regimes [limit cycle at low dilution rate, chaos at intermediate dilution rate, and fixed point at high dilution rate; (35)]. However, no experiment was conducted to verify the exact location of bifurcations. Our analyses shed light on this question, especially regarding the shift from the chaotic to the fixed point regime, which should occur at a dilution rate of 0.62 to 0.68/day

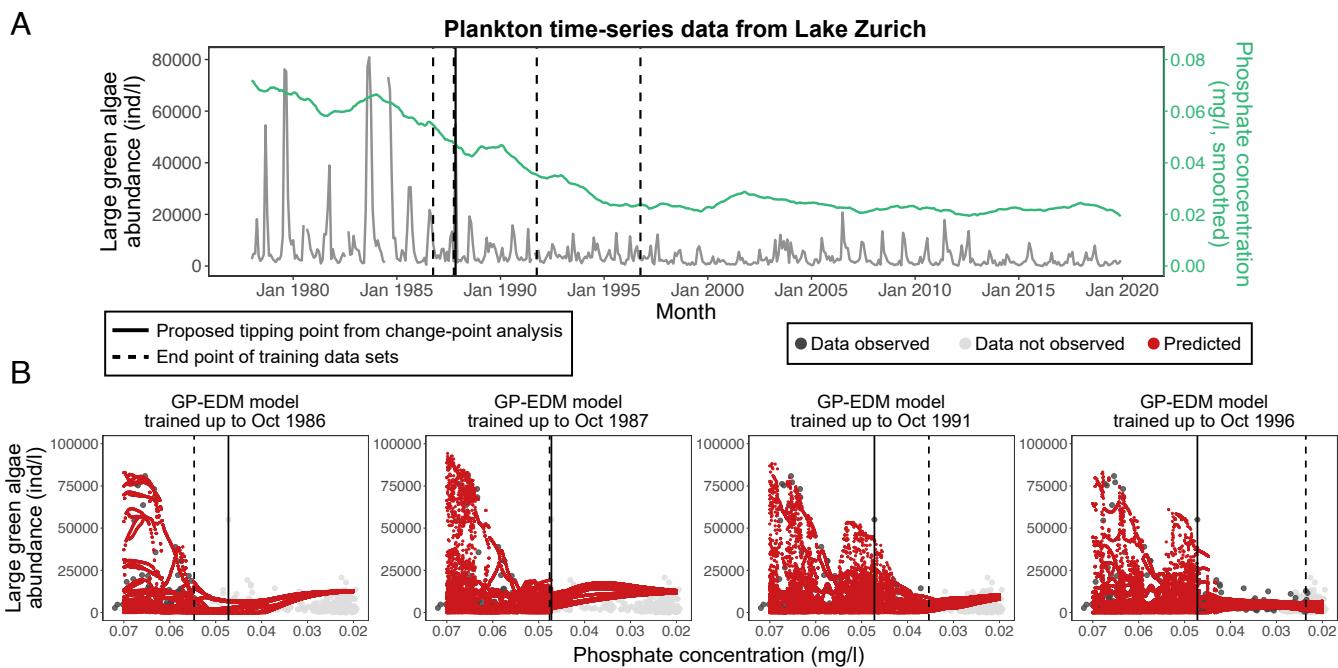


Fig. 5. Anticipating the oligotrophic regime in a lake planktonic food web. (A) Monthly population time series (in gray) of large green algae in Lake Zurich from January, 1978 to December, 2019. The population shows a tipping point on November, 1987 (vertical solid line) according to a change-point analysis. The vertical dashed lines show the end points of the training datasets used with the GP-EDM model. Phosphate concentration (smoothed via a centered moving average), a key environmental driver for oligotrophication, is shown in green. (B) Reconstructed bifurcation diagrams (in red) using the GP-EDM model trained up to each of the four dashed lines. Bifurcation diagrams depict the population abundance after the transient period (y-axis) for a range of phosphate concentration values (x-axis). The training data end points are shown as dates above each plot. Gray points denote population abundance values used to train the GP-EDM model (dark gray) or not yet observed (light gray).

(Fig. 4). Future work could test our prediction by performing experiments under these dilution rates.

Regarding the lake planktonic ecosystem, a large body of research has explored the transitions from oligotrophic to eutrophic and vice versa (4, 16, 44, 45). When analyzing a complex ecosystem such as a lake food web with multiple functional groups, it is challenging to determine whether a shift has occurred, because only certain groups may show a clear pattern of a tipping point (19, 46–48). A key component of GP-EDM is that it uses time-delay embedding to account for hidden state variables. By focusing on large green algae—a highly predictable functional group that showed clear evidence of a tipping point—we were able to anticipate a possible ecosystem-wide regime shift. That is, one month before the potential tipping point, our approach clearly indicated an abrupt shift from a chaotic regime with high-amplitude abundance fluctuations to a cyclic regime with low-amplitude abundance fluctuations for large green algae. Note that this type of regime shift is common in phytoplankton under nutrient reductions but is hard to anticipate with traditional EWS such as rolling variance (*SI Appendix, Fig. S20*) (46, 48). We found that the shift from eutrophic to oligotrophic occurred at a phosphate concentration of 0.0470 mg/l. An important area for future research is to apply GP-EDM to a range of natural ecosystems for which there are also data on a known environmental driver that induces regime shifts to obtain cross-system insights. This type of study could, for example, elucidate whether there is a typical phosphate concentration for the eutrophic-oligotrophic transition in lakes.

In summary, we show that GP-EDM accurately reveals previously unseen dynamical regimes under changing environmental conditions across a range of complexity. Although regime shifts

can be extremely costly, they may also be beneficial or require no intervention [e.g., eutrophic-oligotrophic transition in lakes (16) and fish recruitment success due to hydroclimate regime shift (49)]. Hence, in order to engage in appropriate mitigation actions, it is important to have a plausible estimate of the cost of inaction, which in turn requires us to predict the state of the ecosystem following an impending transition. In addition to ecosystems, many systems in physics, climate, and finance exhibit tipping points (1) and such cost–benefit analyses are useful in these cases as well. The GP-EDM approach developed here applies to any dynamical system modeled as Eq. 1 and future work could use this approach to anticipate unseen regimes in systems from other disciplines. By making accurate predictions of unobserved dynamical regimes, our approach opens the door to rational decision-making and cost–benefit analyses of regime shifts.

Materials and Methods

Dynamical Regimes and Bifurcations. Our goal in this study was to predict population dynamics that occurs under unobserved levels of an environmental driver. By doing so, we also provided predictions for the driver levels under which we should observe tipping points. To formalize these ideas, consider a generic discrete-time model that describes the population dynamics of n species in an ecosystem:

$$\mathbf{x}(t+1) = \mathbf{F}[\mathbf{x}(t), p], \quad [1]$$

where $\mathbf{x}(t) = [x_1(t), \dots, x_n(t)]^\top$ is the vector of abundances at time t , $\mathbf{F} = [F_1, \dots, F_n]^\top$ is a vector-valued function, and p is a *control parameter*. Note that \mathbf{F} typically depends on other parameters, which we consider to be fixed and omit here for simplicity. Systems modeled as Eq. 1 can exhibit different types of

attractor, including fixed points, cycles, and chaos. Each of these types of attractor may have distinct properties such as high or low abundance for fixed points and long or short periods for cycles. We define a *dynamical regime* as an attractor with a given qualitative property. A *bifurcation* occurs when a small change in the control parameter p in Eq. 1 causes a change in its dynamical regime—formally, when this parameter change results in a topologically different phase portrait (14). Specifically, a bifurcation occurs at a critical value (p_c) if we observe distinct dynamical regimes for $p_c - \delta$ and $p_c + \delta$, where δ is a small displacement. Note that we use the terms bifurcation and tipping point interchangeably throughout the text. The number of control parameters that must change to cause a bifurcation is known as the *codimension* of a bifurcation. When working with empirical data, we can leverage knowledge on an environmental driver (e.g., temperature, nutrients, rainfall, fishing pressure) that is known to cause regime shifts and assume that this driver is a proxy for p .

Here, we address the following challenge: Having observed time-series data $\{x_i(t)\}$ ($t = 1, \dots, T$) for a given species i at different dynamical regimes (i.e., at different values of p), can we predict regimes at unobserved values of p (Fig. 1)? We introduce an approach to solve this challenge by focusing on bifurcations of codimension one. In the next section, we describe our approach based on EDM to predict unseen dynamical regimes from time-series data.

Approach to Predict Unseen Dynamical Regimes. To address the problem described above, we combined Gaussian Process regression with time-delay embedding, an approach known as GP-EDM (31). Having observed time series $\{x_i(t)\}$ ($t = 1, \dots, T$) for a given species i at different values of p (Fig. 1A), we would like to learn the map F_i to be able to predict the dynamics of x_i at unobserved values of p :

$$x_i(t+1) = F_i[x_1(t), \dots, x_n(t), p]. \quad [2]$$

However, if we only have data on species i , we propose that we can use time-delay embedding to compensate for unobserved species and instead learn the delay-embedding map G_i (Fig. 1B):

$$x_i(t+1) = G_i[x_i(t), \dots, x_i(t-E), p], \quad [3]$$

where $E + 1$ is the number of lagged versions of x_i , also known as the *embedding dimension*. We can then apply an appropriate nonparametric function approximation method, such as Gaussian Process regression, to learn G_i and interpolate or extrapolate this function to unseen values of p (Fig. 1C and D). We provide a detailed description of GP-EDM in *SI Appendix, section S3*, and we use the R package “GPEDM” to perform our analyses (50).

The rationale for this approach is as follows. Takens’ theorem proves a one-to-one correspondence between an attractor in native coordinates (i.e., $x_1(t), \dots, x_n(t)$) and one that has been reconstructed using time-delay embedding (i.e., $x_i(t), \dots, x_i(t-E)$) (51). The bundle embedding theorem (52) extends time-delay embedding to forced dynamical systems and justifies the inclusion of the control parameter p in the reconstruction. Hence, a bifurcation in native coordinates corresponds to a bifurcation in delay coordinates at the same value of p . Therefore, we hypothesize that learning G_i and using it to interpolate between and extrapolate beyond observed values of p will allow us to reconstruct the bifurcation diagram and to predict the behavior of an ecosystem following a regime shift. Although a bifurcation corresponds to an abrupt change of the attractor, we expect the function G_i to change smoothly with p (Fig. 1B). Note that if we had data on all species in an ecosystem, then we could use Gaussian Process regression (without time-delay embedding) to learn F_i directly (instead of G_i). Also note that, although we use a simple description of our system as having only species abundances as state variables, Takens’ theorem also works when the system contains other unobserved state variables (e.g., life stages, traits). Finally, since Takens’ theorem also applies to continuous-time dynamics (51), our approach does as well. That is, we can use a discretely sampled series $\{x_i(t)\}$ ($t = 1, \dots, T$) to approximate the map that results from integrating the continuous-time dynamics with GP-EDM (34).

Testing the Approach on Model-Generated Data. We tested our approach on four discrete-time and one continuous-time population models exhibiting different types of bifurcations. Discrete-time models included a single-species

model, a two-species competition model with harvesting, a two-species predator-prey model, and a three-species competition model (*SI Appendix, section S1*). Analyses for the continuous-time model are slightly different and are described in *SI Appendix, section S2*. For each discrete-time model, we first determined a range of parameter values, $[p_{\min}, p_{\max}]$, that produced different dynamical regimes. Then, we selected four distinct values of p (p_1, p_2, p_3 , and p_4) within $[p_{\min}, p_{\max}]$ and simulated the dynamics with process noise to generate four time series of length $T = 50$. Note that we iterated each model for 500 time steps and discarded the first 450 points to remove transients. We then stacked these time series together along with additional inputs (i.e., lags of x_i and values of p) to create a data matrix to train the GP-EDM model (*SI Appendix, Fig. S1*). We standardized all inputs to zero mean and unit SD prior to fitting. We selected the value of E for the GP-EDM model based on leave-one-out prediction accuracy, measured as

$$R_E^2 = 1 - \frac{\sum_{t=1}^T [x_i(t) - \hat{x}_i(t)]^2}{\sum_{t=1}^T [x_i(t) - \bar{x}_i]^2}, \quad [4]$$

where $\hat{x}_i(t)$ is the GP-EDM prediction for the left-out observation $x_i(t)$ using a given value of E and \bar{x}_i is the mean of $x_i(t)$ over time. We selected the E value that gave the highest prediction accuracy (i.e., $\max [R_E^2]$ for $E \in \{0, \dots, 8\}$) across all values of p .

After we trained the GP-EDM model, we performed predictions for a grid of 200 values of p within $[p_{\min}, p_{\max}]$ (i.e., interpolated the learned function G_i). To do so for a given value p_j ($j = 1, \dots, 200$), we set initial values of x_i and iterated the learned function G_i forward in time while keeping $p = p_j$ (*SI Appendix, Fig. S1*). Specifically, the prediction $\hat{x}_i(t+1)$ was given by the posterior mean of the GP-EDM model. At every iteration, we used the predicted value as input for the next prediction step. We performed this iteration for 500 time steps and used the last 100 points to characterize the dynamical regime.

We used the JSD to measure the accuracy of our reconstructed bifurcation diagram. The JSD is frequently used to measure the similarity between two probability distributions (e.g., ref. 53). To compute the JSD in a way that is robust to the chosen number of bins, we averaged this metric over several bin numbers (10 to 15 bins). Other procedures such as the Wasserstein distance or the Kolmogorov-Smirnov test could be used here and future work can explore them.

For each value of p_j , we had a sample of two distributions of x_i , one given by the true time series $\{x_i(t)\}$ ($t = 1, \dots, T$) and another given by the predicted time series $\{\hat{x}_i(t)\}$ ($t = 1, \dots, T$), where $T = 100$. Defining P as the discretized (i.e., binned) probability distribution for $\{x_i(t)\}$ and Q as the discretized probability distribution for $\{\hat{x}_i(t)\}$, the JSD is measured as

$$JSD(P||Q) = \frac{1}{2}D(P||M) + \frac{1}{2}D(Q||M), \quad [5]$$

where $M = \frac{1}{2}(P + Q)$ is the mixture distribution of P and Q , and $D(P||M)$ is the Kullback-Leibler divergence of P from M given by

$$D(P||M) = \sum_{x \in \mathcal{X}} P(x) \log_2 \left(\frac{P(x)}{M(x)} \right), \quad [6]$$

where \mathcal{X} is the set of all possible bins and $P(x)$ denotes the probability (i.e., frequency of points) for bin x . Thus, the JSD measures the similarity between P and Q , where 0 denotes very similar distributions and 1 denotes very divergent distributions. Note that the true time series in this case is generated from the model without process noise as opposed to the noisy time series of length $T = 50$ that were used to train the GP-EDM model. That is, we reconstructed the bifurcation diagram from noisy data but compared our reconstruction to the true noise-free diagram.

In addition to the analysis described above, we tested whether our approach can be used as an EWS. To do so, we iteratively applied our approach to a time series generated with the two-species competition model under a slowly changing harvest rate (i.e., the control parameter p). Specifically, we iterated this model while we increased p linearly from 0.42 to 0.48 over 150 time steps. With our parameterization and without process noise, this model shows a fold bifurcation (i.e., a tipping point where the population collapses) at $p = 0.459$.

Before applying our GP-EDM approach, we performed a change-point analysis (54) to determine the location of the tipping point. We did this because, although the deterministic tipping point occurs at $p = 0.459$, the actual tipping point in the data can occur earlier or later due to process noise. To this end, we split the time series of species i into two windows: $\{x_i(1), \dots, x_i(t_{\text{split}})\}$ and $\{x_i(t_{\text{split}} + 1), \dots, x_i(T)\}$ for 100 values of t_{split} within [25, 125]. Then, for each t_{split} , we computed the sum of squares of window j (SS_j) using the within-window average, the sum of squares of the entire time series (SS_{all}) using the global average, and the following ratio $SS_{\text{ratio}} = \frac{SS_1 + SS_2}{SS_{\text{all}}}$. We then selected the t_{split} associated with the lowest value of SS_{ratio} as the tipping point. The lower the value of SS_{ratio} , the greater the difference in mean abundance across the two windows.

To apply our GP-EDM approach, we defined 4 time-series windows: 1) $t = 1, \dots, 90$, 2) $t = 1, \dots, 106$, 3) $t = 1, \dots, 110$, and 4) $t = 1, \dots, 130$. For each window, we trained our GP-EDM model using information on p up to the end of the window and then performed predictions for a grid of 300 values of p within [0.42, 0.48], as we did in our previous analysis. Note that here, we not only interpolate but also extrapolate G_i beyond observed levels of p . This analysis allowed us to verify how well we can anticipate, not only an upcoming tipping point but also the dynamical regime that lies beyond this point. In addition to our GP-EDM prediction of the tipping point location, we computed a well-studied EWS and compared it to our GP-EDM prediction. We used the variance of $x_i(t)$ for successive rolling windows of length 30 as this EWS, which is expected to increase before a fold bifurcation.

Applying the Approach to an Experimental Microbial Ecosystem. Our first analysis using empirical data consisted of reconstructing bifurcation diagrams of an experimental microbial food web containing one ciliate predator (*Tetrahymena* sp.), a preferred bacterial prey (*Pedobacter* sp.), and a less-preferred bacterial prey (*Brevundimonas* sp.) cultivated in a chemostat (35). This dataset contains time series for all three species cultivated under five distinct dilution rate (i.e., mortality) treatments. Dilution rates of 0.9/day and 0.75/day resulted in stable fixed points, 0.5/day generated chaos, and 0.45/day resulted in a limit cycle (35). We used all experimental replicates, which included one times series for the 0.9/day treatment ($T = 23$), one time series for the 0.75/day treatment ($T = 28$), five time series for the 0.5/day treatment ($T = 30$, $T = 42$, $T = 42$, $T = 47$, and $T = 52$), and two time series for the 0.45/day treatment ($T = 32$ and $T = 45$). Because the only environmental condition changing across treatments is the dilution rate, this dataset represents a family of dynamical systems tuned by a single control parameter, precisely analogous to our simulations shown in Fig. 2.

We applied the same analysis that we performed for model-generated data independently to each species in this system. For a given species, we first concatenated all replicate time series across the four dilution treatments (i.e., $p = 0.45, 0.5, 0.75$, and 0.9) and included dilution rate (p) as an additional input. We removed the 5 initial days of transient dynamics from each replicate to ensure that each time series covered a single dynamical regime. We log-transformed abundance as this greatly improved leave-one-out prediction accuracy. Because dilution rate treatments can result in very different mean abundances, we subtracted the mean treatment abundance from each time series. Then, we standardized all inputs and trained the GP-EDM model. We tested $E \in \{0, \dots, 8\}$ and selected the value that maximized leave-one-out prediction accuracy. Note that, because the original number of state variables in this system (e.g., abundances, traits) may be high, large embedding dimensions may be needed to reconstruct the system's attractor via Takens' theorem (51). We then interpolated the GP-EDM model to predict dynamical regimes across a grid of 450 values of p within [0.45, 0.9].

Applying the Approach to a Lake Planktonic Ecosystem. Our second analysis using empirical data consisted of an EWS analysis to anticipate a potential tipping point from an eutrophic to an oligotrophic regime in a lake planktonic food web. We used a dataset of lake planktonic food webs compiled by Merz et al. (16) and selected Lake Zurich to illustrate our approach, as it contained the longest time series and showed signs of a regime shift driven by a gradual decrease in phosphorus. The planktonic food web from Lake Zurich

contained monthly time series for 13 functional groups and 2 environmental variables (phosphate concentration and water temperature) from January, 1978 to December, 2019 ($T = 504$). Because our EWS analysis with model-generated data was based on a linearly decreasing control parameter, we applied a centered moving average smoother (55) with a window size of 24 mo to remove seasonality and retain the trend in phosphate concentration. We included seasonality in our GP-EDM model through water temperature (see below). Before applying our GP-EDM approach, we performed a change-point analysis (54) to determine the location of a potential tipping point for each functional group separately (see *Testing approach on model-generated data*). For each functional group, we split its time series into two windows for 300 values of t_{split} within [102, 402] and computed SS_{ratio} for each t_{split} . We selected the t_{split} associated with the lowest value of SS_{ratio} as a potential tipping point.

The above analyses showed that large green algae had the lowest value of SS_{ratio} of all functional groups at $t_{\text{split}} = 119$ (November, 1987). Thus, we selected the time-series window from $t = 1$ (January, 1978) to $t = 106$ (October, 1986) to train our GP-EDM model and test whether we could predict the unseen oligotrophic regime. We first performed leave-one-out predictions for each functional group separately using the trained GP-EDM model to verify which group was more predictable. We used up to 7 lags of functional group abundance (i.e., $x_i(t), \dots, x_i(t-6)$), up to 2 lags of water temperature (i.e., $w(t), w(t-1)$) to account for seasonality, and 1 lag of phosphate concentration ($p(t)$) (i.e., control parameter) as inputs in our GP-EDM model. We standardized all inputs before training the GP-EDM model. In addition to showing the lowest value of SS_{ratio} , large green algae showed the second-best leave-one-out prediction accuracy ($R^2 = 0.770$). Therefore, we focused on this functional group for subsequent analyses. The best GP-EDM model for large green algae was given by $x_i(t+1) = G_i[x_i(t), \dots, x_i(t-3), w(t), w(t-1), p(t)]$.

We then followed our EWS analysis with model-generated data to make predictions for a range of phosphate concentrations using the time series for large green algae. To do so, we defined 4 time-series windows: 1) one year before the potential tipping point ($t = 1, \dots, 106$), 2) one month before the potential tipping point ($t = 1, \dots, 118$), 3) four years after the potential tipping point ($t = 1, \dots, 166$), and 4) nine years after the potential tipping point ($t = 1, \dots, 226$). For each window, we trained the GP-EDM model defined above and then performed predictions for a grid of 300 values of p within [0.02, 0.07], which correspond approximately to the minimum and maximum values of phosphate concentration in the data. When iterating predictions forward for a given value of p , we fixed p to that value and set w to the average of the corresponding month. By reconstructing the bifurcation diagram for large green algae as more data becomes available, we provide a real-world illustration of our simulations shown in Fig. 3.

Data, Materials, and Software Availability. Data and code have been deposited in GitHub (<https://github.com/lucaspdmedeiros/gpedm-regime-shifts>). Previously published data were used for this work (16, 35).

ACKNOWLEDGMENTS. We thank Vadim Karataev, Tanya Rogers, Vasilis Dakos, and Alan Hastings for useful discussions and helpful comments about this work. We also thank the reviewers for comments that greatly improved this work. This work was supported by the Sustainable Oceans National Science Foundation Research Traineeship (grant number #1734999), the Lenfest Oceans Program, National Oceanic and Atmospheric Administration High Performance Computing and Communications program, NOAA's Quantitative Ecology and Socioeconomics Training program, and the Cooperative Institute for Marine, Earth, and Atmospheric Systems.

Author affiliations: ^aFisheries Collaborative Program, Institute of Marine Sciences, University of California, Santa Cruz, CA 95060; ^bDepartment of Biology, Woods Hole Oceanographic Institution, Woods Hole, MA 02543; ^cDepartment of Applied Mathematics, University of California, Davis, CA 95616; ^dDepartment of Mathematics, University of California, Santa Cruz, CA 95064; ^eDepartment of Mathematics, California State Polytechnic University Humboldt, Arcata, CA 95521; ^fDepartment of Ecology and Evolutionary Biology, University of California, Santa Cruz, CA 95060; and ^gSouthwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Santa Cruz, CA 95060

1. M. Scheffer *et al.*, Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
2. M. Scheffer, S. Carpenter, J. A. Foley, C. Folke, B. Walker, Catastrophic shifts in ecosystems. *Nature* **413**, 591–596 (2001).
3. V. Dakos, S. R. Carpenter, E. H. van Nes, M. Scheffer, Resilience indicators: Prospects and limitations for early warnings of regime shifts. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **370**, 20130263 (2015).
4. S. R. Carpenter, D. Ludwig, W. A. Brock, Management of eutrophication for lakes subject to potentially irreversible change. *Ecol. Appl.* **9**, 751–771 (1999).
5. K. Filbee-Dexter, R. E. Scheibling, Sea urchin barrens as alternative stable states of collapsed kelp ecosystems. *Mar. Ecol. Prog. Ser.* **495**, 1–25 (2014).
6. T. P. Hughes, Catastrophes, phase shifts, and large-scale degradation of a caribbean coral reef. *Science* **265**, 1547–1551 (1994).
7. M. Hirota, M. Holmgren, E. H. Van Nes, M. Scheffer, Global resilience of tropical forest and savanna to critical transitions. *Science* **334**, 232–235 (2011).
8. V. Dakos *et al.*, Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PLoS One* **7**, e41010 (2012).
9. S. H. Strogatz, *Nonlinear Dynamics and Chaos: With Applications to Physics, Biology, Chemistry, and Engineering* (CRC Press, 2018).
10. G. M. Wilkinson *et al.*, Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments. *Ecol. Monogr.* **88**, 188–203 (2018).
11. A. J. Veraart *et al.*, Recovery rates reflect distance to a tipping point in a living system. *Nature* **481**, 357–359 (2012).
12. J. M. Drake, B. D. Griffen, Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459 (2010).
13. L. Dai, D. Vorselen, K. S. Korolev, J. Gore, Generic indicators for loss of resilience before a tipping point leading to population collapse. *Science* **336**, 1175–1177 (2012).
14. Y. A. Kuznetsov, *Elements of Applied Bifurcation Theory* (Springer, 1998), vol. 112.
15. C. Boettiger, N. Ross, A. Hastings, Early warning signals: The charted and uncharted territories. *Thyroid Res.* **6**, 255–264 (2013).
16. E. Merz *et al.*, Disruption of ecological networks in lakes by climate change and nutrient fluctuations. *Nat. Clim. Change* **13**, 389–396 (2023).
17. H. Fujita *et al.*, Alternative stable states, nonlinear behavior, and predictability of microbiome dynamics. *Microbiome* **11**, 63 (2023).
18. E. Weinans *et al.*, Finding the direction of lowest resilience in multivariate complex systems. *J. R. Soc. Interface* **16**, 20190629 (2019).
19. J. J. Lever *et al.*, Foreseeing the future of mutualistic communities beyond collapse. *Ecol. Lett.* **23**, 2–15 (2020).
20. T. M. Bury *et al.*, Deep learning for early warning signals of tipping points. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2106140118 (2021).
21. M. W. Adamson, J. H. Dawes, A. Hastings, F. M. Hilker, Forecasting resilience profiles of the run-up to regime shifts in nearly-one-dimensional systems. *J. R. Soc. Interface* **17**, 20200566 (2020).
22. S. B. Munch, A. Giron-Nava, G. Sugihara, Nonlinear dynamics and noise in fisheries recruitment: A global meta-analysis. *Fish. Fish.* **19**, 964–973 (2018).
23. N. Nova *et al.*, Susceptible host availability modulates climate effects on dengue dynamics. *Ecol. Lett.* **24**, 415–425 (2021).
24. G. Sugihara *et al.*, Detecting causality in complex ecosystems. *Science* **338**, 496–500 (2012).
25. L. P. Medeiros, S. Allesina, V. Dakos, G. Sugihara, S. Saavedra, Ranking species based on sensitivity to perturbations under non-equilibrium community dynamics. *Ecol. Lett.* **26**, 170–183 (2023).
26. C. Hsieh, S. M. Glaser, A. J. Lucas, G. Sugihara, Distinguishing random environmental fluctuations from ecological catastrophes for the north pacific ocean. *Nature* **435**, 336–340 (2005).
27. V. Dakos, S. M. Glaser, C. Hsieh, G. Sugihara, Elevated nonlinearity as an indicator of shifts in the dynamics of populations under stress. *J. R. Soc. Interface* **14**, 20160845 (2017).
28. F. Grziwotz *et al.*, Anticipating the occurrence and type of critical transitions. *Sci. Adv.* **9**, eabq4558 (2023).
29. Y. J. Huang, C. W. Chang, Ch. Hsieh, Detecting shifts in nonlinear dynamics using empirical dynamic modeling with nested-library analysis. *PLoS Comput. Biol.* **20**, e1011759 (2024).
30. T. Säterberg, K. McCann, Detecting alternative attractors in ecosystem dynamics. *Commun. Biol.* **4**, 975 (2021).
31. S. B. Munch, V. Poynor, J. L. Arriaza, Circumventing structural uncertainty: A Bayesian perspective on nonlinear forecasting for ecology. *Ecol. Complex.* **32**, 134–143 (2017).
32. G. Sugihara, Nonlinear forecasting for the classification of natural time series. *Philos. Trans. R. Soc. Lond. Ser. Lond. Math. Phys. Eng. Sci.* **348**, 0106 (1994).
33. G. Sugihara, R. M. May, Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **344**, 7347741 (1990).
34. S. B. Munch, A. Brias, G. Sugihara, T. L. Rogers, Frequently asked questions about nonlinear dynamics and empirical dynamic modelling. *ICES J. Mar. Sci.* **77**, 1463–1479 (2020).
35. L. Becks, F. M. Hilker, H. Malchow, K. Jürgens, H. Arndt, Experimental demonstration of chaos in a microbial food web. *Nature* **435**, 1226–1229 (2005).
36. A. Coreau, G. Pinay, J. D. Thompson, P. O. Cheptou, L. Mermet, The rise of research on futures in ecology: Rebalancing scenarios and predictions. *Ecol. Lett.* **12**, 1277–1286 (2009).
37. O. L. Petchey *et al.*, The ecological forecast horizon, and examples of its uses and determinants. *Ecol. Lett.* **18**, 597–611 (2015).
38. D. Patel, E. Ott, Using machine learning to anticipate tipping points and extrapolate to post-tipping dynamics of non-stationary dynamical systems. *Chaos: An Interdisciplinary J. Nonlinear Sci.* **33**, 023143 (2023).
39. L. W. Kong, H. W. Fan, C. Grebogi, Y. C. Lai, Machine learning prediction of critical transition and system collapse. *Phys. Rev. Res.* **3**, 013090 (2021).
40. D. Duvenaud, "Automatic model construction with Gaussian processes," PhD thesis, University of Cambridge (2014).
41. A. Wilson, R. Adams, "Gaussian process kernels for pattern discovery and extrapolation" in *International Conference on Machine Learning*, S. Dasgupta, D. McAllester, Eds. (PMLR, 2013), pp. 1067–1075.
42. L. Ljung, "System identification" in *Signal Analysis and Prediction*, A. Procházka, J. Uhliř, P. Rayner, N. Kingsbury, Eds. (Birkhäuser, Boston, MA, 1998), pp. 163–173.
43. L. Wiskott, T. J. Sejnowski, Slow feature analysis: Unsupervised learning of invariances. *Neural Comput.* **14**, 715–770 (2002).
44. R. Wang *et al.*, Flickering gives early warning signals of a critical transition to a eutrophic lake state. *Nature* **492**, 419–422 (2012).
45. M. Scheffer, S. H. Hosper, M. L. Meijer, B. Moss, E. Jeppesen, Alternative equilibria in shallow lakes. *Trends Ecol. Evol.* **8**, 275–279 (1993).
46. A. S. Gsell *et al.*, Evaluating early-warning indicators of critical transitions in natural aquatic ecosystems. *Proc. Natl. Acad. Sci. U.S.A.* **113**, E8089–E8095 (2016).
47. A. Aparicio, J. X. Velasco-Hernández, C. H. Moog, Y. Y. Liu, M. T. Angulo, Structure-based identification of sensor species for anticipating critical transitions. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2104732118 (2021).
48. D. A. O'Brien *et al.*, Early warning signals have limited applicability to empirical lake data. *Nat. Commun.* **14**, 7942 (2023).
49. N. Goikoetxea, X. Irigoien, Links between the recruitment success of northern European hake (*Merluccius merluccius* L.) and a regime shift on the NE Atlantic continental shelf. *Fish. Oceanogr.* **22**, 459–476 (2013).
50. S. B. Munch, T. L. Rogers, GPEDM: Gaussian Process regression for Empirical Dynamic Modeling (R package version 0.0.0.9010, 2025). <https://tanyalrogers.github.io/GPEDM>. Accessed 21 May 2025.
51. F. Takens, "Detecting strange attractors in turbulence" in *Lecture Notes in Mathematics Dynamical Systems and Turbulence, Warwick 1980*, D. Rand, L-S. Young, Eds. (Springer Berlin, Heidelberg, 1981), pp. 366–381.
52. J. Stark, Delay embeddings for forced systems. I. Deterministic forcing. *J. Nonlinear Sci.* **9**, 255–332 (1999).
53. S. V. Scarpino, G. Petri, On the predictability of infectious disease outbreaks. *Nat. Commun.* **10**, 898 (2019).
54. L. J. Gilarranz, A. Narwani, D. Odermatt, R. Siber, V. Dakos, Regime shifts, trends, and variability of lake productivity at a global scale. *Proc. Natl. Acad. Sci. U.S.A.* **119**, e2116413119 (2022).
55. G. Shmueli, J. Polak, *Practical Time Series Forecasting with R: A Hands-On Guide* (Axelrod schnall publishers, 2024).

Supporting Information

Revealing unseen dynamical regimes of ecosystems from population time-series data

Lucas P. Medeiros, Darian K. Sorenson, Bethany J. Johnson, Eric P. Palkovacs, and Stephan B. Munch

S1 Discrete-time population dynamics models

For our main set of analyses, we used four discrete-time population dynamics models to test our GP-EDM approach. The first model consists of the classic single-species logistic model given by (1):

$$x_1(t+1) = x_1(t)p[1 - x_1(t)] \exp [z(t)], \quad [\text{S1}]$$

where p is the intrinsic growth rate of the population and $z(t)$ is process noise sampled independently at each time t from a normal distribution with mean $-\frac{s^2}{2}$ and variance s^2 . This noise term is the same for all models (see below) and guarantees that $\exp [z(t)]$ has a log-normal distribution with mean one. For our main set of simulations, we used $[p_{\min}, p_{\max}] = [2.5, 4.0]$, $p_1 = 2.5$, $p_2 = 2.8$, $p_3 = 3.25$, $p_4 = 4.0$, and $s = 0.02$.

The second model consists of a two-species competition model with harvesting given by:

$$\begin{aligned} x_1(t+1) &= x_1(t) \exp \left[r \left(1 - \frac{x_1(t)}{k} - cx_2(t) \right) - \frac{px_1(t)}{a^2 + x_1^2(t)} \right] \exp [z_1(t)] \\ x_2(t+1) &= x_2(t) \exp \left[r \left(1 - \frac{x_2(t)}{k} - bx_1(t) \right) \right] \exp [z_2(t)], \end{aligned} \quad [\text{S2}]$$

where $r = 2.7$ is the intrinsic growth rate, $k = 1$ is the carrying capacity, $b = 0.1$ and $c = 0.2$ are competition strengths, p is the harvest rate, $a = 0.1$ is the harvest saturation coefficient, and $z_i(t)$ is the process noise for species i . For our main set of simulations, we used $[p_{\min}, p_{\max}] = [0.4, 0.5]$, $p_1 = 0.4$, $p_2 = 0.44$, $p_3 = 0.47$, $p_4 = 0.5$, and $s = 0.02$.

The third model consists of a two-species predator-prey model given by (2):

$$\begin{aligned} x_1(t+1) &= x_1(t) \exp \left[r \left(1 - \frac{x_1(t)}{k} \right) - \frac{px_2(t)}{1 + phx_1(t)} \right] \exp [z_1(t)] \\ x_2(t+1) &= x_2(t) \exp \left[\frac{epx_1(t)}{1 + phx_1(t)} - m \right] \exp [z_2(t)], \end{aligned} \quad [\text{S3}]$$

where $r = 2.5$ is the prey intrinsic growth rate, $k = 4$ is the prey carrying capacity, p is the predator attack rate, $h = 0.1$ is the predator handling time, $e = 0.5$ is the predator conversion rate, $m = 2$ is the predator mortality rate, and $z_i(t)$ is the process noise for species i . For our main set of simulations, we used $[p_{\min}, p_{\max}] = [2.4, 3.4]$, $p_1 = 2.4$,

21 $p_2 = 2.6$, $p_3 = 3.15$, $p_4 = 3.4$, and $s = 0.02$.

22 The fourth model consists of a three-species competition model given by (3):

$$\begin{aligned} x_1(t+1) &= x_1(t) \exp \left[p(1 - x_1(t) - ax_2(t) - bx_3(t)) \right] \exp [z_1(t)] \\ x_2(t+1) &= x_2(t) \exp \left[p(1 - cx_1(t) - x_2(t) - dx_3(t)) \right] \exp [z_2(t)] \\ x_3(t+1) &= x_3(t) \exp \left[p(1 - ex_1(t) - fx_2(t) - x_3(t)) \right] \exp [z_3(t)], \end{aligned} \quad [\text{S4}]$$

23 where p is the intrinsic growth rate and $a = 0.5$, $b = 0.5$, $c = 0.1$, $d = 0.2$, $e = 0.2$,
24 $f = 0.1$ are competition strengths, and $z_i(t)$ is the process noise for species i . For our
25 main set of simulations, we used $[p_{\min}, p_{\max}] = [1.5, 3.0]$, $p_1 = 1.5$, $p_2 = 2.1$, $p_3 = 2.4$,
26 $p_4 = 3.0$, and $s = 0.02$.

27 S2 Continuous-time population dynamics model

28 In addition to the four discrete-time models described above, we also tested our approach using a continuous-time model. Specifically, we used a stochastic version of the
29 Hastings-Powell three-species food chain model (4):

$$\begin{aligned} dx_1 &= x_1 \left[(1 - x_1) - \frac{a_1 x_2}{p + x_1} \right] dt + \sigma dW_1 \\ dx_2 &= x_2 \left[\frac{a_1 x_1}{p + x_1} - d_1 - \frac{a_2 x_3}{b_2 + x_2} \right] dt + \sigma dW_2 \\ dx_3 &= x_3 \left[\frac{a_2 x_3}{b_2 + x_2} - d_2 \right] dt + \sigma dW_3, \end{aligned} \quad [\text{S5}]$$

31 where $a_1 = 5.0$ is the consumer attack rate, $a_2 = 0.1$ is the predator attack rate, p is the
32 half-saturation coefficient of the consumer, $b_2 = 2.0$ is the half-saturation coefficient of
33 the predator, $d_1 = 0.4$ is the consumer natural mortality, $d_2 = 0.01$ is the predator natural
34 mortality, and $\sigma = 0.001$ is the variance of the Wiener process dW_i . The Wiener process
35 has mean zero and is uncorrelated across species. Note that this additive noise term is
36 used for simplicity and, if σ is too large, can generate negative abundances. However, this
37 was not a problem for our simulated time series because of the relatively small σ (Fig.
38 S11). Except for the stochastic term, we use the same parameter values as (4). For our
39 simulations, we used $[p_{\min}, p_{\max}] = [2.0, 3.0]$, $p_1 = 2.0$, $p_2 = 2.3$, $p_3 = 2.6$, and $p_4 = 3.0$.

40 To conduct our analyses with this model, we first integrated equation [S5] for each
41 p_i separately using an integration time step of 0.01 and setting $\sigma = 0$ to generate a
42 noise-free time series with 100,000 points. Then, we removed the first half of the time
43 series to eliminate transients, obtaining a times series with 50,000 points. We computed
44 the local maxima and minima of the resulting time series to build the true bifurcation
45 diagram for this system (see (4) for details). We then repeated this procedure but using
46 $\sigma = 0.001$ to obtain a noisy time series with 50,000 points. In this case, we integrated
47 equation [S5] using the Euler-Maruyama method. Next, we sub-sampled equally spaced

48 points to obtain a final time series $\{x_i(t)\}$ ($t = 1, \dots, T$) with 100 points (i.e., $T = 100$)
 49 for each p_i . We then followed the same procedure described in the main text (*Materials*
 50 and *Methods*) to train a GP-EDM model and predict the dynamics for unseen values of
 51 p . We used the time series for the resource species (i.e., x_1 in equation [S5]) to perform
 52 time-delay embedding and conduct our analyses with GP-EDM. After generating our
 53 predicted time series $\{\hat{x}_i(t)\}$ ($t = 1, \dots, T$) for a given value of p as described in *Materials*
 54 and *Methods*, we computed the local maxima and minima of the predicted time series to
 55 build a bifurcation diagram and compare to the true bifurcation diagram. The results of
 56 this analysis are shown in Fig. S11.

57 **S3 Gaussian Process regression with time-delay em-** 58 **bedding (GP-EDM)**

59 As described in the main text, a key aspect of our approach is to learn the following
 60 function G :

$$x(t+1) = G[x(t), \dots, x(t - E_x), p(t), \dots, p(t - E_p)], \quad [\text{S6}]$$

61 where $x(t)$ is the abundance of a given species at time t , $p(t)$ is the control parameter
 62 at time t , $E_x + 1$ is the number of lagged versions of x , and $E_p + 1$ is the number of
 63 lagged versions of p . Note that we dropped the subscript i to denote the i -th species
 64 in the ecosystem (i.e., x_i) to simplify the notation. Also note that in the main text we
 65 focused mostly on a single lag for p , but the approach can also accommodate multiple
 66 lags of p , which is particularly useful when the environmental driver changes through
 67 time (see *Applying the approach to a lake planktonic ecosystem* in the main text). We
 68 define $\mathbf{x}(t) = [x(t), \dots, x(t - E_x), p(t), \dots, p(t - E_p)]^\top$ as the vector containing all lagged
 69 versions of x and p . If we consider the approximation error $\epsilon(t) \sim \mathcal{N}(0, v)$, we arrive at
 70 the following equation:

$$x(t+1) = G[\mathbf{x}(t)] + \epsilon(t). \quad [\text{S7}]$$

71 Note that for a standard implementation of EDM, $\mathbf{x}(t)$ simply includes lags of the state
 72 variable we would like to predict (i.e., $\mathbf{x}(t) = [x(t), \dots, x(t - E_x)]^\top$). However, in this
 73 study, we also include one or more lags of the control parameter. In what follows, we
 74 describe how we can use Gaussian Process regression to learn G and extrapolate this
 75 function to unseen values of p . A similar approach is also described in previous studies
 76 such as (5, 6).

77 A Gaussian Process (GP) is a continuous generalization of the multivariate normal
 78 distribution. It is completely specified by a mean function, $m(\mathbf{x}) = \mathbb{E}[G(\mathbf{x})]$, and a
 79 covariance function, $C(\mathbf{x}(t), \mathbf{x}(s)) = \mathbb{E}[(G(\mathbf{x}(t)) - m(\mathbf{x}(t)))(G(\mathbf{x}(s)) - m(\mathbf{x}(s)))]$ of the
 80 process $G(\mathbf{x})$, where $\mathbf{x}(t)$ and $\mathbf{x}(s)$ represent two different input vectors. A GP is typically
 81 written as $G(\mathbf{x}) \sim GP(m(\mathbf{x}), C(\mathbf{x}(t), \mathbf{x}(s)))$ (7). Since we do not have information on the
 82 characteristics of $G(\mathbf{x})$ *a priori*, we assumed a constant prior mean function $m(\mathbf{x}) = \mathbf{0}$.
 83 Given that we always standardize our data to mean zero and unit standard deviation,

84 setting the prior mean to zero corresponds to setting it to the average species abundance.
 85 That is, in the absence of data, the GP-EDM model would just predict $x(t+1)$ to be the
 86 average species abundance. Following previous applications of GP-EDM (5, 6), we used
 87 a squared-exponential covariance function, given by:

$$C(\mathbf{x}(t), \mathbf{x}(s)) = \tau^2 \exp \left[-\sum_{j=1}^D \phi_j (\mathbf{x}_j(t) - \mathbf{x}_j(s))^2 \right], \quad [\text{S8}]$$

88 where τ^2 is the prior variance in $G(\mathbf{x})$ and D is the total number of input variables (i.e.,
 89 dimension of the input vectors). In our case, $D = E_x + E_p + 2$. The ϕ_j 's are the inverse-
 90 length scale parameters, which govern how wiggly the function $G(\mathbf{x})$ is in the direction
 91 of the j th input. Importantly, if $\phi_j = 0$, then the j th input has no effect on the output
 92 (i.e., $x(t+1)$).

93 In keeping with previous applications of GP-EDM (5, 6), we used minimally infor-
 94 mative priors for the hyperparameters and selected values for these by maximizing the
 95 marginal posterior distribution after integrating out the unknown function, G . We then
 96 performed predictions for new values of \mathbf{x} using the posterior distribution for G condi-
 97 tioned on the data and these MAP (maximum *a posteriori*) estimates of the hyperpa-
 98 rameters (see (5) for further details and justification).

99 Our fully specified GP-EDM regression model is given by:

$$\begin{aligned} p(x(t+1) | G, \mathbf{x}(t), v) &\sim \mathcal{N}(G(\mathbf{x}(t)), v) \\ p(G | \tau^2, \boldsymbol{\phi}) &\sim GP(\mathbf{0}, C) \\ p(\boldsymbol{\phi}, \tau^2, v), \end{aligned} \quad [\text{S9}]$$

100 where $p(\boldsymbol{\phi}, \tau^2, v)$ is the prior specification for the hyperparameters $\boldsymbol{\phi}$, τ^2 , and v and the
 101 vector $\boldsymbol{\phi} = [\phi_1, \dots, \phi_D]^\top$ contains all of inverse-length scale parameters. We used indepen-
 102 dent, half-normal priors for $\boldsymbol{\phi}$, that is, $p(\phi_j) \sim \frac{2}{\sqrt{2\pi\lambda}} \exp[-\phi_j^2/2\lambda]$ which has a mode at 0
 103 to encourage irrelevant inputs to drop out of the model. This sparsity-inducing prior is
 104 called automatic relevance determination (ARD, (7)) and is analogous to regularization
 105 in linear regression. We set $\lambda = \frac{\pi}{2}$ so that the function has approximately one local
 106 extremum on average over the range of the data (5). We set Beta(1.1, 1.1) priors for
 107 $\tau^2/Var(x)$ and $v/Var(x)$, which only restricts the total variance in the predicted output
 108 to be less than twice the observed variance in the data (5) and facilitates identifiability.

109 To perform leave-one-out predictions for x using the GP-EDM model, we first deter-
 110 mined the hyperparameters that maximize the marginal log-likelihood using the time-
 111 series data containing $x(t)$ and $p(t)$ for $t = 1, \dots, T - E + 1$, where T is the total number
 112 of data points. Defining the vector of hyperparameters as $\boldsymbol{\theta} = [\boldsymbol{\phi}, \tau^2, v]^\top$, the marginal
 113 log-likelihood is given by:

$$\ln p(\boldsymbol{\theta} | \text{data}) = -\frac{1}{2} \ln |\mathbf{C} + v\mathbf{I}| - \frac{1}{2} \mathbf{y}^\top (\mathbf{C} + v\mathbf{I})^{-1} \mathbf{y} - \ln p(\boldsymbol{\theta}), \quad [\text{S10}]$$

114 where \mathbf{y} is the vector of observed outputs containing $x(t+1)$ for $t = 1, \dots, T$ and \mathbf{I} is
 115 the $T \times T$ identity matrix. This maximization was done using the R-prop algorithm (8).
 116 Given the estimated hyperparameters, we performed predictions for a set of new inputs
 117 \mathbf{x}_{new} by computing the conditional mean (m_c) and covariance (C_c) functions evaluated
 118 using these new inputs. This was done via the following updating rule:

$$\begin{aligned}
 m_c(\mathbf{x}_{\text{new}}) &= C(\mathbf{x}_{\text{new}}, \mathbf{X}^\top)(\mathbf{C} + v\mathbf{I})^{-1}\mathbf{y} \\
 C_c(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{new}}^\top) &= C(\mathbf{x}_{\text{new}}, \mathbf{x}_{\text{new}}^\top) - C(\mathbf{x}_{\text{new}}, \mathbf{X}^\top)(\mathbf{C} + v\mathbf{I})^{-1}C(\mathbf{X}, \mathbf{x}_{\text{new}}^\top),
 \end{aligned} \quad [\text{S11}]$$

119 where \mathbf{C} is the $T \times T$ covariance matrix obtained by evaluating the covariance function
 120 at all pairs of input points.

121 To perform leave-one out cross-validation, we modified the conditional mean as fol-
 122 lows. Let y_j be the point we are trying to predict and the \mathbf{x}_j be the corresponding input.
 123 Let \mathbf{y}_{-j} and \mathbf{X}_{-j} be the arrays of observed targets and inputs with the j^{th} row removed
 124 and let \mathbf{C}_{-j} be the covariance matrix with the corresponding rows and columns deleted.
 125 The prediction for y_j is given by:

$$\hat{y}_j = C(\mathbf{x}_j, \mathbf{X}_{-j}^\top)(\mathbf{C}_{-j} + v\mathbf{I}_{-j})^{-1}\mathbf{y}_{-j}. \quad [\text{S12}]$$

126 We used the R package “GPEDM” (9) to train our GP-EDM models and to perform
 127 predictions.

128 We used equation [S11] to predict the dynamical regime of x for unseen values of the
 129 control parameter p . For simplicity, we explain this procedure when using a single lag
 130 of p . To predict the dynamical regime for a given unseen value \tilde{p} , we first estimate the
 131 hyperparameters using the entire data set for all observed regimes (i.e., all time series of
 132 x for all observed values of p). Then, we set $\mathbf{x}_{\text{new}} = [x(t^*), \dots, x(t^* - E_x), \tilde{p}]^\top$ using the last
 133 $E_x + 1$ values of the time series. If we had time series of x for different values of p (e.g., as
 134 in Figs. 1 and 2 in the main text), we chose the time series for which the observed p was
 135 closest to \tilde{p} . Then, we used $m_c(\mathbf{x}_{\text{new}})$ as our prediction for $x(t^* + 1)$ using equation [S11].
 136 Next, we performed a prediction for $x(t^* + 2)$ using $\mathbf{x}_{\text{new}} = [x(t^* + 1), \dots, x(t^* + 1 - E_x), \tilde{p}]^\top$,
 137 where $x(t^* + 1)$ is our previous prediction. We repeated this procedure for 500 time steps
 138 for our analyses with models and 1000 time steps for our analyses with empirical data.
 139 We then used the 100 last predicted abundances as our predicted dynamical regime as
 140 explained in the main text.

141 A final point to note is that, when the inverse-length scale parameter (ϕ_j) for the
 142 control parameter is close to zero, the fitted GP-EDM model is insensitive to changes
 143 in p . This occurred in our initial analyses of the experimental microbial ecosystem (see
 144 *Applying the approach to an experimental microbial ecosystem* in the main text). As
 145 a consequence, the reconstructed bifurcation diagram for the less-preferred prey species
 146 showed little variation in the dynamics of x as we varied p (Alternative GP-EDM set up 2
 147 in Fig. S15). For this species, the GP-EDM model gave a strong weight to process noise

148 while decreasing the importance of the control parameter. To circumvent this problem,
 149 we fixed the inverse length scale associated with p to 1 for the experimental data and
 150 optimized the remaining ϕ_j parameters as described above. Our rationale for doing this is
 151 that, since we know that the control parameter must have an effect on x , then our prior
 152 knowledge for it can be added to the GP-EDM model. For the experimental system,
 153 we are sure that the control parameter (dilution rate) is important, so fixing the ϕ_j
 154 corresponding to the control parameter to 1 is well justified.

155 S4 Power spectrum and largest Lyapunov exponent

156 To investigate which type of regime is being predicted through our GP-EDM ap-
 157 proach, we computed the power spectrum and largest Lyapunov exponent of true and
 158 predicted time series. We performed this analysis for two cases: (1) true and predicted
 159 time series from the two-species predator-prey model from Fig. 2 in the main text (equa-
 160 tion [S3]) and (2) predicted time series from the lake planktonic food web from Fig.
 161 5 in the main text. For each case, we performed these analyses for each time series
 162 under different values of the control parameter p . For the two-species predator-prey
 163 model, we performed these analyses for $p = 2.7, 2.8, 2.9, 3.0, 3.1, 3.2, 3.3, 3.4$, where p is
 164 the predator attack rate. For the lake planktonic food web, we performed these analyses
 165 for $p = 0.02, 0.03, 0.04, 0.05, 0.06, 0.07$, where p is the phosphate concentration. For a
 166 given species i under a given value of p , we will refer to the true time series as $\{x_i(t)\}$
 167 ($t = 1, \dots, T$) and to the predicted time series as $\{\hat{x}_i(t)\}$ ($t = 1, \dots, T$), where $T = 100$.
 168 Note that we have the true time series for the two-species predator-prey model but not
 169 for the lake planktonic food web. Results for the two-species predator-prey model are
 170 shown in Fig. S8 and for the lake planktonic food web in Fig. S22.

171 To compute power spectra, we followed the procedure from (10). That is, we computed
 172 power spectra for each time series via ridge regression onto sine and cosine basis functions
 173 with frequencies $\frac{2\pi k}{T}$, where $k = 1, \dots, T/2$. Each time series was standardized to mean
 174 zero and unit standard deviation prior to decomposition and the ridge regression penalty
 175 was 0.01. Power at each frequency was calculated from the sine and cosine coefficients.

176 The computation of the largest Lyapunov exponent relies on the sequence of Jacobian
 177 matrices $\mathbf{J}(t)$ evaluated at successive states (10, 11). For the two-species predator-prey
 178 model, these matrices can be analytically computed from the model equations (i.e., equa-
 179 tion [S3]) when performing the analyses with the true time series. However, for predicted
 180 time series, we estimated the sequential Jacobian matrices from the fitted GP-EDM
 181 model, as we will explain in the next paragraph. For the true time series, we evaluated
 182 the following 2×2 Jacobian matrix $\mathbf{J}(t)$ at each state (i.e., $x_1(t)$ and $x_2(t)$ at each t):

$$\mathbf{J}(t) = \begin{bmatrix} \frac{\partial x_1(t+1)}{\partial x_1(t)} & \frac{\partial x_1(t+1)}{\partial x_2(t)} \\ \frac{\partial x_2(t+1)}{\partial x_1(t)} & \frac{\partial x_2(t+1)}{\partial x_2(t)} \end{bmatrix}, \quad [S13]$$

183 where the matrix coefficients were calculated analytically from equation [S3].

184 For predicted time series, we estimated the sequential Jacobian matrices from the
 185 fitted GP-EDM model. We did so by extracting the gradient $\frac{\partial x_i(t+1)}{\partial x_i(t-j)}$ at time t for the
 186 j th model input, where $j = 0, \dots, E - 1$. For the two-species predator-prey model, these
 187 inputs were four lags of prey abundance: $x_i(t), x_i(t - 1), x_i(t - 2), x_i(t - 3)$ with $i = 1$.
 188 For the lake planktonic food web, these inputs were four lags of green algae abundance:
 189 $x_i(t), x_i(t - 1), x_i(t - 2), x_i(t - 3)$ (Table S2). From the GP-EDM gradient at a given time
 190 t , we built the following $E \times E$ Jacobian matrix:

$$\mathbf{J}(t) = \begin{bmatrix} \frac{\partial x_i(t+1)}{\partial x_i(t)} & \frac{\partial x_i(t+1)}{\partial x_i(t-1)} & \cdots & \frac{\partial x_i(t+1)}{\partial x_i(t-E-1)} \\ 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 1 & 0 \end{bmatrix}. \quad [\text{S14}]$$

191 This is a standard structure of the Jacobian matrix when conducting analyses with time-
 192 delay embedding (10, 11).

193 To compute the largest Lyapunov exponent, we followed a QR decomposition ap-
 194 proach using the sequence of Jacobian matrices from time $t = 1$ to $t = 100$ (12). We used
 195 the same approach for analytical (i.e., true time series) and inferred (i.e., predicted time
 196 series) Jacobian matrices. The largest Lyapunov exponent corresponds to the asymptotic
 197 growth rate of an infinitesimal perturbation $\mathbf{y}(t)$ to the state $\mathbf{x}(t)$ of the system. The
 198 state of the system can be either in native (i.e., $\mathbf{x}(t) = [x_1(t), x_2(t), \dots, x_n(t)]$) or delay
 199 (i.e., $\mathbf{x}(t) = [x_i(t), x_i(t - 1), \dots, x_i(t - E)]$) coordinates. Importantly, the transformation
 200 from native to delay coordinates preserves the largest Lyapunov exponent (12, 13). Be-
 201 cause the dynamics of the perturbation $\mathbf{y}(t)$ are linear, we can use the following matrix
 202 difference equation: $\mathbf{Y}(t + 1) = \mathbf{J}(t)\mathbf{Y}(t)$, where $\mathbf{J}(t)$ is the Jacobian matrix evaluated at
 203 $\mathbf{x}(t)$ and $\mathbf{Y}(t)$ is a matrix with each column containing a given perturbation vector $\mathbf{y}(t)$.
 204 The solution of this system from time 0 to T is given by a product of $T - 1$ successive
 205 Jacobian matrices:

$$\mathbf{Y}(T) = \mathbf{J}(T - 1) \cdot \mathbf{J}(T - 2) \cdot \dots \cdot \mathbf{J}(1) \cdot \mathbf{J}(0) \cdot \mathbf{Y}(0). \quad [\text{S15}]$$

206 If the initial condition is the identity matrix ($\mathbf{Y}(0) = \mathbf{I}$), then the parallelepiped spanned
 207 by the perturbation vectors has sides with norm 1 (i.e., $\|\mathbf{y}(0)\| = 1$). After T time steps,
 208 the parallelepiped will be deformed and the rates at which each of its sides get stretched
 209 correspond to the Lyapunov exponents (12).

210 To quantify the rates at which the parallelepiped sides get stretched, we can lever-
 211 age the QR decomposition. Specifically, the QR decomposition of $\mathbf{Y}(t)$ is given by:
 212 $\mathbf{Y}(t) = \mathbf{Q}(t)\mathbf{R}(t)$, where $\mathbf{Q}(t)$ is an orthonormal matrix and $\mathbf{R}(t)$ is an upper triangu-
 213 lar matrix with non-negative diagonal elements. Because $\mathbf{Q}(t)$ only describes rotations
 214 and reflections, the diagonal elements of $\mathbf{R}(t)$ contain all the information about how

the columns vectors of $\mathbf{Y}(t)$ get stretched. For large t , the i th Lyapunov exponent will be given by $\frac{1}{t} \ln[r_{ii}(t)]$, where $r_{ii}(t)$ is the i th diagonal element of $\mathbf{R}(t)$ (12). To avoid numerical issues of multiplying $T - 1$ Jacobian matrices and then performing the QR decomposition, we performed the decomposition at each time step and then computed the time-average of each diagonal element of the \mathbf{R} matrices. That is, we performed the following sequential QR decompositions:

$$\begin{aligned}\mathbf{Y}(T) &= \mathbf{J}(T-1) \cdot \dots \cdot \mathbf{J}(1) \cdot \mathbf{J}(0) \cdot \mathbf{Y}(0) \\ &= \mathbf{J}(T-1) \cdot \dots \cdot \mathbf{J}(1) \cdot \mathbf{Q}(1) \cdot \mathbf{R}(1) \\ &= \mathbf{J}(T-1) \cdot \dots \cdot \mathbf{Q}(2) \cdot \mathbf{R}(2) \cdot \mathbf{R}(1) \\ &= \mathbf{Q}(T) \cdot \mathbf{R}(T) \cdot \dots \cdot \mathbf{R}(3) \cdot \mathbf{R}(2) \cdot \mathbf{R}(1),\end{aligned}\quad [\text{S16}]$$

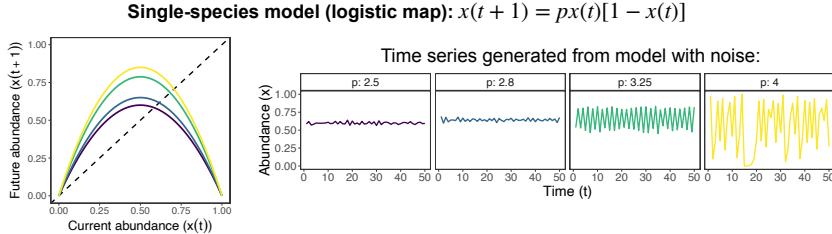
where $\mathbf{Q}(1)\mathbf{R}(1)$ is the decomposition of $\mathbf{J}(0)\mathbf{Y}(0)$, $\mathbf{Q}(2)\mathbf{R}(2)$ is the decomposition of $\mathbf{J}(1)\mathbf{Q}(1)$, and so on. Using the fact that the diagonal elements of a product of triangular matrices is the product of their diagonal elements, we computed $\frac{1}{t} \ln[r_{ii}(t)]$ as:

$$\frac{1}{T} \ln \left[\prod_{j=1}^T r_{ii}(j) \right] = \frac{1}{T} \sum_{j=1}^T \ln[r_{ii}(j)].\quad [\text{S17}]$$

The largest Lyapunov exponent corresponds to the largest value of $\frac{1}{t} \ln[r_{ii}(t)]$ across all i , where $i = 1, \dots, n$ under native coordinates and $i = 1, \dots, E$ under delay coordinates. If the largest Lyapunov exponent is greater than zero, an infinitesimal perturbation will grow in the long term, which implies chaotic dynamics.

228 **References**

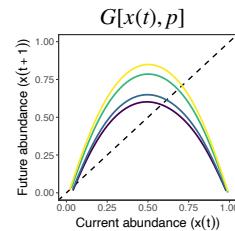
- 229 [1] RM May, Simple mathematical models with very complicated dynamics. *Nature*
230 **261**, 459–467 (1976).
- 231 [2] ML Rosenzweig, RH MacArthur, Graphical representation and stability conditions
232 of predator-prey interactions. *The American Naturalist* **97**, 209–223 (1963).
- 233 [3] WE Ricker, Stock and recruitment. *Journal of the Fisheries Board of Canada* **11**,
234 559–623 (1954).
- 235 [4] A Hastings, T Powell, Chaos in a three-species food chain. *Ecology* **72**, 896–903
236 (1991).
- 237 [5] SB Munch, V Poynor, JL Arriaza, Circumventing structural uncertainty: a bayesian
238 perspective on nonlinear forecasting for ecology. *Ecological Complexity* **32**, 134–143
239 (2017).
- 240 [6] TL Rogers, SB Munch, Hidden similarities in the dynamics of a weakly synchronous
241 marine metapopulation. *Proceedings of the National Academy of Sciences* **117**, 479–
242 485 (2020).
- 243 [7] CE Rasmussen, CKI Williams, *Gaussian process for machine learning*. (The MIT
244 Press), (2006).
- 245 [8] M Blum, MA Riedmiller, Optimization of gaussian process hyperparameters using
246 rprop. in *ESANN*. Vol. 339, p. 344 (2013).
- 247 [9] S Munch, T Rogers, Gpedm: Gaussian process regression for empirical dynamic
248 modeling. *R package version 0.0. 0.9005* (2021).
- 249 [10] TL Rogers, SB Munch, SiS Matsuzaki, CC Symons, Intermittent instability is
250 widespread in plankton communities. *Ecology Letters* **26**, 470–481 (2023).
- 251 [11] TL Rogers, BJ Johnson, SB Munch, Chaos is not rare in natural ecosystems. *Nature
252 Ecology & Evolution* **6**, 1105–1111 (2022).
- 253 [12] G Datseris, U Parlitz, Dynamical systems in *Nonlinear Dynamics: A Concise Intro-
254 duction Interlaced with Code*. (Springer), pp. 1–19 (2022).
- 255 [13] JH Argyris, G Faust, M Haase, R Friedrich, *An exploration of dynamical systems
256 and chaos: completely revised and enlarged second edition*. (Springer), (2015).
- 257 [14] E Merz, et al., Disruption of ecological networks in lakes by climate change and
258 nutrient fluctuations. *Nature Climate Change* **13**, 389–396 (2023).



Step 1 Fit a Gaussian Process regression model with lagged abundances and p as inputs
(e.g., GPEDM package in R) to approximate $x(t+1) = G[x(t), \dots, x(t-E), p]$

data
(50 points for each level of p , 200 points in total)
(In this example, $E = 0$)

t	x	x_1	p
1	0.587	NA	2.5
2	0.620	0.587	2.5
3	0.571	0.620	2.5
...
1	0.691	NA	2.8
2	0.598	0.691	2.8
3	0.678	0.598	2.8
...



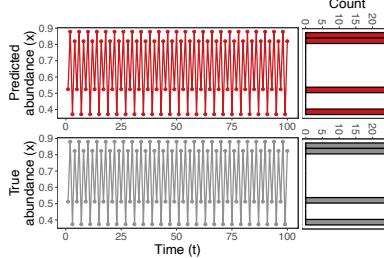
```
G <- fitGP(data = data, y = "x", x = c("x_1", "p"))
```

(See SI Appendix Section S3 for a mathematical description of the approximated function G)

Step 2 Sequentially predict x by interpolating function G to an unseen level of p (e.g., $p = 3.52$)

```
for (t in 1:100) {
  newdata <- data.frame(x_1 = 0.524, p = 3.52)
  x[t] <- predict(object = G, newdata = newdata)$outsampresults$predmean
  newdata <- c(x[t], newdata$p)
}
```

Step 3 Compare true and predicted time series of x (e.g., Jensen-Shannon divergence) for each level of p



(In this example, $p = 3.52$)

Jensen-Shannon divergence is computed by comparing these two histograms of x

Fig. S1. Worked out example of our GP-EDM approach. This figure is similar to Fig. 1 in the main text, but provides more details on the computational steps of our approach using the single-species logistic model. The top panel shows the time series generated from the model at four different levels of the intrinsic growth rate (p). On the left, we show the relationship between $x(t+1)$ and $x(t)$ for different values of p . **Step 1:** fit a Gaussian Process regression model with lagged abundances and p as inputs to approximate the function G . In this example, since we only have one species, we only need one lag, that is: $x(t+1) = G[x(t), p]$. We also show the structure of the data that is used to fit the GP-EDM model and the command to do so in the GPEDM R package. On the right, we show the function G that was inferred from the data. **Step 2:** Sequentially predict x by interpolating function G to an unseen level of p . For instance, for $p = 3.52$ (which is not in the observed data) and a given initial condition $x(0) = 0.524$, we make a prediction for $x(1)$ using the fitted model. We then use the predicted point $x(1)$ to make a prediction for $x(2)$ and so on until we reach $t = 100$. We show a simplified version of the commands to do so with the GPEDM R package. **Step 3:** Compare true and predicted time series of x for each level of p . As an example, the plots show true and predicted time series (i.e., dynamical regimes) for $p = 3.52$. We can then compare the two distributions using a metric such as the Jensen-Shannon divergence.

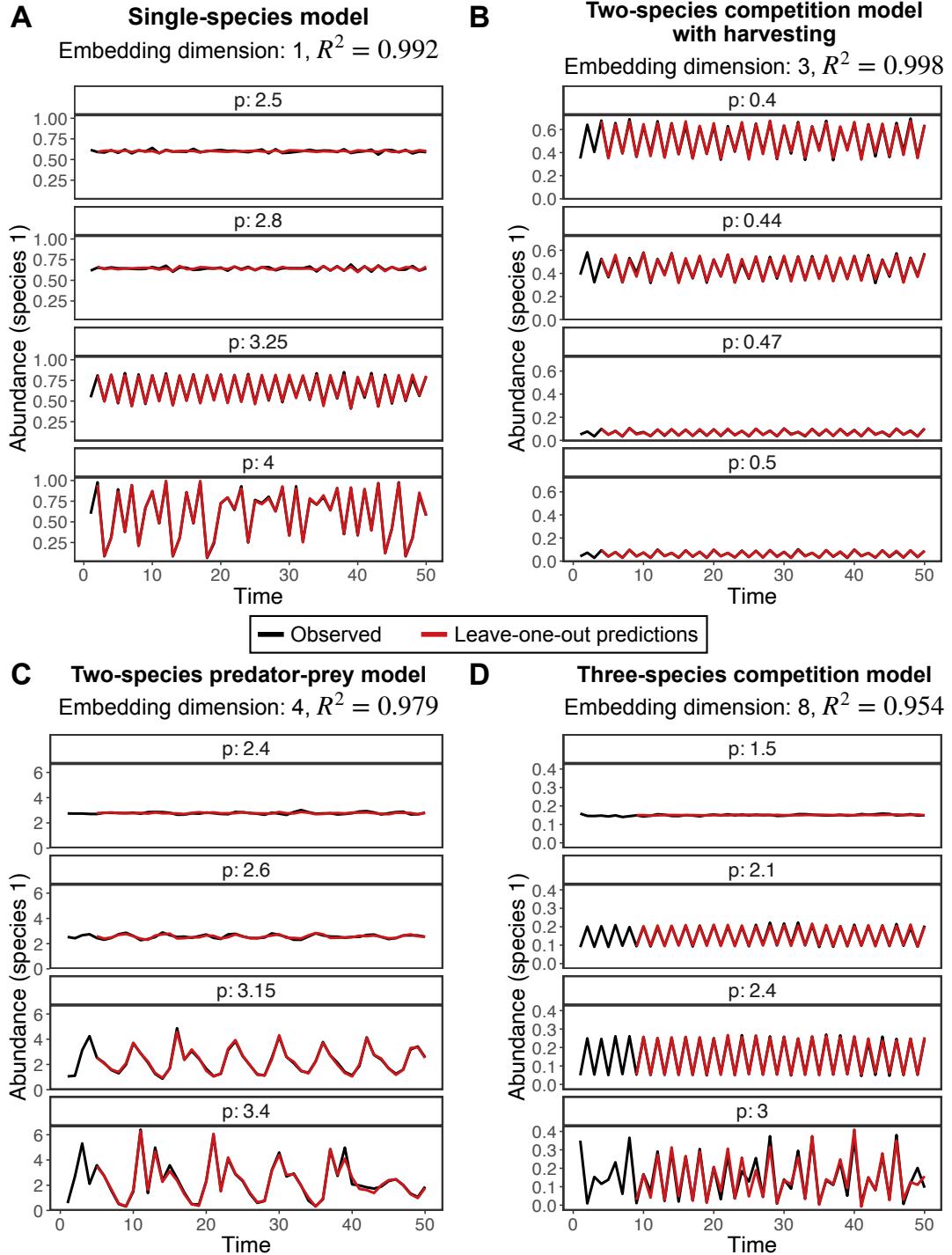


Fig. S2. Time series of species 1 (x_1 , in black) generated from four discrete-time population dynamics models (equations [S1] to [S4]) under four different levels of a control parameter (p) and the leave-one-out predictions (in red) from GP-EDM. Above each plot, we list the model used as well as the embedding dimension ($E + 1$) that led to the best prediction accuracy and the associated R^2 value.

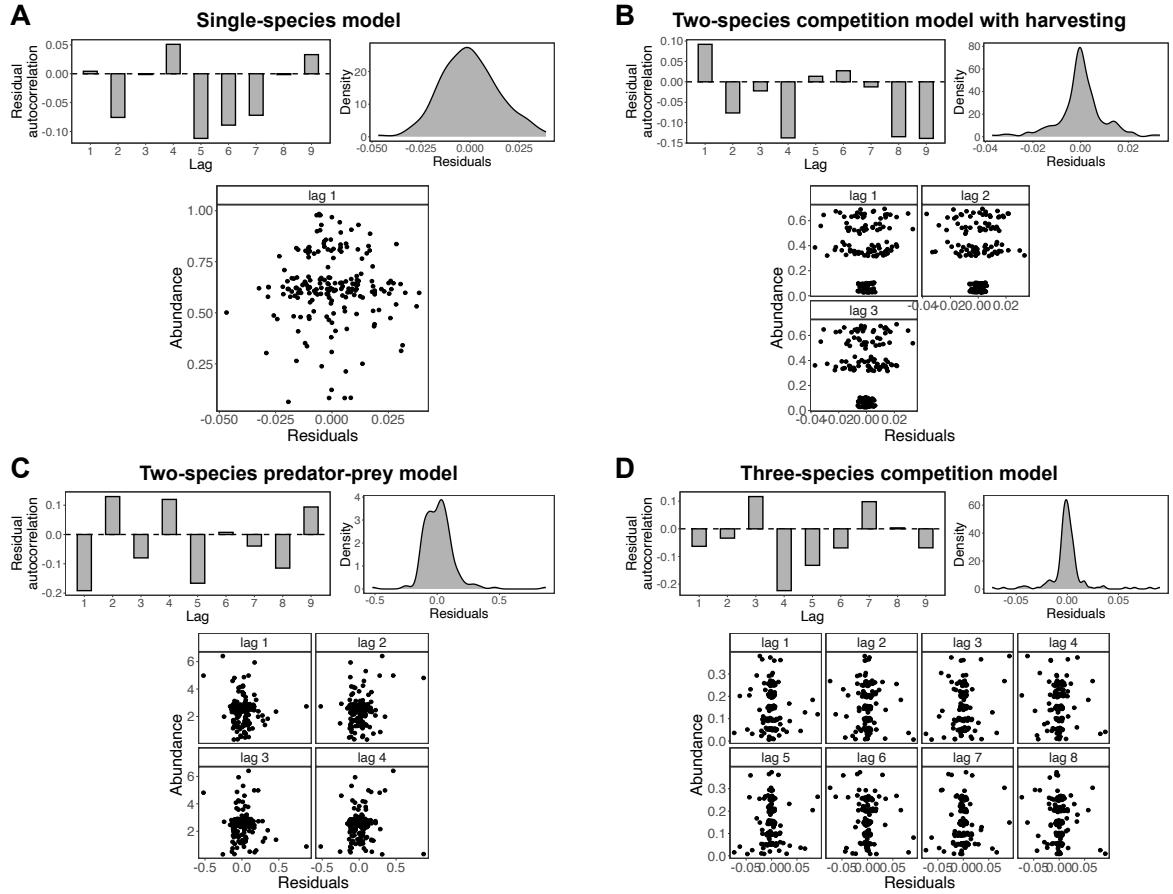


Fig. S3. Plots of GP-EDM model residuals (i.e., observed minus predicted abundance) for the four discrete-time population dynamics models. For each panel, the plot on the top left shows the autocorrelation of residuals at different lags. The plot on the top right shows the distribution of residuals. The plot on the bottom shows the residuals as a function of model inputs (i.e., lagged species abundances). All plots show results for species 1 (i.e., x_1).

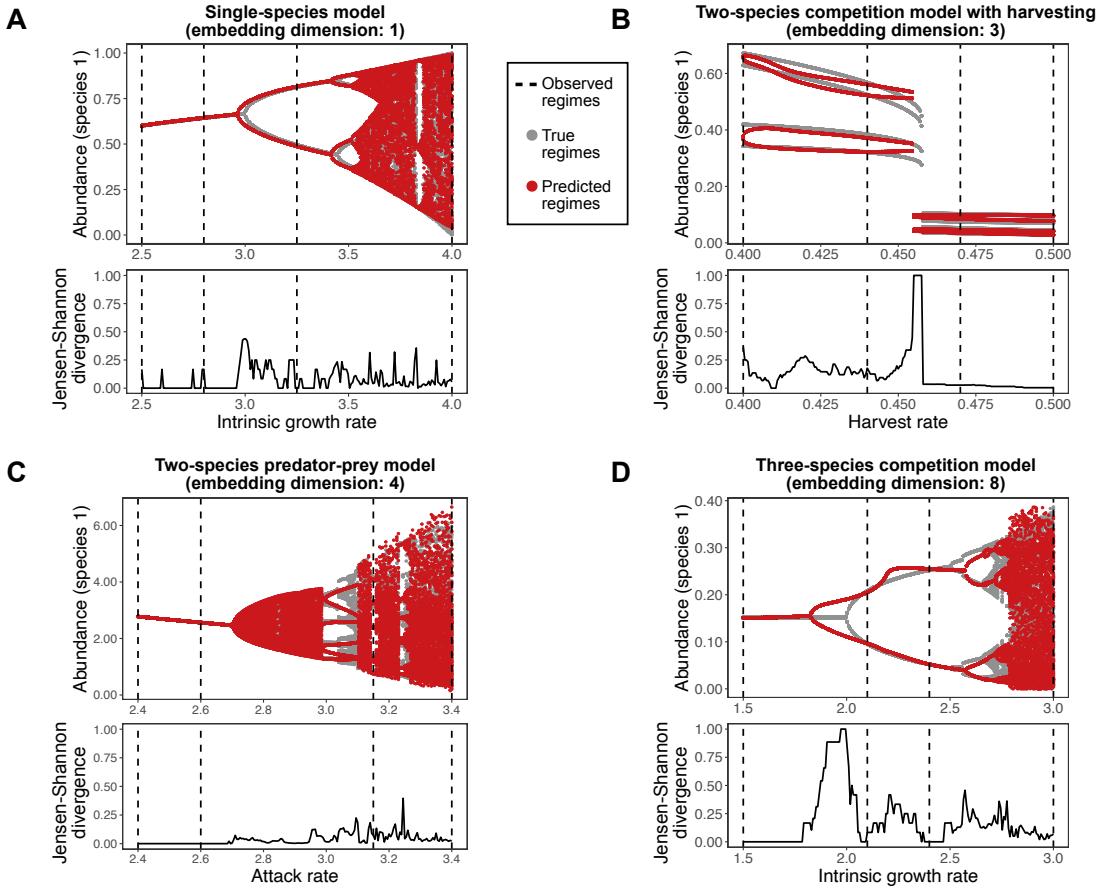


Fig. S4. Reconstruction of bifurcation diagrams for all four discrete-time population dynamics models. This figure is the same as Fig. 2 in the main text, but includes all four models (*SI Appendix Section S1*). **A-D**, Top panels show the true (in gray) and predicted (in red) bifurcation diagrams. Bifurcation diagrams depict the population abundance of a given species after the transient period (y-axis) for a range of control parameter values (p , x-axis). Bottom panels show the Jensen-Shannon divergence (JSD) computed at each value of p . The JSD ranges from 0 to 1, where 0 represents a close match between true and predicted dynamics. The average JSD across all values of p is: single-species model, 0.075 (**A**); two-species competition model with harvesting, 0.125 (**B**); two-species predator-prey model, 0.037 (**C**); three-species competition model, 0.310 (**D**). In all panels, the vertical dashed lines depict the values of p for which we observed population time series to train the GP-EDM model.

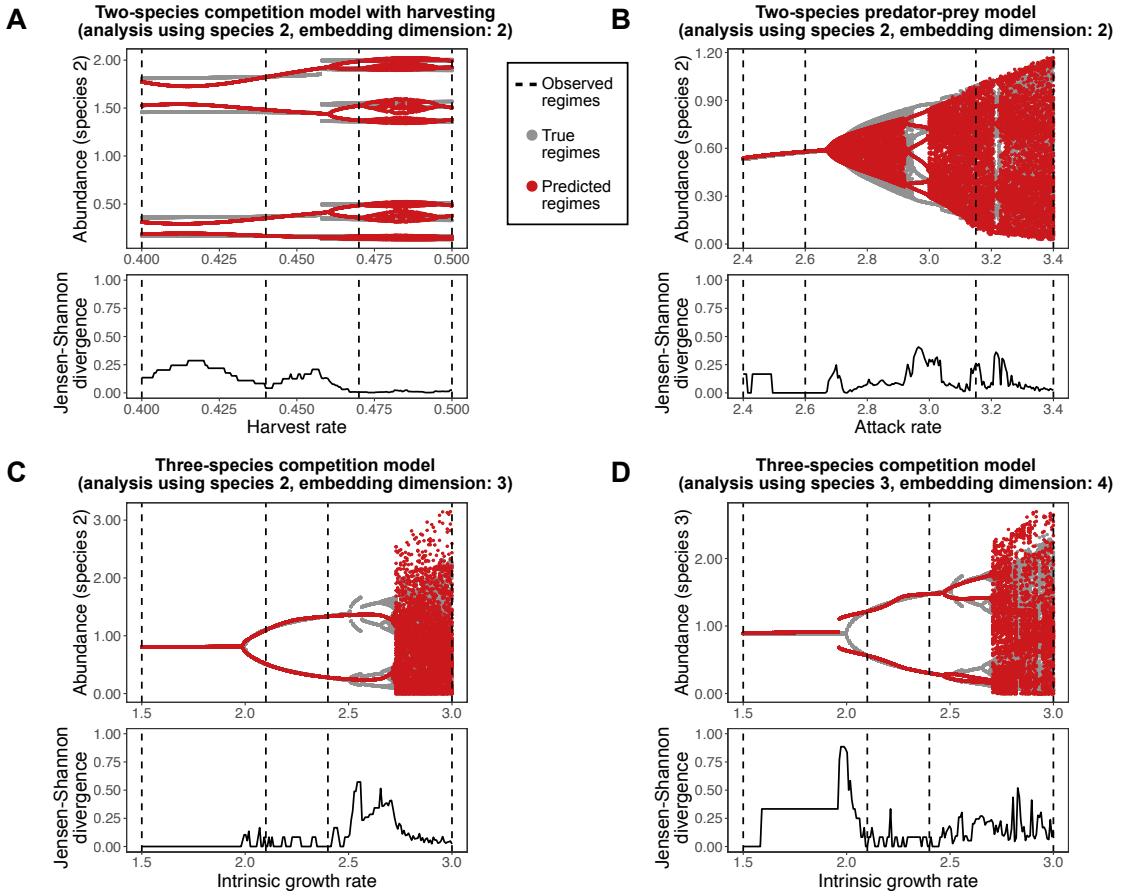


Fig. S5. Reconstruction of bifurcation diagrams for all population dynamics models but performing time-delay embedding with alternative species. In Fig. S4 above, we performed our analysis using species 1 for all models. Here we show results using species 2 with the two-species competition model with harvesting (**A**), using species 2 with the two-species predator-prey model (**B**), using species 2 with the three-species competition model (**C**), and using species 3 with the three-species competition model (**D**).

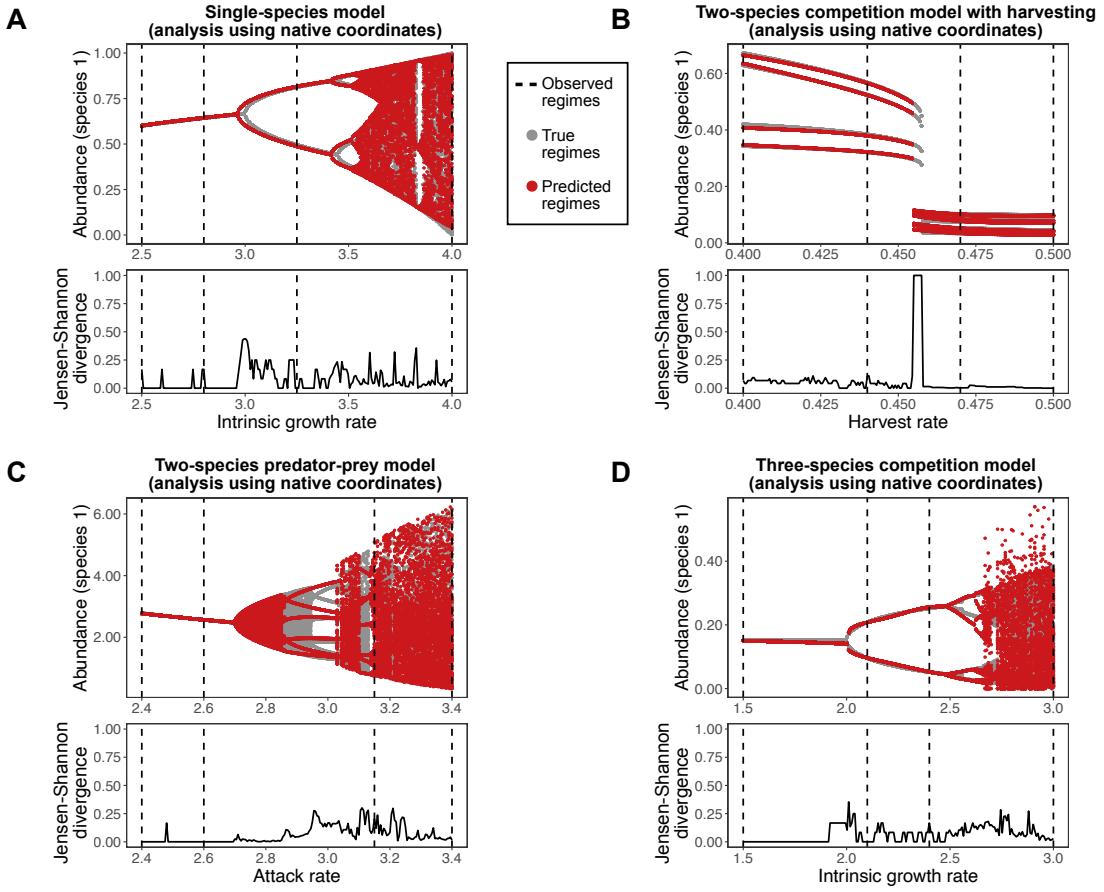


Fig. S6. Reconstruction of bifurcation diagrams for all population dynamics models when we use native instead of delay coordinates. For example, for the two-species predator-prey model (**C**), we use Gaussian Process regression to approximate the functions $x_1(t + 1) = F_1[x_1(t), x_2(t), p]$ and $x_2(t + 1) = F_2[x_1(t), x_2(t), p]$ (i.e., native coordinates) instead of the function $x_1(t + 1) = G_1[x_1(t), x_1(t - 1), x_1(t - 2), x_1(t - 3), p]$ (i.e., delay coordinates with $E = 3$ as in Fig. 2 in the main text). As expected, the average Jensen-Shannon divergence between true and predicted dynamics was lower with native than with delay coordinates for most models (single-species model, 0.075 vs 0.075 (**A**); two-species competition model with harvesting, 0.059 vs 0.125 (**B**); two-species predator-prey model, 0.061 vs 0.037 (**C**); three-species competition model, 0.077 vs 0.310 (**D**)). Results for delay coordinates are shown in Fig. S4 above.

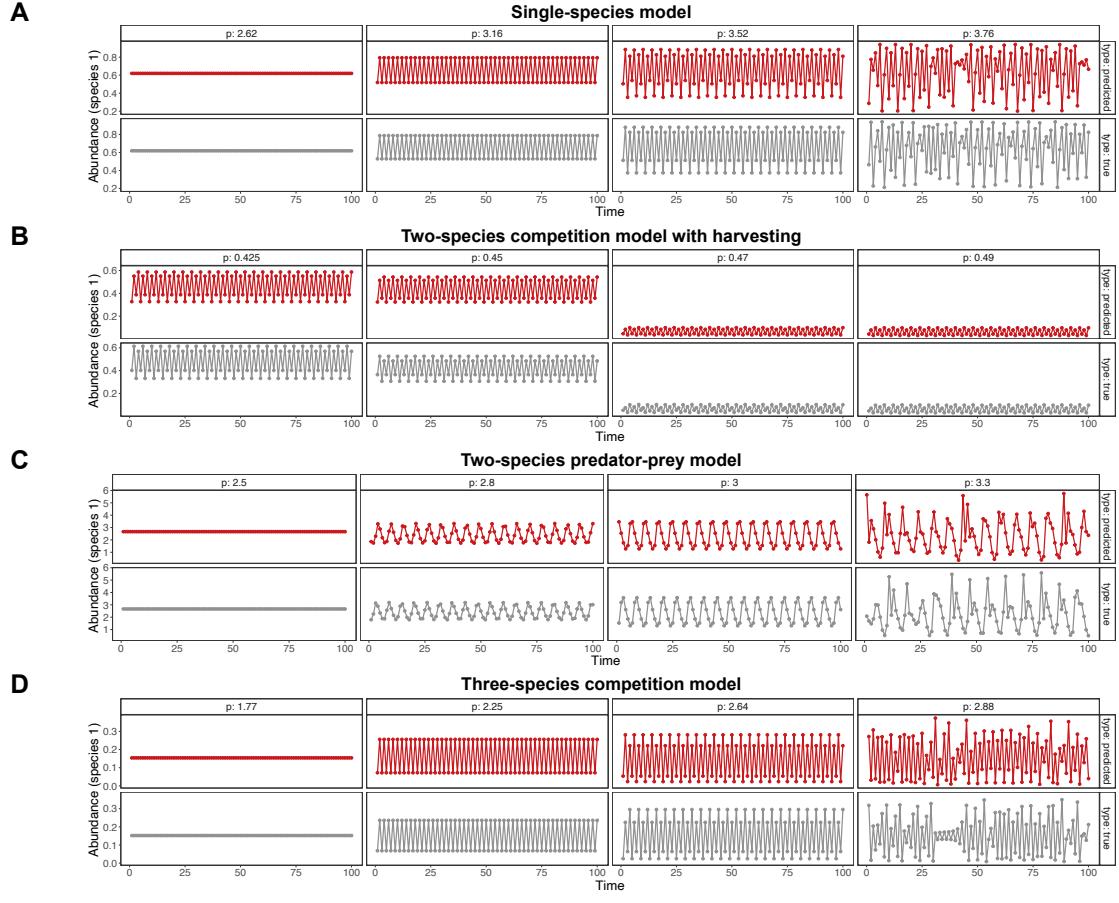
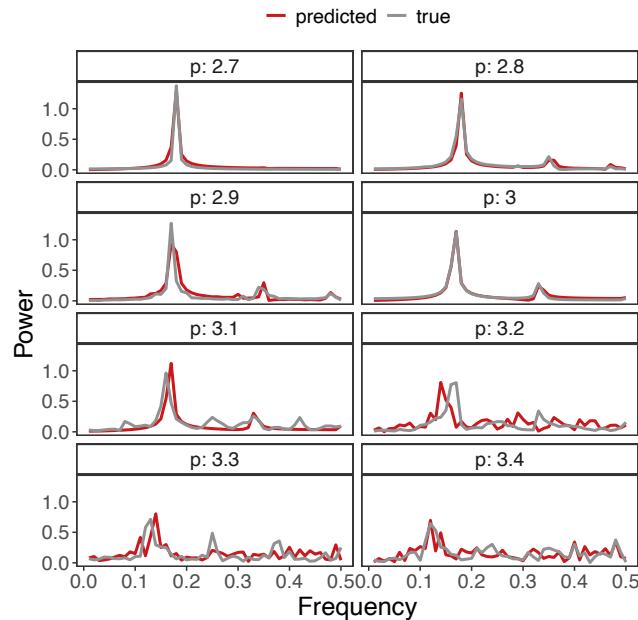


Fig. S7. Examples of predicted time series for all population dynamics models at selected control parameter (p) values. In Fig. 1D in the main text, we show these predicted time series for the two-species predator-prey model. Here, we show examples of these time series for all models under certain values of p . In all cases, the periodicity and time sequence of points of predicted regimes (in red) closely match true regimes (in gray).

A

Power spectrum of true and predicted dynamics

**B**

Largest Lyapunov exponent of true and predicted dynamics

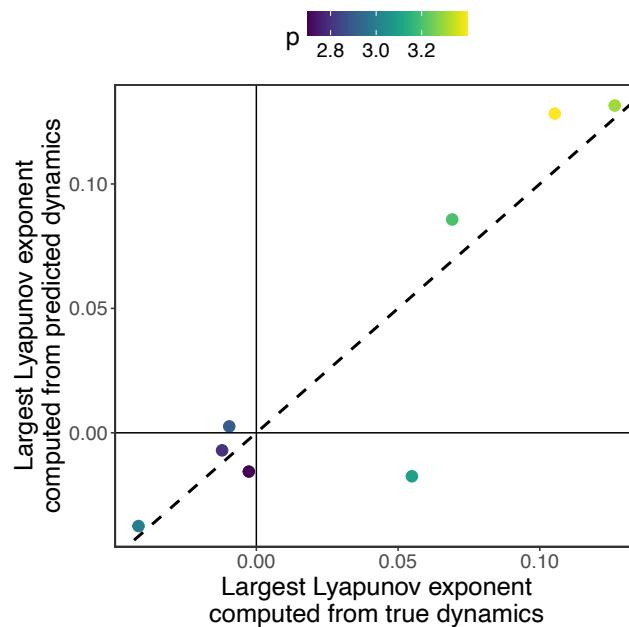


Fig. S8. Power spectrum and largest Lyapunov exponent computed for true and predicted dynamics of species 1 from the two-species predator-prey model under 8 different values of the control parameter p (*SI Appendix Section S4*). **A**, Each panel shows the power spectrum of the true and predicted time series under a given value of p . The peaks of the power spectrum (i.e., main periodicities) are similar for true and predicted dynamics for all values of p . **B**, Each point shows the value of the largest Lyapunov exponent computed using the true (x axis) or predicted (y axis) time series under a given value of p (point color). The dashed 1:1 line denotes a perfect match between true and predicted. An exponent greater than zero denotes chaos. For most values of p , true and predicted values are close to each other. Some of the time series used for these analyses can be seen in Fig. S7 above.

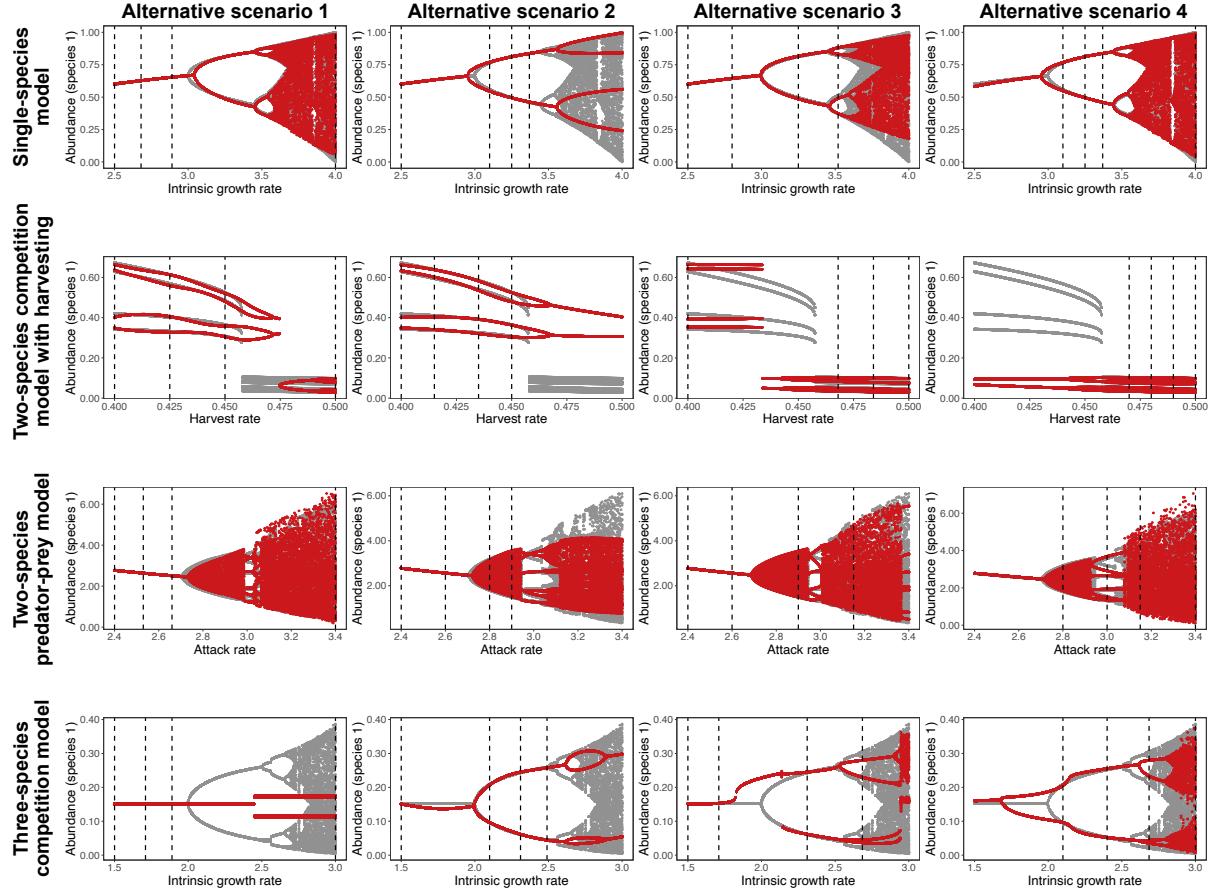


Fig. S9. Reconstruction of bifurcation diagrams for all population dynamics models under four different alternative scenarios of observed dynamical regimes as the training data for GP-EDM. For the single-species model (first row), we use the following scenarios: (1) $p = 2.5, 2.68, 2.89, 4$; (2) $p = 2.5, 3.1, 3.25, 3.37$, (3) $p = 2.5, 2.8, 3.25, 3.52$, and (4) $p = 3.1, 3.25, 3.37, 4$. For the two-species competition model with harvesting (second row), we use the following scenarios: (1) $p = 0.4, 0.425, 0.45, 0.5$; (2) $p = 0.4, 0.415, 0.435, 0.45$, (3) $p = 0.4, 0.468, 0.484, 0.5$, and (4) $p = 0.47, 0.48, 0.49, 0.5$. For the two-species predator-prey model (third row), we use the following scenarios: (1) $p = 2.4, 2.53, 2.66, 3.4$; (2) $p = 2.4, 2.6, 2.8, 2.9$, (3) $p = 2.4, 2.6, 2.9, 3.15$, and (4) $p = 2.8, 3, 3.15, 3.4$. For the three-species competition model (fourth row), we use the following scenarios: (1) $p = 1.5, 1.71, 1.89, 3$; (2) $p = 1.5, 2.1, 2.31, 2.49$, (3) $p = 1.5, 1.71, 2.31, 2.685$, and (4) $p = 2.1, 2.4, 2.685, 3$. In all plots, regimes present in the training data are depicted as vertical dashed lines. The true dynamics are shown in gray and the predicted dynamics in red.

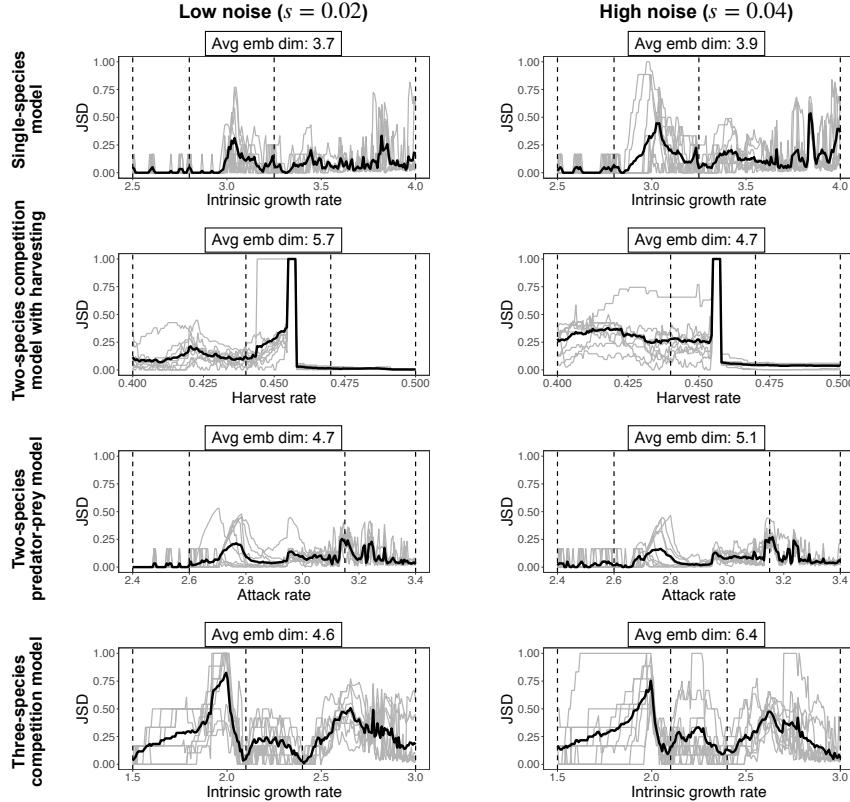
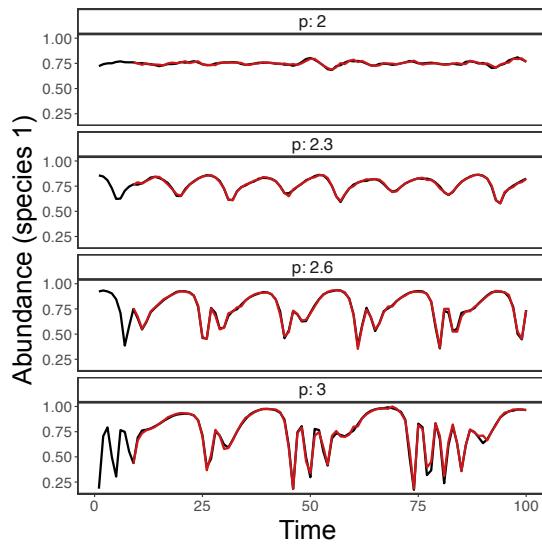


Fig. S10. Jensen-Shannon divergence (JSD) between reconstructed and true bifurcation diagrams for 10 different training data sets and for two levels of process noise for each population dynamics model. The set up for these analyses is exactly the same as the one described in the main text (see Figs. 2 and S4). However, the training data is different because each time we generate the data with process noise, we obtain a different time series. Plots on the left show results for low process noise ($s = 0.02$) and plots on the right show results for high process noise ($s = 0.04$). Each plot shows the JSD computed at each value of the control parameter (p) for a given training data set (each gray line corresponds to one training data set). The black line corresponds to the average JSD across all training data sets. The JSD ranges from 0 to 1, where 0 represents a close match between true and predicted dynamics. In all panels, the vertical dashed lines depict the values of p for which we observed time series to train the GP-EDM model. We also report, at the top of each plot, the average embedding dimension across all 10 GP-EDM models.

A Resource species (species 1), Embedding dimension: 8, $R^2 = 0.986$



B

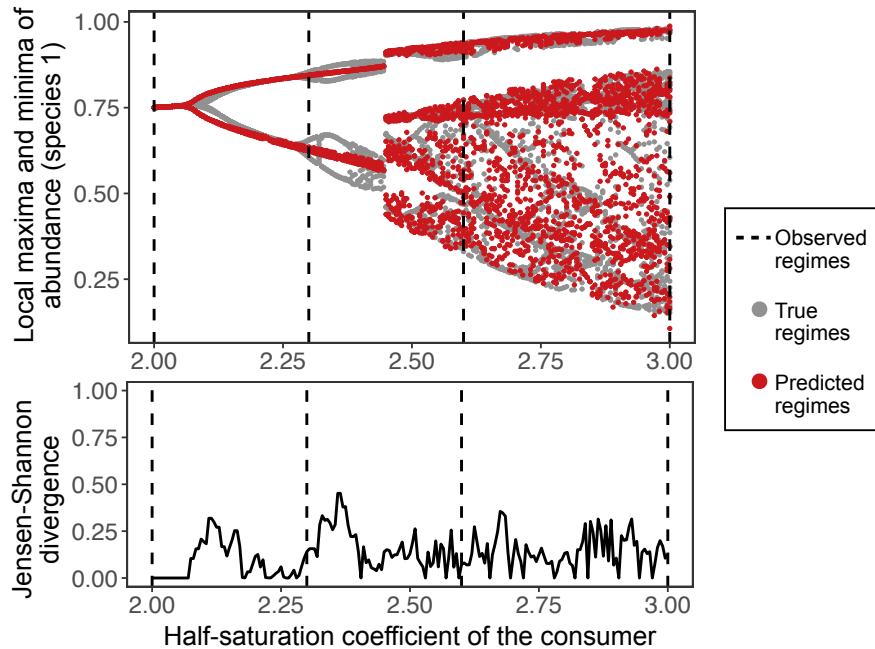


Fig. S11. Results for the continuous-time three-species food chain model (*SI Appendix Section S2*). **A**, Time series of species 1 (x_1 , in black) under four different levels of the half-saturation coefficient of the consumer species (p) and the leave-one-out predictions (in red) from GP-EDM. On top of the plot, we list the embedding dimension ($E + 1$) that led to the best prediction accuracy and the associated R^2 value. **B**, The top plot shows true (in gray) and predicted (in red) bifurcation diagrams. Each bifurcation diagram depicts the local maxima and minima of the abundance of species 1 after the transient period (y-axis) for a range of control parameter values (p , x-axis). The bottom plot shows the Jensen-Shannon divergence (JSD) computed at each value of p . The JSD ranges from 0 to 1, where 0 represents a close match between true and predicted dynamics. In all panels, the vertical dashed lines depict the values of p for which we observed population time series to train the GP-EDM model.

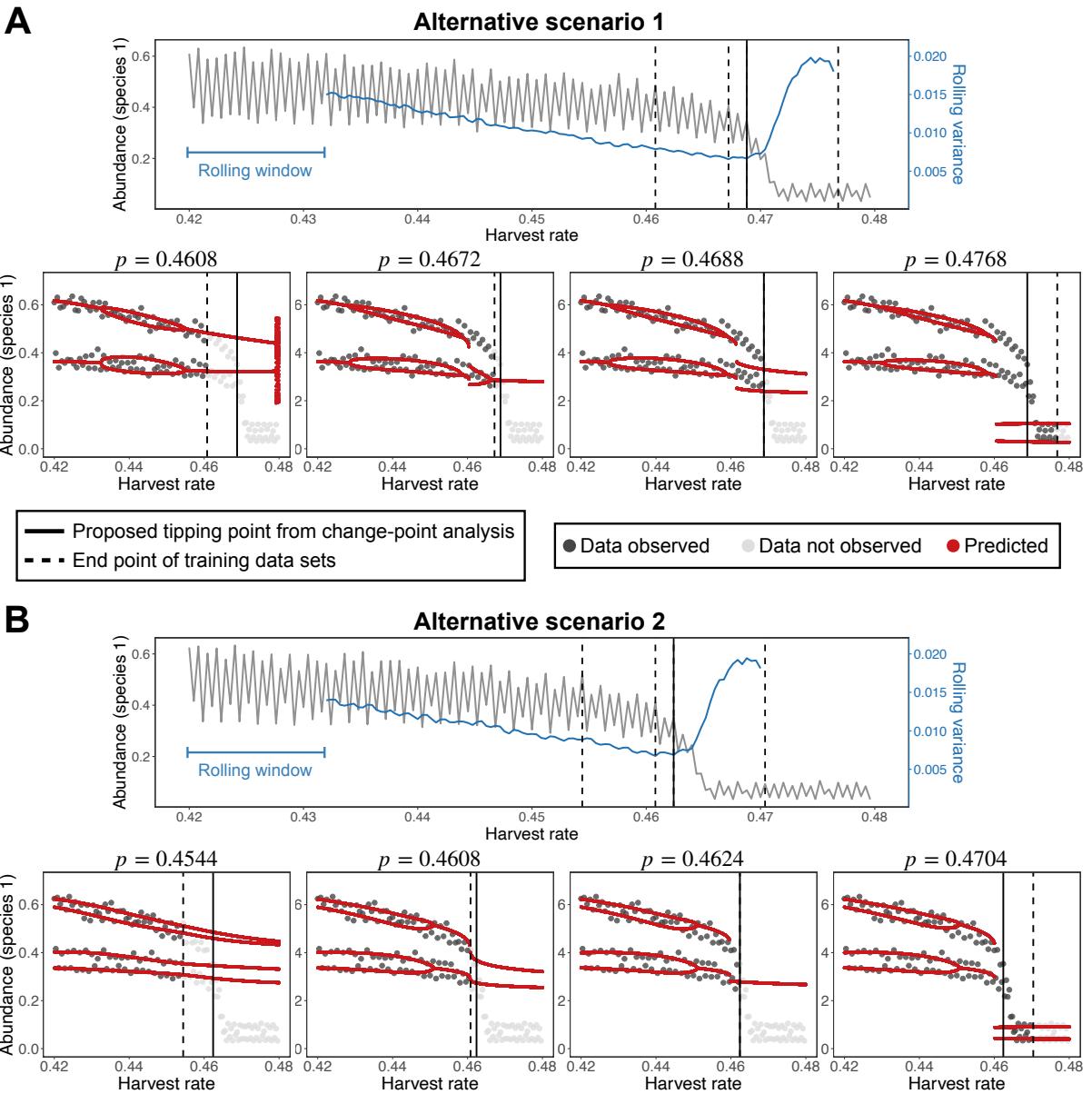


Fig. S12. Illustration of GP-EDM approach as an Early Warning Signal using the two-species competition model with harvesting under two alternative training data sets. The set up for these analyses is exactly the same as the one described in the main text (see Fig. 3). However, the training data is different because the location of the tipping point will change each time we generate the time series due to process noise. The population (in gray) shows a tipping point at $p = 0.4688$ in **A** and at $p = 0.4624$ in **B** (vertical solid lines) according to a change-point analysis. The vertical dashed lines show the end points of the training data sets used with the GP-EDM model. The rolling variance computed using a window with 30 points is shown in blue. The plots below each time series show the predicted bifurcation diagrams (in red) using the GP-EDM model trained up to each of the four dashed lines. The value of p at the end of each training data set is shown above each plot. Gray points denote population abundance values used to train the GP-EDM model (dark gray) or not yet observed (light gray).

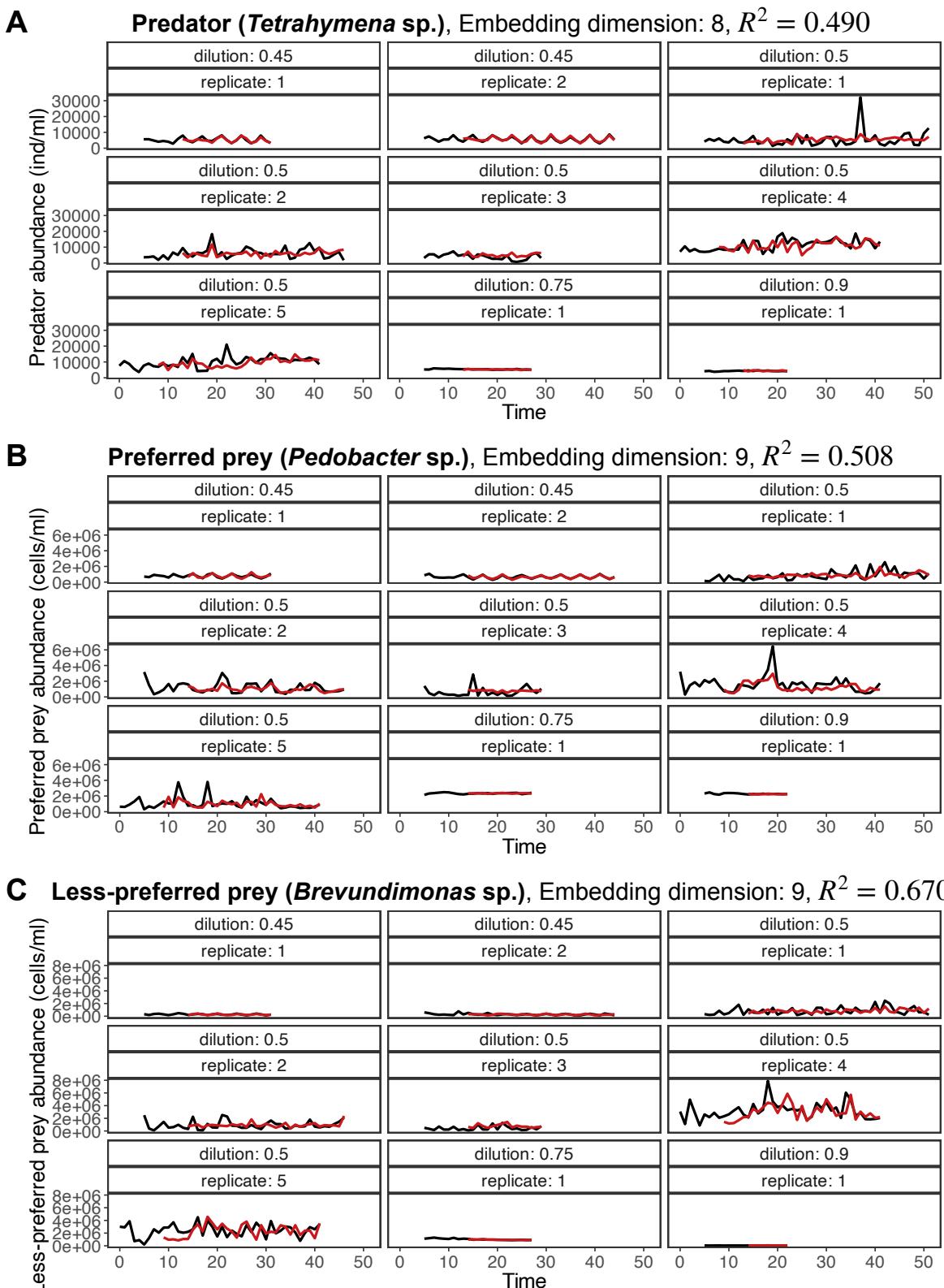


Fig. S13. Time series of each species (A-C) in the experimental microbial food web (in black) under four different dilution rate treatments (0.45/day, 0.5/day, 0.75/day, and 0.9/day) and the leave-one-out predictions (in red) from GP-EDM. For each species, we report the embedding dimension ($E + 1$) that led to the best prediction accuracy and the associated R^2 value.

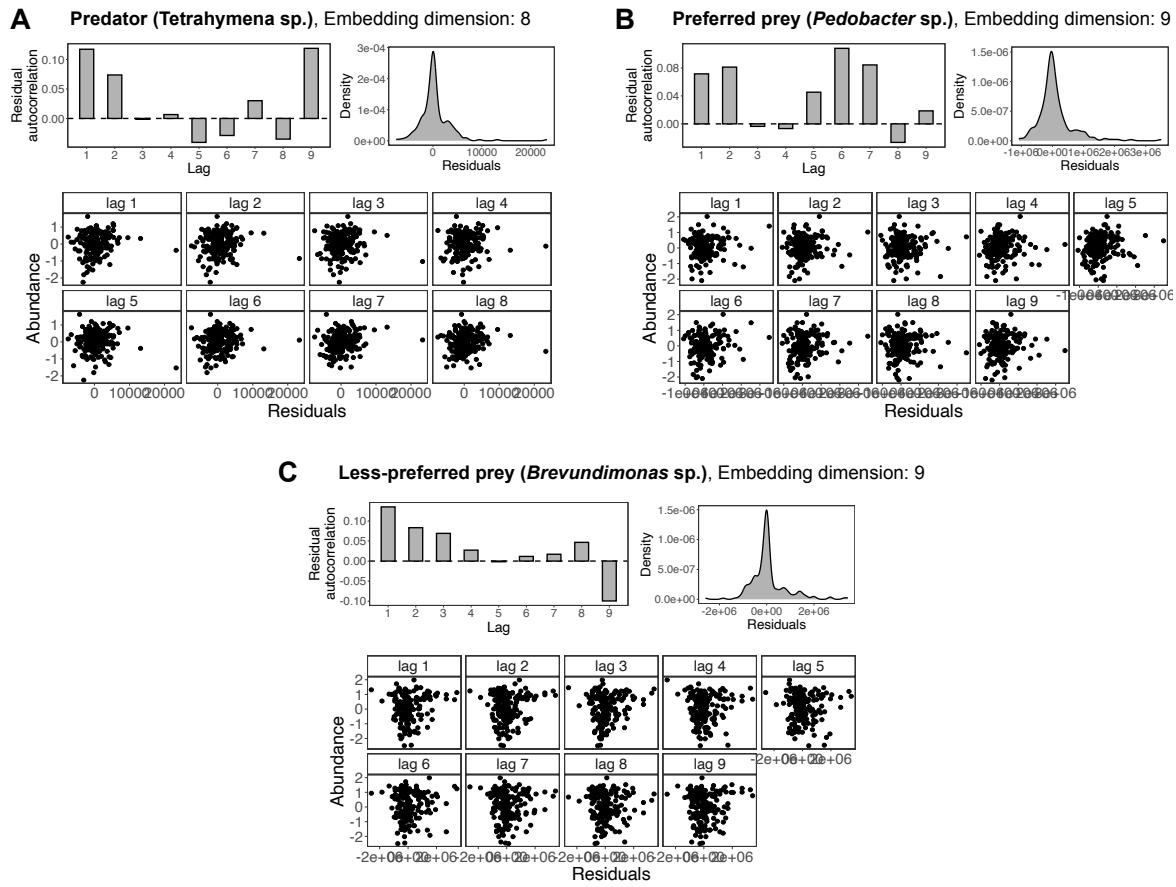


Fig. S14. Plots of GP-EDM model residuals (i.e., observed minus predicted abundance) for the experimental microbial food web. Each panel (A-C) shows the results using a given species to perform time-delay embedding. For each panel, the plot on the top left shows the autocorrelation of residuals at different lags. The plot on the top right shows the distribution of residuals. The plot on the bottom shows the residuals as a function of model inputs (i.e., lagged species abundances).

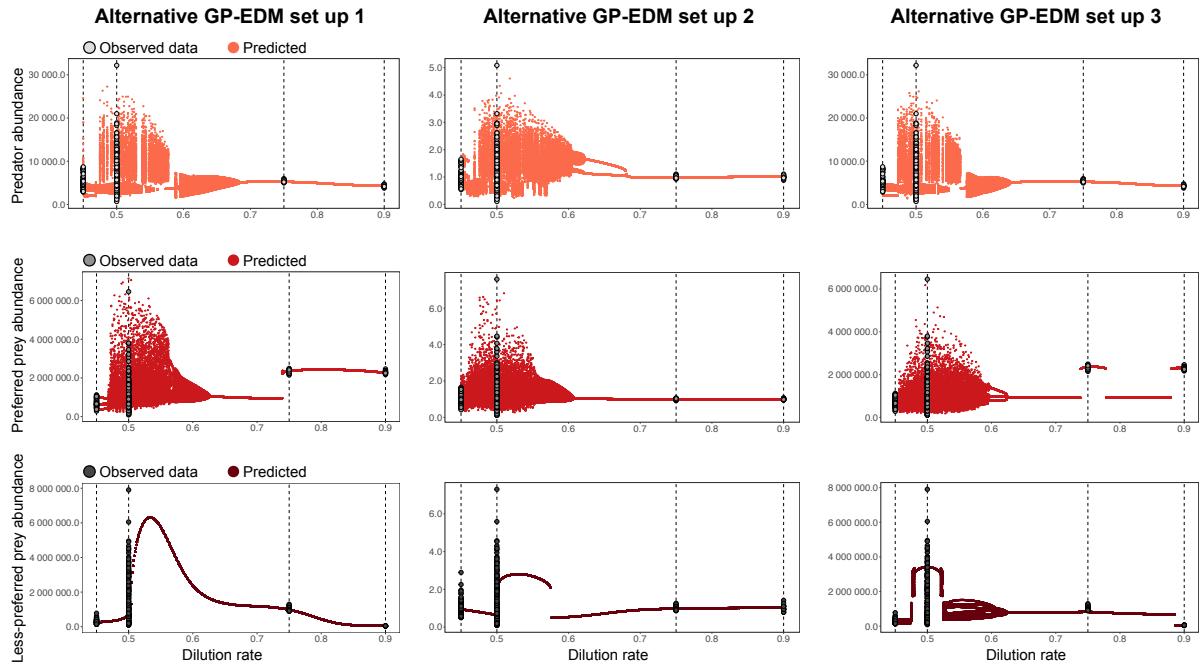


Fig. S15. Reconstructed bifurcation diagrams for the experimental microbial food web using different GP-EDM set ups. Results are qualitatively the same as in Fig. 4 in the main text for the predator and preferred prey species, but not for the less-preferred prey species. The alternative set up 1 is exactly the same as our main text set up (see *Materials and Methods* in the main text), but we do not subtract treatment means from each time series prior to standardizing the data. The alternative set up 2 is exactly the same as our main text set up, but we do not fix the inverse-length scale parameter (ϕ_j) of dilution rate to 1 (see *SI Appendix Section S3*). The alternative set up 3 is exactly the same as our main text set up, but we do not subtract treatment means from each time series prior to standardizing the data and do not fix the inverse-length scale parameter of dilution rate to 1. Note that this last set up is the one used in our analyses with models (i.e., Figs. 2 and S4). Each row of plots shows the analysis for a given species. In all plots, gray points denote the data used to train the GP-EDM model and red points denote predicted dynamical regimes. Vertical dashed lines represent the dilution rates of the experimental treatments (0.45/day, 0.5/day, 0.75/day, and 0.9/day).

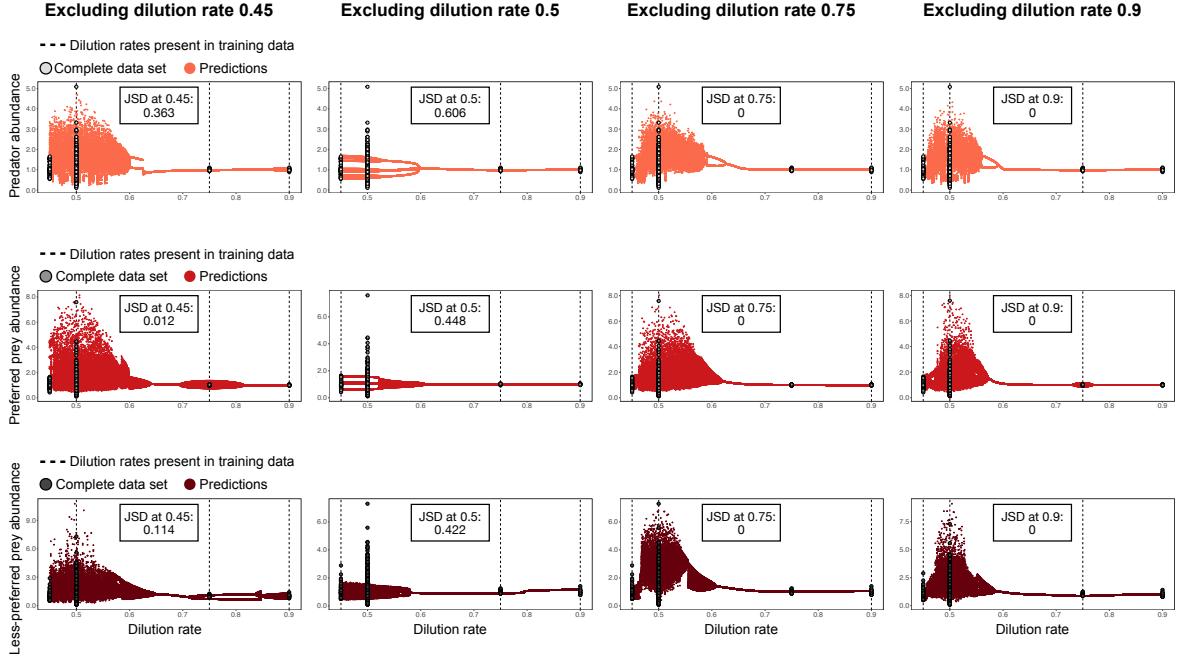


Fig. S16. Reconstructed bifurcation diagrams for the experimental microbial food web when excluding a given dilution rate (p) from the training data prior to fitting the GP-EDM model. For example, in the first column, we excluded all the data for $p = 0.45/\text{day}$ prior to fitting the GP-EDM model. We then reconstructed the entire bifurcation diagram, which included predicting the dynamical regime at $p = 0.45/\text{day}$. The rest of the procedure is exactly the same as the one used for Fig. 4 in the main text. Each row of plots shows the analysis for a given species. In all plots, gray points denote the complete data set and red points denote predicted dynamical regimes. Vertical dashed lines represent the dilution rates that were present in the training data. For each plot, we also report the Jensen-Shannon divergence at the excluded dilution rate.

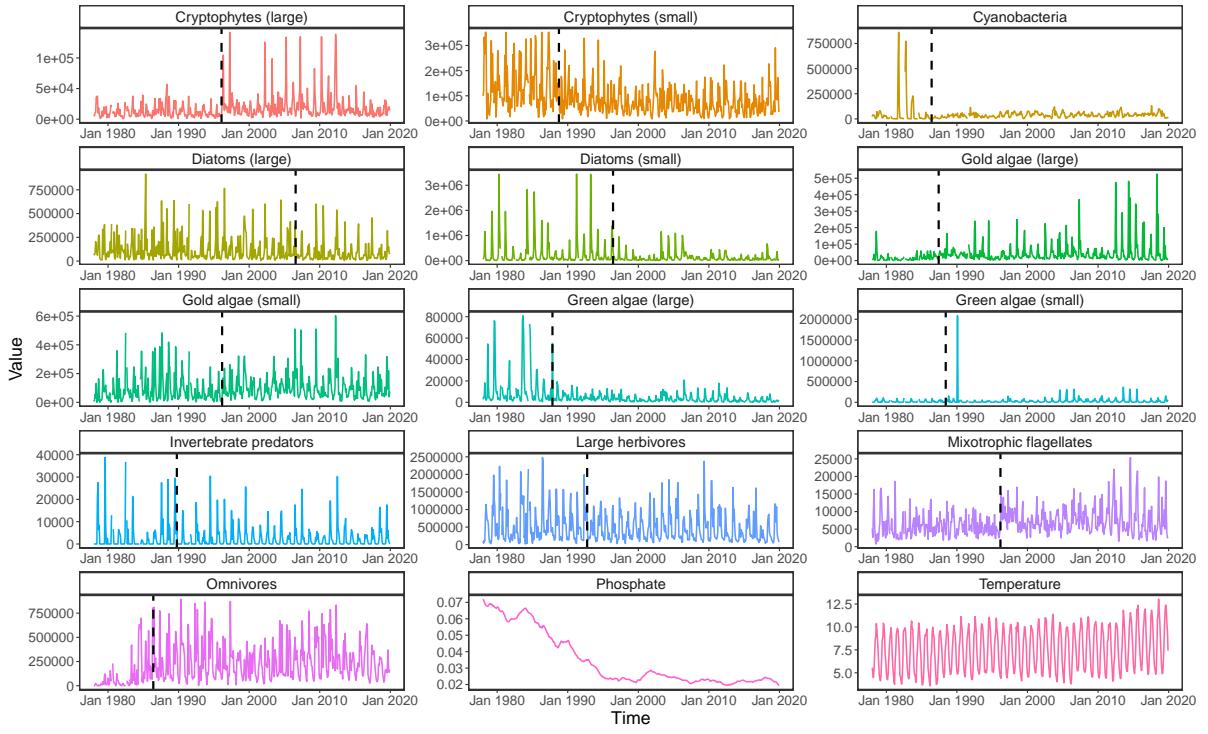
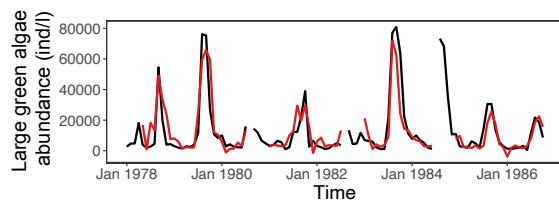


Fig. S17. Monthly time series for all 13 plankton functional groups (in ind/l), phosphate concentration (in mg/l), and water temperature (in Celsius) from Lake Zurich between January 1978 and December 2019. As explained in the *Materials and Methods* section in the main text, we performed a change-point analysis to determine the location of a potential tipping point for each functional group independently. This potential tipping point is shown for each functional group as a black dashed line and Table S1 gives additional information. For Cryptophytes, small refers to $\leq 700 \mu\text{m}^3$ and large to $> 700 \mu\text{m}^3$. For Diatoms, small refers to $\leq 800 \mu\text{m}^3$ and large to $> 800 \mu\text{m}^3$. For Gold algae, small refers to $\leq 250 \mu\text{m}^3$ and large to $> 250 \mu\text{m}^3$. For Green algae, small refers to $\leq 550 \mu\text{m}^3$ and large to $> 550 \mu\text{m}^3$. Additional information about this data set can be found in (14).

ALeave-one-out predictions ($R^2 = 0.77$)**B**

Plots of GP-EDM model residuals

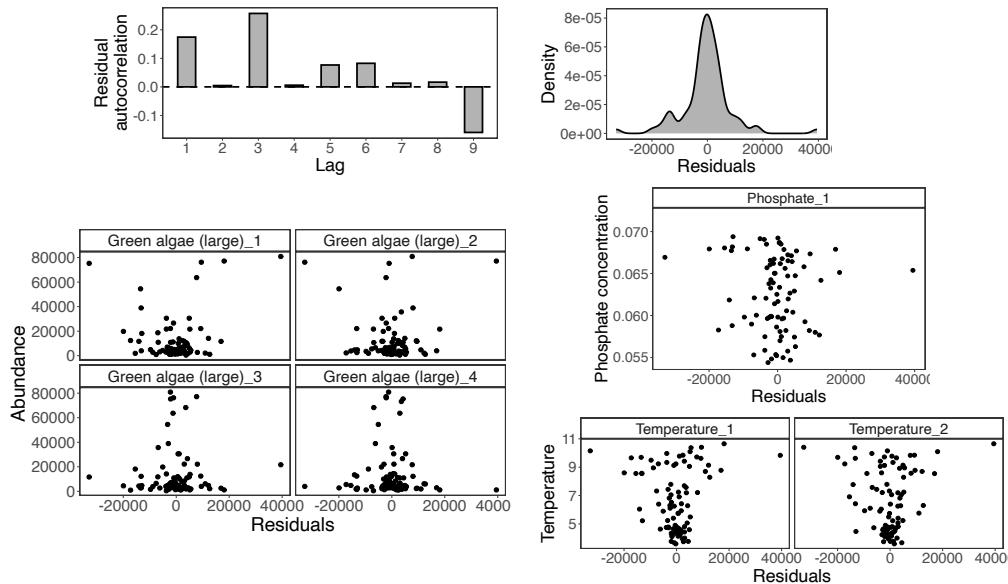


Fig. S18. **A**, Abundance time series of large green algae (in ind/l) for Lake Zurich from January, 1978 to October, 1986 (in black) and leave-one-out predictions (in red) from GP-EDM. The inputs used in this GP-EDM model were 4 lags of large green algae abundance, one lag of phosphate, and two lags of temperature (Table S2). **B**, GP-EDM model residuals (i.e., observed minus predicted abundance) for the leave-one-out predictions shown in **A**. The plot on the top left shows the autocorrelation of residuals at different lags. The plot on the top right shows the distribution of residuals. The plots on the bottom show the residuals as a function of each model input.

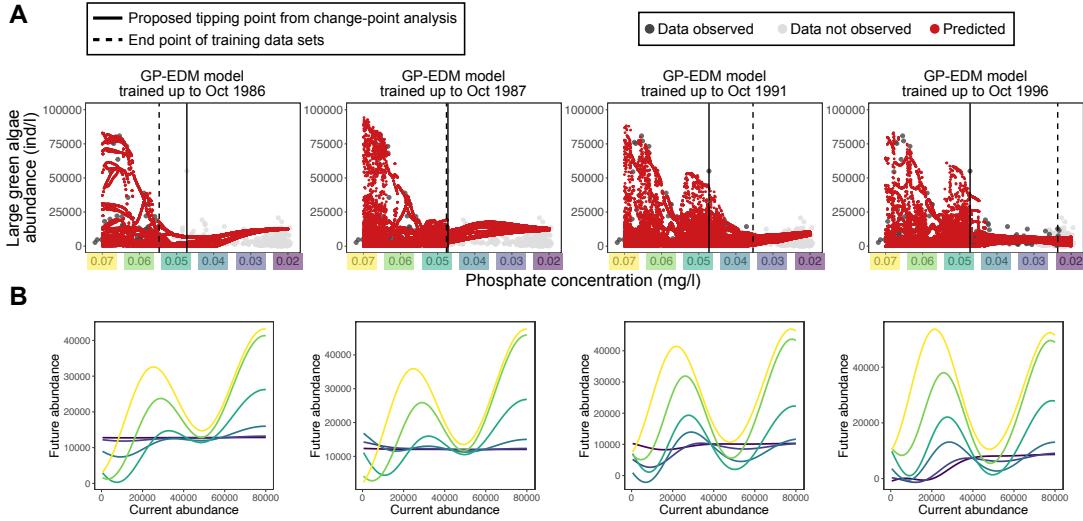


Fig. S19. Relationship between future ($x(t + 1)$) and current ($x(t)$) abundance inferred from the GP-EDM model for different levels of phosphate concentration (different colors) for large green algae in Lake Zurich. **A**, Predicted bifurcation diagrams (in red) using the GP-EDM model trained up to each of the four dashed lines. Bifurcation diagrams depict the population abundance after the transient period (y-axis) for a range of phosphate concentration values (x-axis). The training data end points are shown as dates above each plot. Gray points denote population abundance values used to train the GP-EDM model (dark gray) or not yet observed (light gray). These plots are exactly the same as the plots in Fig. 5 in the main text. **B**, Each plot shows the inferred relationship between $x(t + 1)$ and $x(t)$ using the training data set specified in the plots in **A**. Note that, as in Fig. 1 in the main text, other model inputs are also important for function $x(t + 1) = G[x(t), x(t - 1), x(t - 2), x(t - 3), w(t), w(t - 1), p(t)]$, but are held constant here for simplicity. These plots suggest that the relationship between $x(t + 1)$ and $x(t)$ becomes more nonlinear at higher phosphate concentrations, resulting in chaotic regimes with large-amplitude abundance fluctuations.

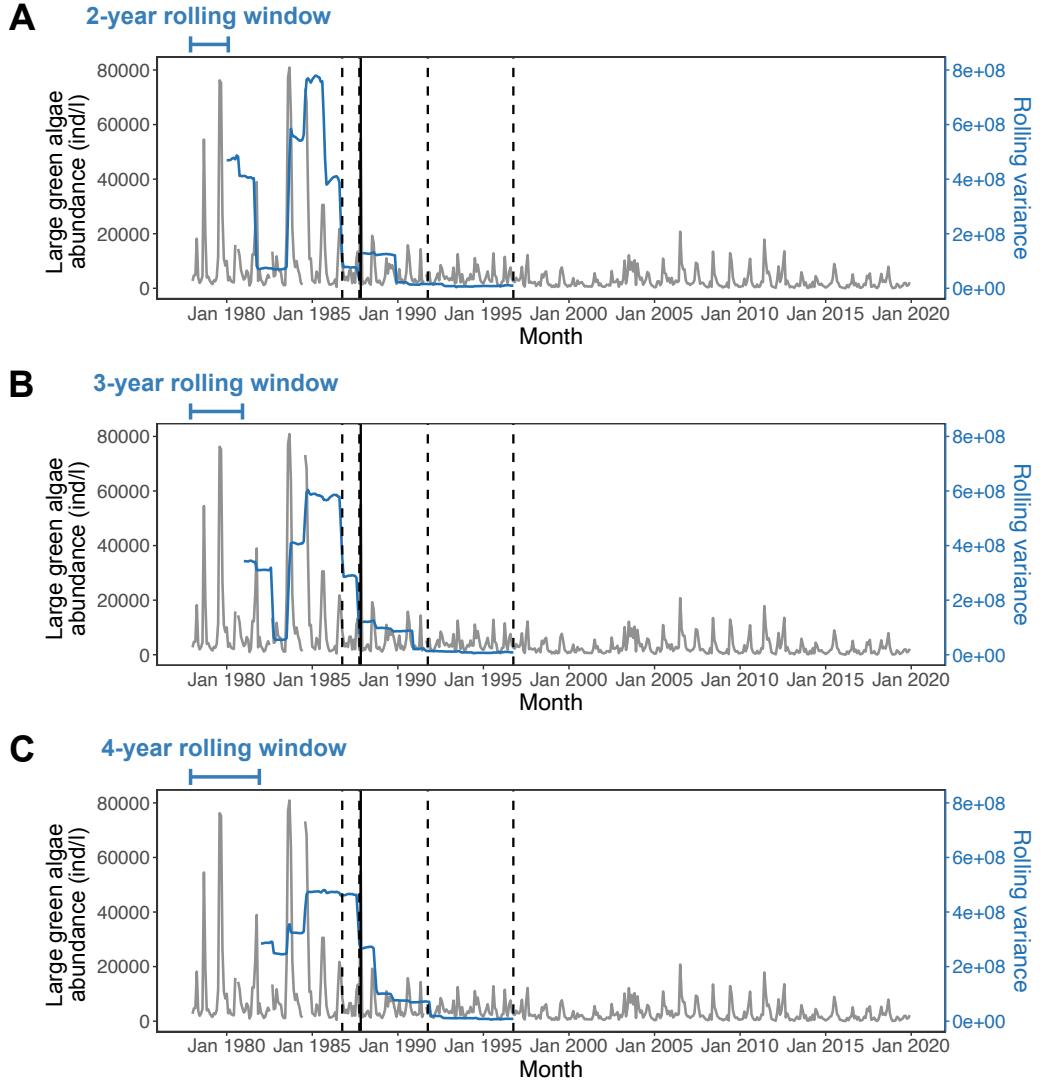


Fig. S20. Monthly population time series (in gray) of large green algae and its rolling variance (in blue) in Lake Zurich from January, 1978 to December 2019. The population shows a tipping point on November, 1987 (vertical solid line) according to a change-point analysis. The vertical dashed lines show the end points of the training data sets used with the GP-EDM model. The rolling variance was computed using three different window sizes: 2 years (**A**), 3 years (**B**), and 4 years (**C**).

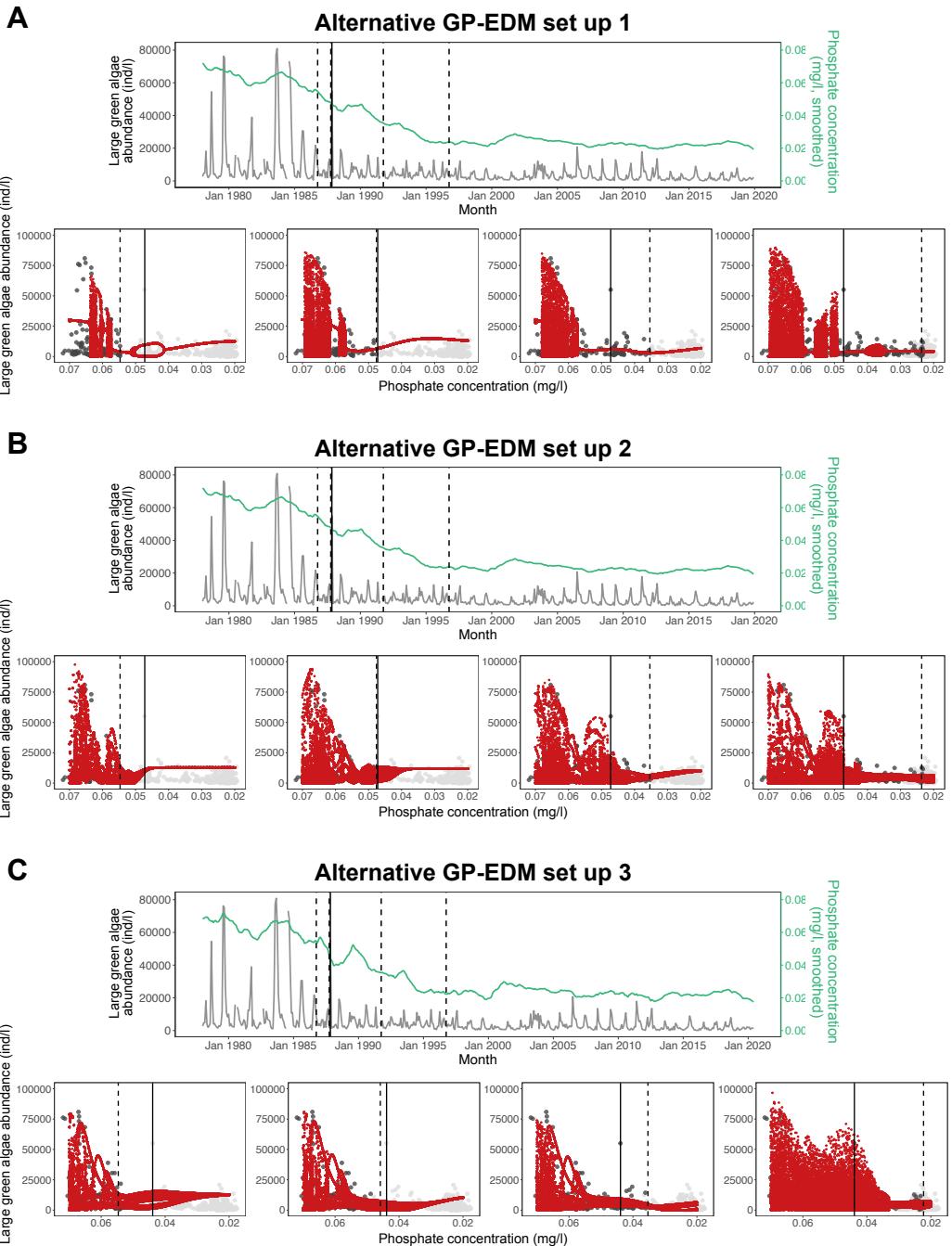


Fig. S21. Reconstructed bifurcation diagrams of large green algae in Lake Zurich using different GP-EDM set ups. Results are qualitatively the same as in Fig. 5 in the main text. **A**, Alternative set up 1 is exactly the same as our main text set up, but we do not include temperature as a model input. **B**, Alternative set up 2 is exactly the same as our main text set up, but we fix the inverse-length scale parameter (ϕ_j) of phosphate to 1. **C**, Alternative set up 3 is exactly the same as our main text set up, but we use a smaller window to smooth the phosphate time series (window with 12 instead of 24 months). In **A-C**, the abundance of large green algae is shown in gray and phosphate concentration is shown in green. The potential tipping point is shown as a vertical solid line and vertical dashed lines show the end points of the training data sets. In the bifurcation diagram plots, red points represent our predictions and gray points denote abundance values used to train the GP-EDM model (dark gray) or not yet observed (light gray).

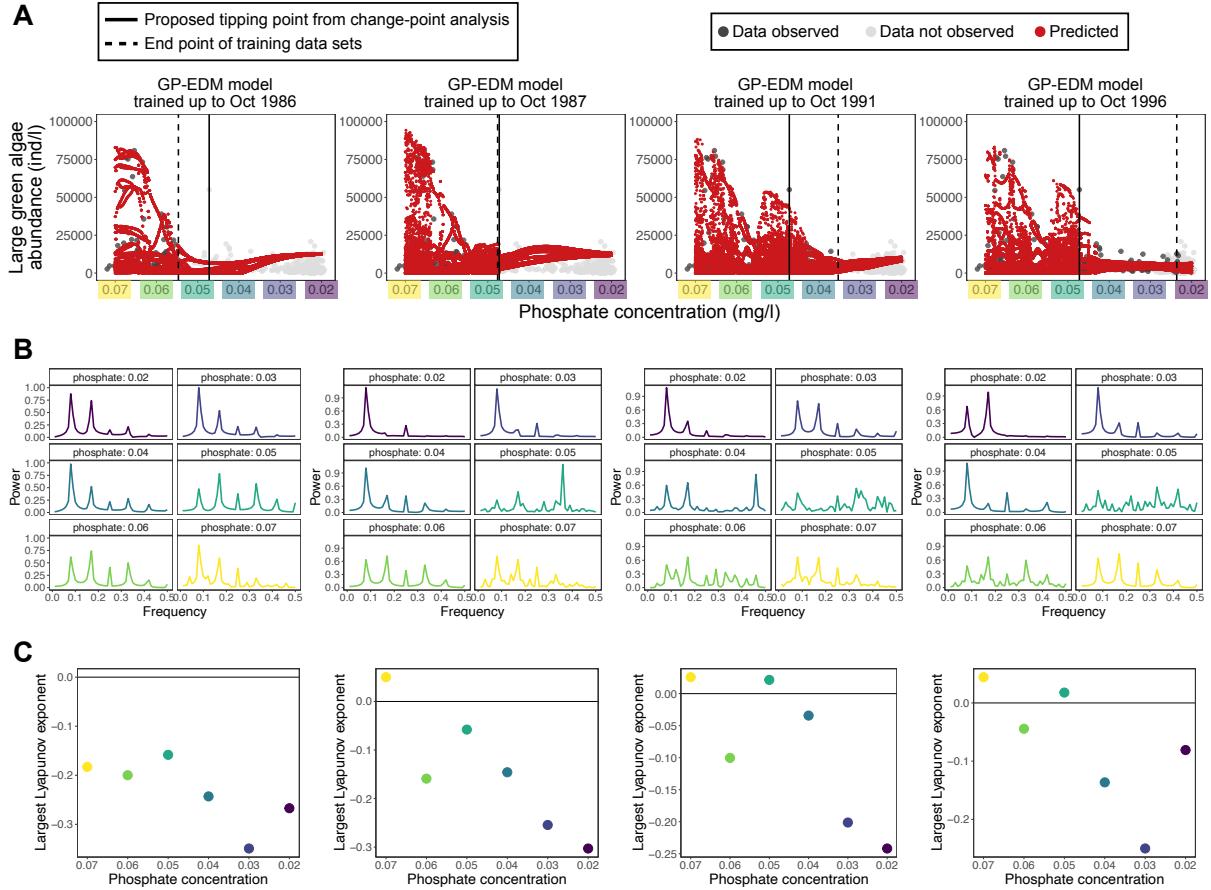


Fig. S22. Power spectrum and largest Lyapunov exponent computed for the predicted dynamics of large green algae in Lake Zurich under 6 different values of phosphate concentration (*SI Appendix Section S4*). **A**, Predicted bifurcation diagrams (in red) using the GP-EDM model trained up to each of the four dashed lines. Bifurcation diagrams depict the population abundance after the transient period (y-axis) for a range of phosphate concentration values (x-axis). The training data end points are shown as dates above each plot. Gray points denote population abundance values used to train the GP-EDM model (dark gray) or not yet observed (light gray). These plots are exactly the same as the plots in Fig. 5 in the main text. **B**, Each panel shows the power spectrum of the predicted time series under a given value of phosphate concentration. The power spectrum suggest periodic dynamics for low phosphate concentration and aperiodic dynamics (i.e., chaos) for high phosphate concentration. **C**, Each point shows the value of the largest Lyapunov exponent computed under each value of phosphate concentration. Chaotic dynamics (i.e., largest Lyapunov exponent greater than zero) appears more likely at high values of phosphate concentration. In all plots, colors denote different values of phosphate concentration.

Table S1. Results from the change-point analysis for all plankton functional groups from Lake Zurich (see details in the *Materials and Methods* section in the main text). TP time refers to the time point t for the potential tipping point. TP month refers to the month of the potential tipping point. SS_{ratio} refers to the ratio between $SS_1 + SS_2$ and SS_{all} , where SS_i is the sum of squares of the i th time series window and SS_{all} is the sum of squares of the entire time series. Phosphate refers to the phosphate concentration (in mg/l) at the potential tipping point. We highlight in gray the 3 functional groups with the strongest evidence of a shift in the dynamics, that is, with the lowest values of SS_{ratio} .

Functional group	TP time	TP month	SS_{ratio}	Phosphate
Cryptophytes (large)	218	Feb-96	0.940	0.024
Cryptophytes (small)	130	Oct-88	0.855	0.043
Cyanobacteria	102	Jun-86	0.991	0.055
Diatoms (large)	344	Aug-06	0.976	0.023
Diatoms (small)	222	Jun-96	0.961	0.023
Gold algae (large)	114	Jun-87	0.961	0.049
Gold algae (small)	219	Mar-96	0.989	0.024
Green algae (large)	119	Nov-87	0.853	0.047
Green algae (small)	126	Jun-88	0.996	0.043
Invertebrate predators	142	Oct-89	0.995	0.046
Large herbivores	178	Oct-92	0.980	0.035
Mixotrophic flagellates	219	Mar-96	0.931	0.024
Omnivores	102	Jun-86	0.903	0.055

Table S2. Results from the leave-one-out prediction analyses for all plankton functional groups from Lake Zurich (see details in the *Materials and Methods* section in the main text). For each functional group, we show the details of the selected GP-EDM model, that is, the model with the highest R^2 value. Embedding dimension refers to the total number of lagged inputs in the GP-EDM model (lags of abundance + lags of temperature + lags of phosphate). ϕ refers to the set of inverse length scale parameters (ϕ_j) for all model inputs. R^2 is the leave-one-out prediction accuracy (equation [4] in the main text). We highlight in gray the 3 most predictable functional groups, that is, with the highest values of R^2 .

Functional group	Embedding dimension	ϕ	R^2
Cryptophytes (large)	$6 + 2 + 1 = 9$	0, 0.022, 0, 0, 0, 0.388, 0.402, 0	0.405
Cryptophytes (small)	$6 + 2 + 1 = 9$	0.397, 0, 0, 0.075, 0.122, 0.572, 0, 0.331, 0.04	0.394
Cyanobacteria	$1 + 1 + 1 = 3$	0.405, 0.179, 0	0.696
Diatoms (large)	$4 + 2 + 1 = 7$	0, 0, 0.417, 0, 0.236, 1.158, 0	0.186
Diatoms (small)	$1 + 2 + 1 = 4$	2.392, 1.556, 0, 0.055	0.492
Gold algae (large)	$5 + 1 + 1 = 7$	0.269, 0, 0.18, 0.437, 0, 0.424, 0	0.174
Gold algae (small)	$7 + 1 + 1 = 9$	0.453, 0, 0, 0.068, 0.072, 0.006, 0.609, 0.015	0.271
Green algae (large)	$4 + 2 + 1 = 7$	0.713, 0.164, 0.196, 0.621, 0, 0.352, 0.066	0.77
Green algae (small)	$6 + 2 + 1 = 9$	0.289, 0.077, 0, 0.176, 0, 0.23, 0.165, 1.011, 0.022	0.801
Invertebrate predators	$6 + 2 + 1 = 9$	1.368, 0.005, 0, 0.037, 0, 0, 0.471, 1.494, 0.644	0.325
Large herbivores	$6 + 2 + 1 = 9$	0.015, 0, 0, 0, 0.017, 0.337, 0.111, 0.424	0.526
Mixotrophic flagellates	$6 + 2 + 1 = 9$	0.157, 0, 0, 0, 0.841, 0, 0.222, 0.11, 0	0.46
Omnivores	$4 + 2 + 1 = 7$	0.042, 0.219, 1.173, 1.145, 0, 0.326, 0	0.399