Recent developments in empirical dynamic modelling

Stephan B. Munch, Tanya L. Rogers, George Sugihara

SEDAR87-RD-14

November 2024

Received: 9 May 2022 | Accepted: 22 August 2022
DOI: 10.1111/2041-210X.13983

Methods in Ecology and Evolution

REVIEW

Methods in Ecological Forecasting

Recent developments in empirical dynamic modelling

Stephan B. Munch[1,2] | Tanya L. Rogers[1] | George Sugihara[3]

[1]Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Santa Cruz, California, USA
[2]Department of Applied Mathematics, University of California, Santa Cruz, California, USA
[3]Scripps Institution of Oceanography, University of California San Diego, La Jolla, California, USA

Correspondence
Stephan B. Munch
Email: smunch@ucsc.edu

Funding information
Department of Defense; Department of the Interior, Grant/Award Number: NPS-P20AC00527; Lenfrest Oceans Program; National Oceanic and Atmospheric Administration; National Science Foundation, Grant/Award Number: NSF ABI-1667584 and NSF DEB-1655203

Handling Editor: Sydne Record

Abstract

1. Ecosystems are complex and sparsely observed making inference and prediction challenging.
2. Empirical dynamic modelling (EDM) circumvents the need for a parametric model and complete observations of all system variables. Classical univariate approaches, which require time-series observations of only a single focal variable, can produce verifiable out-of-sample forecasts; however, they can sometimes require long time series that may be difficult to obtain. More importantly, classical approaches limit the depth of mechanistic understanding that can be gained and the generalizability of forecasts to non-analogue futures.
3. We review the main ideas of EDM and more recent extensions that expand their capabilities for improving forecasts and understanding mechanism.
4. Algorithms are now available that allow for missing data, unequal sampling intervals and combining short time series, which increase the number of datasets that can be used. Recent extensions of EDM to multivariate time series substantially expand the range of applications and mechanistic questions that can be addressed, including detecting causal coupling, tracking changing interactions in real time, leveraging short time series from information shared in coupled variables, modelling dynamically changing stability, scenario exploration, and management applications involving optimal control.

KEYWORDS
convergent cross-mapping, early warnings, empirical dynamic modelling, gaussian process, scenario exploration, simplex, S-map, stability

1 | INTRODUCTION

Ecosystems are complex, involving tens to thousands of species that interact with each other and the environment in nonlinear ways and on multiple time-scales ranging from seconds to millennia. As a consequence, it is not surprising that models of ecosystem dynamics are challenged in their ability to make valid predictions. Critically, although we wish to achieve a comprehensive and holistic understanding of ecosystem dynamics, nearly all systems of interest are sparely observed, meaning we typically have data on a relatively small fraction of species' abundance, traits, environmental drivers and other variables that might contribute to ubiquitous fluctuations in ecosystem state. This 'partially observed system' problem is not unique to ecology; incomplete observations impede advances in domains as disparate as genetics, neurobiology, climate science and finance.

There are three classes of approaches for addressing the partially observed system dilemma. The first, and arguably most common approach is simply to ignore it: to use simplified parametric models that focus only on the observed state variables. Examples include the single-species models widely used in wildlife management and conservation (FAO, 2020; IUCN, 2020). This approach sidesteps the problem, essentially treating unobserved variables as process noise. While there are circumstances (e.g. time-scale separation) where the 'process noise approximation' can be made rigorous, doing so requires strict assumptions about the dynamics that are difficult to justify. Most importantly, by ignoring unobserved variables, this expedient approach foregoes potential gains in predictability and can produce an oversimplified conception of the system dynamics.

The second class of approaches is to use a well-vetted model for the system to infer missing state variables and parameters. This tactic requires a system model that is a good approximation of the underlying dynamics (i.e. hidden Markov or state-space models, see e.g. De Valpine, 2002; Holmes et al., 2012; Millar & Meyer, 2000). Although these models often make compelling in-sample predictions, if the approximation is poor, this approach will generate unreliable estimates of missing state variables and parameters, and fall short in terms of out-of-sample prediction (e.g. Deyle, May, et al., 2016; Judd et al., 2008).

In contrast to the prior two, the third class of approaches does not assume a particular functional form, but instead uses time-series data in a nonparametric fashion and implicitly accounts for unobserved state variables. These data-centric methods of 'attractor reconstruction' have appeared in the literature under various monikers including 'nonlinear forecasting', '(time) delay embedding', 'state-space reconstruction', and most recently, 'empirical dynamic modelling' (EDM). Delay embedding was first used in ecology by Schaffer (1984) as a step in constructing a unimodal map, but applications of the method for forecasting and exploration of dimensionality were first introduced to ecology in the 1990s by Sugihara and May (1990).

Since the 1990s, there have been many conceptual and statistical advances in EDM that expand both the range of questions that can be addressed and the datasets that can be analysed. Our review is focused on these methodological developments, primarily in the ecological literature. Before reviewing these, we briefly summarize the ideas behind EDM and its implementation (for more details, see Munch et al. (2020) and Chang et al. (2017)). We close with a discussion of some important areas for future development.

2 | OVERVIEW OF EDM

EDM views time series as 'observation functions' of a dynamical system (a brief video summary can be found here, http://tinyurl.com/EDM-intro). In its simplest form, the dynamical system can be thought of as a homogeneous (i.e. time-invariant) nonlinear system of N state variables, say $\boldsymbol{x}(t) = \left\{x_1(t), \dots, x_N(t)\right\}$, whose dynamics are governed by $\frac{dx}{dt} = f(\boldsymbol{x})$, and the observed time series y is some function of the system state, that is, $y(t) = P[\boldsymbol{x}(t)]$. In ecological applications, an observation function is often thought of as a state variable, say $y(t) = x_1(t)$, but this is not required. In general, observation functions can be any variable that records displacement of the system as it evolves (e.g. total abundance of several age classes or species in a functional group).

Natural systems are typically thought to be dissipative, meaning they collapse from a potentially very high N-dimensional space to a much lower dimensional 'attractor' with dimension d, to which trajectories converge. The attractor can be a point ($d = 0$), a closed loop as in a stable limit cycle ($d = 1$), a torus ($d = 2$) or a more complex shape with a fractal dimension (e.g. a 'strange attractor'). The embedding dimension, E, corresponds to the number of variables (or coordinate axes) required to resolve the attractor. The Whitney embedding theorem states that $E \geq 2d$ is sufficient. Assuming the dynamics are reasonably smooth (most ecological models fall in this category), trajectories that start from nearby points on the attractor tend to remain close together, at least for a period of time. EDM exploits this fact when making forecasts based on the fates of near-by analogues. Say we are starting from a point, $\boldsymbol{x}(t)$, and wish to predict the state h steps into the future, $\boldsymbol{x}(t+h)$. Smoothness implies that we can find a collection of historical points close to $\boldsymbol{x}(t)$, say $\boldsymbol{x}\left(t_1\right), \dots, \boldsymbol{x}\left(t_n\right)$, and use their future states to make a prediction. The simplest way to do so would be to use the average of their future states, $\hat{\boldsymbol{x}}(t+h) = 1/n \sum_{i=1}^{n} \boldsymbol{x}\left(t_j + h\right)$. Repeating this process over a wide range of initial states, $\boldsymbol{x}(t)$, provides multiple estimates of the mapping from $\boldsymbol{x}(t)$ to $\boldsymbol{x}(t+h)$. Interpolating among them results in an empirically derived discrete time model for the system, that is, $\hat{\boldsymbol{x}}(t+h) = \tilde{F}[\boldsymbol{x}(t)]$.

The second, possibly more important, property of dynamical systems exploited in EDM is described in Takens' (1981) delay embedding theorem, which shows how any one system variable can contain information about the others. Takens' theorem shows 'generically' that there is a one-to-one correspondence between the attractor in the original coordinate system, defined by the collection of vectors $\boldsymbol{x}(t) = \left\{x_1(t), \dots, x_N(t)\right\}$, and the collection of 'delay coordinate vectors' defined by $\left\{x_j(t), x_j(t-\tau), \dots x_j(t-E\tau)\right\}$, where τ is the time delay and E is the embedding dimension. This is profoundly important because if any one state variable contains information about the others, lags of that variable can act as a substitute for (and indirectly account for the dynamics of) unobserved state variables. More generally, Takens' says that delay coordinates can be constructed from any observation function of the system, that is, $\boldsymbol{y}(t) = \{y(t), y(t-\tau), \dots, y(t-E\tau)\}$. The original proof for smooth manifolds has been generalized to stochastic and fractal attractors (Stark et al., 2003). The one-to-one correspondence between $\boldsymbol{x}(t)$ and $\boldsymbol{y}(t)$, provides a justification for constructing discrete time models of the form $\hat{y}(t+h) = G[\boldsymbol{y}(t)]$. In most applications of EDM, doing so involves two steps: choosing E and τ, then approximating G from data using some flexible approach. We consider these steps in more detail below.

It is important to note that Takens' theorem was originally restricted to time-invariant, deterministic systems, but was later extended to forced and stochastic systems (Stark, 1999; Stark et al., 2003). In a stochastic system, a rigorous delay-embedding map is given by $x_t = f\left(x_{t-\tau}, \dots, x_{t-E\tau}, \varepsilon_{t-\tau}, \dots, \varepsilon_{t-E\tau}\right)$. However, the approximation $f\left(x_{t-\tau}, \dots, x_{t-E\tau}\right) + \varepsilon_t$ is implicit in most ecological applications. Simulations show that EDM successfully captures the conditional mean for an age-structured stochastic population model (Munch et al., 2020).

2.1 | Selecting embedding parameters

A variety of methods for selecting τ and E have been proposed and are well described elsewhere (Chang et al., 2017; Kantz & Schreiber, 2003; Li et al., 2021). Since the dynamics in the original state space are assumed to be smooth, nearby trajectories should not cross. The objective in selecting τ and E is to find values such that trajectories in delay coordinate space also do not cross. This is referred to as 'unfolding' or 'embedding' the attractor. Intuitively, any point in delay coordinates where trajectories cross is a point where prediction based on the behaviour of nearby points will be poor. This can happen, for instance, if E is too small and trajectories that are not actually nearby in a higher-dimensional space (and head in different directions) appear to be close together in a lower-dimensional space. This is known as the problem of 'false neighbours' (Kennel et al., 1992) and is addressed in forecasting (Sugihara & May, 1990).

Early approaches to delay coordinate embedding (Chan & Tong, 2001; Sauer et al., 1991) suggested selecting τ based on minimizing time-series autocorrelation or mutual information, then choosing E based on nearest neighbour forecasting (Sugihara & May, 1990) or an equivalent false nearest neighbours algorithm (Abarbanel & Kennel, 1993). One straightforward approach is to evaluate prediction accuracy over a grid of E and τ and simply choose the pair that produces the most accurate forecasts τ time steps ahead (Sugihara, 1994; Sugihara & May, 1990), preferentially choosing smaller values in the case of statically indistinguishable results. In most ecological applications to date, τ is usually fixed to 1 since the sampling intervals are often fairly coarse (e.g. monthly or annual surveys), making E the only free parameter, which is noteworthy in terms of simplicity. However, as more high-frequency ecological data become available, estimates of τ will provide meaningful information on the relevant time-scale in the system.

Heuristically, when the time series exhibits some near-periodic behaviour, the product $E\tau$ should be close to the recurrence time (i.e. average time it takes for the system to return to a nearby state, see e.g. Kantz & Schreiber, 2003). However, many systems of interest involve dynamics that play out on disparate time-scales. When multiple time-scales are present, the use of a constant τ can be less desirable, as different values correspond to modelling dynamics on different time-scales of interest (Judd & Mees, 1998). Intuitively, if τ is based on the short time-scale, many coordinates (i.e. large E) might be needed. Conversely, if τ is based solely on the long time-scale, the high-frequency dynamics will be treated as noise. Selecting a single τ based on prediction accuracy assumes embedding relative to the dominant time-scale. Constructing delay vectors with several different values of τ can ameliorate this difficulty (Judd & Mees, 1998) and is justified by the multivariate embedding theorem (Deyle & Sugihara, 2011). Alternatively, the time-scale problem can be circumvented using an ensemble of randomly generated delay vectors (see e.g. Tajima et al., 2015).

2.2 | Fitting the model

Given a collection of delay coordinate vectors, the next step is to approximate the delay coordinate map. Many different function approximation schemes have been applied to EDM, each with strengths and weaknesses, including locally constant models (Simplex; Sugihara & May, 1990), locally linear models (S-map; Sugihara, 1994), Gaussian processes (GPs; Munch et al., 2017), polynomial 'response surface' models (Turchin & Taylor, 1992), neural network models (Nychka et al., 1992) and generalized additive models (Benincà et al., 2015). Bhat and Munch (2022) show that the algebra of delay embedding leads naturally to a representation as a recurrent neural network, opening the way for partially specified models. It is worth noting that, while some methods are clearly too stiff to be broadly applicable (e.g. polynomial models), the 'best' approach (in terms of prediction accuracy) depends on many case-specific factors including the length of the time series, the amount of noise present and the shape of the attractor, which a priori is unknown. We highlight a few features of Simplex, S-map and GP below.

The most widely used models in ecology are the locally constant (zeroth order) 'Simplex' (Sugihara & May, 1990) and the locally linear (first order) 'S-map' (Sugihara, 1994). These function approximations intentionally have a minimum number of free parameters to reduce the possibility of overfitting and aid transparency.

The zeroth-order Simplex is a nearest neighbour forecasting method. A major advantage of Simplex is its extreme simplicity: to make a prediction from given an initial state, Simplex uses a weighted average of the future values of the focal state's nearest neighbours. Since the minimum number of points needed to surround the focal state in E dimensions is $E+1$, Simplex uses $E+1$ nearest neighbours, eliminating this as a tunable parameter. Indeed, given that τ is typically set to the sampling interval in ecological applications ($\tau = 1$), in practice E is often the only parameter that is fit to data. As such, Simplex provides a simple and computationally efficient first pass in data exploration to test for the presence of determinism using prediction skill for validation.

The first-order S-map is a locally linear least-squares algorithm that is applied sequentially to each point on the attractor with weights that decay exponentially with Euclidean distance from the focal point in delay coordinate space (Sugihara, 1994). Thus, S-map contains only one additional free parameter, θ, which controls how fast the weights decay. Specifically, for any pair of delay vectors at times t and s, the weight is $w_{t,s} = \exp\left[-\theta \Sigma_i \left(y_{t-i} - y_{s-i}\right)^2\right]$. Note that this is defined over all pairs, not just the $E+1$ nearest neighbours. When $\theta = 0$, the weights are constant across the whole time series (all points have equal weight) so the resulting model is equivalent to an autoregressive (AR) model of order E. Critically, the fact that the global AR model is a special case of S-map enables a clear test of the hypothesis that the dynamics are nonlinear (Sugihara, 1994). Applications of this test find evidence that nonlinear dynamics are ubiquitous in nature (Anderson et al., 2008; Clark & Luis, 2020; Glaser, Fogarty, et al., Glaser, Fogarty, et al., 2014; Hsieh et al., 2005; Klein et al., 2016; Sugihara et al., 1999). To encourage finding a parsimonious model, 'regularized' versions of S-map have been developed, though the ideal regularization scheme depends on the modelling objective (Cenci et al., 2019).

To relax the rigid minimalism of the zeroth- and first-order function approximations, Munch et al. (2017) introduced a Bayesian approach to EDM based on GP regression (GP-EDM). GP regression provides a probabilistic, Bayesian framework for EDM that readily incorporates prior information and extends to hierarchical modelling (Munch et al., 2017). In GP-EDM, the covariance function (and its associated inverse length scales) controls the degree of nonlinearity that is analogous to the weighting kernel in S-map. Typical applications of GP-EDM set $\mathrm{Cov}\left[G(\boldsymbol{y_t}), G(\boldsymbol{y_s})\right] = \tau \exp\left[-\Sigma_{i=1}^{E} \phi_i \left(y_{t-i} - y_{s-i}\right)^2\right]$ though other forms are certainly possible. It is also possible to define the GP using Euclidean distance with a single inverse length scale parameter (e.g. all $\phi_i = \overline{\phi}$), analogous to use of the single θ parameter in S-map which gives equal weighting in all directions. However, greater flexibility is obtained by allowing each input to have its own length scale which is particularly useful for optimizing predictability with multivariate embedding.

Although the increased flexibility of GP-EDM increases the possibility of overfitting, this can be substantially minimized through Bayesian prior specification. In contrast to frequentist 'wiggliness penalties' which generally involve an unknown Lagrange multiplier that must be determined out of sample, the GP-EDM uses 'automatic relevance determination' (ARD) priors (Neal, 1996) to control the wiggliness of the estimated function. Specifically, Munch et al. (2017) set the ARD prior such that—in the absence of data—the modal inverse length scale is zero (i.e. the most likely model is flat), and the mean inverse length scale is set such that, on average, the model will have one local maximum over the range of the data. Thus, the prior shrinks the inverse length scales towards 0 for irrelevant inputs, effectively removing them from the model. Using ARD results in parsimonious collections of inputs. In addition, using input-specific length scales simplifies extending EDM to multiple types of inputs where Euclidean distance is not optimal. This is particularly useful in allowing for unequal lag spacing (Munch et al., 2017), such as when the dynamics occur on multiple time-scales, for example, $\hat{y}_{t+1} = G\left[y_t, y_{t-3}, y_{t-12}\right]$ for monthly data from a system with strong seasonal and annual dynamics.

2.3 | Evaluating the model

Regardless of the function approximation scheme chosen, some measure of model quality/validity is needed. The most widely used metrics are related to prediction accuracy, such as the Pearson correlation between predicted and observed values, the mean squared prediction error, mean absolute prediction error, or when appropriate, the percentage correct sign (e.g. positive/negative growth rate). Because these models are almost entirely data driven, predictions must be evaluated 'out of sample'—an idea introduced to ecology in the 1990s (Sugihara & May, 1990). For small datasets and where large deviations in the data (steep peaks and deep valleys) are an important focus, mean squared error in leave-one-out forecasts is commonly applied. When longer time series are available, leaving out more of the data or using explicit training and testing datasets (e.g. Sugihara & May, 1990) is more robust. When we are interested in using EDM to make predictions in real time, it may be more appropriate to use sequential updating where the forecast for each year depends only on data from earlier in the time series. In this case, predictions will improve as the training set (library) increases with time (Giron-Nava et al., 2017; Johnson et al., 2021; Munch et al., 2017), that is, as the attractor becomes denser and nearest neighbour analogues become more similar—a phenomenon known as convergence (Sugihara et al., 2012).

Evaluating prediction accuracy over a range of time-scales can help differentiate linear from nonlinear dynamics, as scale dependence is an important feature of nonlinearity (Sugihara et al., 1999). When sampling intervals are short relative to the system dynamics, the series will be highly autocorrelated. Consequently, for one-step ahead prediction, the constant predictor will do well, and the 'best' model will be a linear AR model of low order (S-map with $\theta = 0$). With unstable, nonlinear dynamics, prediction accuracy will decay exponentially as predictions are made further into the future, and in some cases nonlinear models with higher embedding dimensions can emerge. Along these lines, 'trajectory matching', that is, minimizing the distance between observed and predicted segments, rather than single points, improves estimation for nonlinear models (Hooker & Ellner, 2015; Shertzer et al., 2002), and Judd et al. (2008) have shown that training models using multi-step predictions can substantially improve out of sample forecast accuracy.

3 | EXTENSIONS OF EDM

Many extensions to univariate EDM have been developed that address obstacles to practical application in ecology, such as short time-series and missing data, and that allow for greater applicability and inference when using multivariate data. Here we discuss several extensions that open up a wider range of possibilities.

3.1 | Short time series: Leveraging replicates

Many ecological time series are short relative to the time-scale of the system, which presents a major obstacle to the successful application of EDM: if the time series does not adequately cover the range of possible dynamics and make a sufficient number of 'cycles' around the attractor, it can be difficult for EDM to make reliable predictions. Empirical, results for fish suggest that to obtain appreciable gains in forecast performance, EDM requires time series spanning 10 times the maturation age of the focal organism (Munch et al., 2018), though shorter time series may still provide useful predictions (e.g. Giron-Nava et al., 2017). The maximum estimable embedding dimension is limited by time-series length, and the use of lags eliminates the first $E\tau$ data points as training data, which can be a substantial fraction of short time series. Hence, approaches for making better use of short series are critical for ecological applications.

Concatenating delay matrices for time series that share similar dynamics is one solution. This includes combining information across species with similar dynamics or across spatially replicated series for a single species. Hsieh et al. (2008) proposed a scheme for doing this designed to limit false-positive assessments of nonlinearity—concatenating time series for species with similar dynamics such that the fitting set (library) and test set (predicted) are maximally different from each other, but internally similar. Here dynamic similarity was determined by how well each time series could predict another, later referred to as 'co-prediction' (Kuriyama et al., 2020; Liu et al., 2012).

Often, replicate time series are collected for multiple populations over space. Here, the simplest approach is to assume that all spatial replicates have identical dynamics so that delay matrices from each spatial replicate can be combined (Glaser, Ye, & Sugihara, 2014). Another possibility for leveraging spatially replicated data is to include lags of neighbouring sites as additional predictors (Johnson et al., 2021). Simulations indicate that 'mixed-lag' embeddings outperform concatenation when there is substantial spatial variation in dynamics, but that concatenation is superior when the dynamics are spatially uniform (Johnson et al., 2021). The best approach is likely to vary with the study system, but all can improve performance with short time series.

Another option is to model the dynamics for multiple populations in a common hierarchical model, as implemented in GP-EDM (Munch et al., 2017). This allows for information to be shared across populations, but does not require the dynamics to be identical. To account for differences, the hierarchical model has one additional parameter, referred to as the 'dynamic correlation' (Rogers & Munch, 2020), which estimates the linear similarity between pairs of (nonlinear) delay maps. Specifically, for two time series x_t and y_t, with dynamics $x_{t+1} = f(x_t)$ and $y_{t+1} = g(y_t)$, the dynamic correlation measures $corr[f(s), g(s)]$ over states s, rather than $corr[x_t, y_t]$ over times t. Thus, the dynamic correlation provides information on the similarity of the dynamics irrespective of whether populations are correlated through time. For example, populations with the same underlying dynamics (i.e. $f \approx g$), but that are out of phase, will have high dynamic correlation, but low temporal correlation, particularly when the dynamics are chaotic. The dynamic correlation can reveal hidden spatial structure in population dynamics and dynamic similarities among populations that are temporally asynchronous. For instance, Rogers and Munch et al. (2020) used the dynamic correlation to rule out spatial differences in dynamics as a source of asynchrony in a crab metapopulation.

Combining data across similar species or spatial replicates can be leveraged to obtain better global predictions than would be obtainable using just a single short time series. This requires that the delay embedding maps are similar (usually a reasonable assumption for populations of the same species) but not strongly synchronized (i.e. each series provides some independent information). Combining series in this way could facilitate prediction of extreme events like population crashes for species or locations whose time series do not contain a crash, provided that crashes are found in series with similar dynamics.

3.2 | Missing data, variable step sizes

Most ecological datasets have at least a few missing observations, and many long-term sampling programmes have some variability in the sampling intervals due to weather, equipment failures, funding lapses and so on. The recent global pandemic, for instance, interrupted sampling in many long-term monitoring programmes (Viglione, 2020).

In a large dataset with only a handful of missing observations, ignoring delay vectors containing missing values is of little consequence (Johnson & Munch, 2022) and experience indicates that small variations in sampling interval do not cause serious problems (McGowan et al., 2017). However, when missing observations are more common or steps sizes more variable, some alternative is needed. This is particularly relevant in short series since every missing value results in E missing delay vectors. One obvious solution is to interpolate, either to obtain uniform sampling or fill in data gaps. However, interpolating can create artefacts and is best when restricted to small changes in smoothly varying data (Johnson & Munch, 2022).

Variable step size EDM (VS-EDM; Johnson & Munch, 2022) circumvents the problems of missing data and variable step sizes by expanding the delay vector to include the sampling interval. To provide some intuition for this, recall that the solution for an ordinary differential equation (ODE), $\frac{dx}{dt} = f(x)$, can be written as $x_{t+h} = F(x_t, h)$. Applying this to delay embedding, VS-EDM fits $y(t_i) = F\left[y(t_{i-1}), t_i - t_{i-1}, \ldots, y(t_{i-E}), t_{i-E+1} - t_{i-E}\right]$. Although this expansion of the delay coordinates can be implemented with any function approximation scheme (Simplex, S-Map, etc.), this solution doubles the dimension of the input space, so some regularization is warranted. Using GP-EDM with ARD priors, Johnson and Munch (2022) showed that VS-EDM substantially outperforms standard interpolation and dropping missing data, particularly when time series are relatively short. Importantly, when the step sizes are fixed, ARD eliminates dependence on the time interval (Johnson & Munch, 2022). Multi-step-ahead forecasts can be generated directly from the resulting model by varying the sampling interval.

3.3 | Multivariate embeddings

Univariate delay embedding based on Takens theorem can provide basic information on predictability, nonlinearity and determinism, but the resulting delay embedding map is difficult to interpret mechanistically. However, if data on other state variables or relevant covariates are available, multivariate EDM reconstructions that are ecologically relevant can be used directly to explore mechanism. Deyle and Sugihara (2011) opened this path formally by providing generalizations of Takens' theorem to multivariate embeddings that justify mixed lag, multi-variable models where the lags do not have to be fixed or consecutive. Earlier extensions of Takens' theorem to driven systems (e.g. by environmental or stochastic drivers) justify the inclusion of external drivers as additional covariates (Stark, 1999; Stark et al., 1997). Dixon et al. (1999) provided the first ecological example of a mechanistic multivariate embedding using trial and error forecasting with physical drivers to understand spikes in reef fish spawning.

Multivariate models can potentially improve performance and provide more information about the underlying system dynamics and relationships among variables. In particular, they can provide more interpretable information on time-dependent interactions with other state variables (e.g. predators, competitors or abiotic drivers). As with univariate models, lags of multiple variables can be used to compensate for any remaining unobserved deterministic variables, for example, $x_{1,t} = F\left[x_{1,t-1}, x_{1,t-2} \ \dots, x_{2,t-1}, x_{2,t-2}, \ \dots\right]$. Conversely, dependence on lagged values can indicate that relevant state variables are missing. In addition, when a key driver is essentially stochastic, a univariate reconstruction may yield poor predictions. For example, McGowan et al (2016) found that explicitly including stochastic environmental drivers that were determined to be causal (and whose effects propagated with a time lag detectable with the prediction horizon), improved prediction of red tides well beyond that attainable with lags of the target variable, chlA, alone.

Practical implementations of multivariate embeddings using Simplex or S-map typically normalize the values of each time series to have 0-mean and unit variance, effectively giving each coordinate of the embedding equal weight/relevance. This has proven to be effective but can be relaxed in GP-EDM by assigning each coordinate its own length scale, permitting the ARD to determine the relevance of different inputs.

Before we discuss multivariate embeddings as a route to mechanistic understanding of ecological dynamics, we first address some of the more counter-intuitive implications of the multivariate embedding theorems.

3.4 | Information leverage with multiview

Multivariate embeddings can be highly advantageous when time series are short, since many different combinations of variables and their lags can be used to reconstruct the attractor (Deyle & Sugihara, 2011). Since each of these reconstructions provides a different view of the underlying dynamical system, there is enormous potential for leveraging information from multiple data streams. In fact, the number of possible reconstructions grows combinatorially (Ye & Sugihara, 2016): Given l lags for each of n variables, the number of E-dimensional variable combinations is $m = \binom{nl}{E} - \binom{n(l-1)}{E}$. For example, the number of distinct three-dimensional combinations (three-dimensional embeddings) for a system with up to 3 lags of 10 variables is nearly 3000.

Ye and Sugihara (2016) introduced 'multiview' embedding to take advantage of this and showed that multiple short time series can produce very good forecasts even in high-dimensional systems. Some promising recent implementations of this idea for improving forecasts are the random embeddings discussed by Ma et al. (2018) or using state-dependent weightings to improve forecasts (Okuno et al., 2019). This approach leverages the fact that although there are many theoretically equivalent embeddings, they each stretch or shrink the attractor differently in different regions of the state space, meaning that some may produce better forecasts than others in different regions.

Although multiview leverages the multiplicity of possible reconstructions to improve forecasts, many—if not most—are difficult to interpret mechanistically. Moreover, under the multivariate embedding theorem, the idea of identifying a uniquely 'best' model that is *the* mechanistic representation of the system no longer makes sense. As a simple example, consider a two species model in which $x_{t+1} = rx_t\left(1 - x_t\right) - x_t y_t$ and $y_{t+1} = cx_t y_t - my_t$. This system can be rearranged to find that $x_{t+1} = rx_t\left(1 - x_t\right) - x_t\left(cx_{t-1} - m\right)\left[r\left(1 - x_{t-1}\right) - x_t / x_{t-1}\right]$. Since these two representations are algebraically equivalent, estimating maps of the form $x_{t+1} = F_1\left[x_t, x_{t-1}\right]$ and $x_{t+1} = F_2\left[x_t, y_t\right]$ should lead to statistically indistinguishable fits to time series. However, the nonlinearity in the delay embedding map is more severe (containing terms like x_t / x_{t-1} which blow up as $x_{t-1} \rightarrow 0$) compared to the original system which is much smoother, containing only linear and bilinear terms. As a consequence of increased nonlinearity, we expect convergence of the delay map to be slower than the equivalent dynamics in native coordinates. This is particularly important in stochastic systems where the delay map includes the noise sequence implying that causal drivers in native coordinates contain more information than delay coordinates.

As the dimensionality increases and more series are available, the model selection problem becomes more difficult, as there could be potentially thousands of multivariate embeddings with equivalent, or near equivalent, fits. In these cases, we clearly need some other means of identifying mechanistically relevant predictors (the causal drivers of the focal variable—what May (2020) called the 'active variables').

3.5 | Causality

Identifying causal variables and their linkages is central to all of science. The variety of different methods proposed to assess causality from observational (as opposed to experimental) data reflect fundamentally different views of the underlying system (e.g. linear stochastic vs. nonlinear deterministic or some combination), with each approach being more or less suitable for different problems and domains (e.g. see review by Runge et al. (2019)). Approaches that rely on correlation to identify causal links can have difficulty with nonlinear dynamic systems where, contrary to popular belief, causally coupled causal variables can show no long-term correlation with each other, or have correlations that spontaneously switch sign (mirage correlation Sugihara et al., 2012). These phenomena make studying causality in ecological systems more difficult.

Regardless of whether conditional mutual information, prediction accuracy or some other metric (M) is used, most approaches are based on some form Granger causality (Granger, 1986), which involves comparing M(X|Y,Z) and M(X|Z) to determine whether Y causes X. Obviously, this fails if Z contains the same information as Y. But in dynamical systems where Takens' theorem applies, if Z contains lags of X (or other variables) that are sufficient to reconstruct the attractor, why should we need Y at all? As an example, consider the two species model from the previous section and let $Z = \{x_t, x_{t-1}\}$ and $Y = y_t$. All of the relevant information in Y is already contained in Z, so no causal relationship between Y and X would be found. However, in stochastic systems, where the rigorous delay coordinate map must include the noise history (Stark et al., 1997), contemporaneous values of Y contain more information than lagged values of X and we can expect $M(x_{t+1}|x_t, y_t) > M(x_{t+1}|x_t, x_{t-1})$. Indeed, in the original exposition of Granger causality involving co-integration, the prescient disclaimer was made that 'this may not apply to dynamic systems'.

Convergent cross-mapping (CCM; Sugihara et al., 2012) provides a solution to these problems. Since the delay coordinates for each observed variable form a one-to-one projection of the whole attractor, observables from the same system should produce one-to-one projections of each other as well; that is, coordinates contain information about each other. Hence, if Y causes X, we should be able to reconstruct Y using the delay coordinates for X. This procedure, known as cross-mapping, uses the affected target to estimate contemporaneous states of the proposed causal driver. Unidirectional and bidirectional causal effects can be identified: if Y causes X, but X does not cause Y, delays of X will predict Y, but delays of Y will not predict X. That is, the direction of causation is opposite to the direction of cross-mapping (Cummins et al., 2015). If causality is bidirectional, predictability will be present in both directions. This procedure works well in systems with weak to moderate coupling where the variables are not synchronized. When variables are strongly coupled (e.g. they are synchronized), CCM will indicate bidirectional causation regardless of which variable drives the other, a case explicitly excluded from the original exposition (Sugihara et al., 2012). Ye et al. (2015) extend CCM to cover this case by cross-mapping with lags: using delay coordinates of X to predict values of Y before and after contemporaneous states, and noting that effects can never precede causes (though synchronization may still be problematic). Cross-mapping with lags can also help reveal causal effects on mixed timescales (Saberski et al., 2021).

Because causality is transitive, CCM is not able to differentiate direct and indirect causality unless there are detectable time lags in causal effect (Ye et al., 2015). However, if data are available for potential intermediate variables, relative cross-map strength can in principle help distinguish between direct and indirect coupling. In cases where two interacting variables are forced by a common external variable, such as with seasonal forcing, convergence is generally not observed beyond cross-correlation. However when cross-mapping and cross-correlation are strong and convergence is ambiguous, the use of null surrogates based on seasonality can be used to distinguish cross-map skill that goes beyond the shared seasonal signal (Deyle, Maher, et al., 2016; Sugihara et al., 2017). Alternatively, Leng et al. (2020) propose extending the classical notion of partial correlations to CCM, involving 'partial CCM' scores.

As with other EDM approaches, we can concatenate delay matrices from spatially replicated series (e.g. experimental plots) to obtain more robust CCM results from short time series (Glaser, Fogarty, et al., 2014). Multispatial CCM involves drawing bootstrapped samples from the pool of all spatially replicated observations and uses samples weighted by their dynamic similarity to estimate expected dynamics (Clark et al., 2015).

3.6 | S-map coefficients and interaction strength

Given a collection of causally interacting variables (determined via CCM or through biological observation), it may be possible to infer how their interaction strength varies with the state of the system. The coefficients of the S-map model have been shown to provide an interpretable measure of state-dependent interactions (Deyle, May, et al., 2016). Assume that the dynamics in the native coordinate space are given by $\mathbf{x}_{t+1} = F[\mathbf{x}_t]$ where $\mathbf{x}_t = \{x_{1,t}, \ldots, x_{n,t}\}$. Then in the neighbourhood of a focal point $\mathbf{x}^*$ on the attractor, the dynamics may be approximated as $\mathbf{x}_{t+1} = \mathbf{F}[\mathbf{x}^*] + \mathbf{J}^*(\mathbf{x}_t - \mathbf{x}^*)$ where the Jacobian matrix $\mathbf{J}^*$ is given by $J_{i,j} = \partial F_i / \partial x_j$ evaluated at $\mathbf{x}^*$. Since each $J_{i,j}$ indicates how much $x_{i,t+1}$ will change for a unit change in $x_{j,t}$, it can be interpreted as a measure of interaction strength. Importantly, the regression coefficients obtained by fitting a set of local linear models provide estimates of the $J_{i,j}$'s and can be used to empirically estimate interaction strengths and how they vary with the state of the system (Deyle, May, et al., 2016). Methods for quantifying uncertainty in interaction strength parameters have been proposed using ensembles of regularized S-map models (Cenci & Saavedra, 2018).

It is worth considering the parallel between the Jacobian interaction coefficients mentioned here and the coefficients of the classical community matrix. Here, we define the community matrix as the Jacobian that arises from a first-order Taylor approximation of the full system, evaluated at equilibrium (May, 1972), that is, $J_{i,j} = \partial F_i / \partial x_j$ evaluated at $\mathbf{x}^*$. The result is a matrix of constant interaction coefficients. Ives et al. (2003) proposed using vector autoregression (VAR/MAR) to estimate a per capita community matrix assuming linear dynamics around a single equilibrium point. The S-map Jacobian (Sugihara, 1994), however, is computed sequentially at each point as the system travels along its attractor and provides information on how interactions change with system state, that is, $J_{i,j}(t) = \partial F_i / \partial x_j$ evaluated at $\mathbf{x}_t$. Thus, in contrast to the equilibrium case, S-map coefficients depend on the global model structure realized in the attractor (Song & Saavedra, 2021). The coefficients represent net interactions integrated over the course of a time step so that with short time steps they will closely match expected instantaneous interactions (Deyle, May, et al., 2016), noting that coefficients estimated on larger time steps can encompass indirect effects and unobserved variables in addition to direct effects (Rogers et al., 2020).

The idea of estimating sequential Jacobians can appear superficially similar to dynamic linear models (DLM; e.g. Lamon III et al., 1998) in that both are derived from local linear approximations. However, S-map and DLM make very different assumptions. In a DLM, parameters vary such that points nearby in time have similar parameters. In S-map, parameters vary such that points 'nearby on the attractor' (similar states which may be far apart in time) have similar parameters. Thus, with linear dynamics or with extreme oversampling where local coefficients change slowly relative to the sampling rate, DLM can give results similar to S-map. However, with nonlinear dynamics where states change across successive samples, S-map provides vastly better estimates of the Jacobian coefficients than DLM (fig. 2 in Deyle, May, et al., 2016).

It is important to keep in mind that many combinations of state variables and lags are theoretically equivalent embeddings, and the sequence of values obtained for S-map coefficients depends on the embedding used. So, although the interpretation of S-map coefficients as interaction strengths is sound, it is conditional on the embedding chosen. In light of this, we suggest evaluating the sensitivity of any specific interaction coefficient to the choice of embedding. This can be done exhaustively, through a random projection procedure (Ma et al., 2018), or some other ensemble method (e.g. Cenci & Saavedra, 2018). Consistent results across a range of embeddings point towards the robustness of the interpretation. Another approach suggested by Chang et al. (2021) is the 'multiview distance regularized S-map', which combines CCM, multiview embedding and regularized S-map to provide better estimates of interaction strengths in high-dimensional systems. This procedure has two steps: (1) using CCM to identify a set of causally coupled variables, potentially reducing the number of inputs to S-map and (2) estimating S-map coefficients with weights determined by the multiview distance metric rather than standard Euclidean distance. The authors also employed the elastic net (Cenci et al., 2019) to shrink the S-map coefficients towards 0. This approach produced striking improvements in estimating coefficients for large systems (Chang et al., 2021).

3.7 | Stability and early warning signals

Ecologists have long used the Jacobian matrix to evaluate stability of fixed points in theoretical studies (Allesina & Tang, 2015; May, 1974). Classical stability analysis uses eigenvalues to characterize the long-run growth of a perturbation (May, 1974), while singular values (e.g. 'reactivity'; Caswell & Neubert, 2005) describe the initial growth of a perturbation. Similarly, if $\mathbf{x}_{t+1} = F\left[\mathbf{x}_t\right]$, then a small change in state, say from $\mathbf{x}_t$ to $\mathbf{x}_t + \Delta_t$ will be propagated to the next step as $\Delta_{t+1} = \mathbf{J}_t \Delta_t$ where $\mathbf{J}_t$ is the Jacobian evaluated at $\mathbf{x}_t$. Similarly, structural stability, say $\mathbf{s}_t$, that is, sensitivity of the future state to a small change in a parameter q, is also driven by the Jacobian. That is, if $\mathbf{s}_t = \partial \mathbf{x}_t / \partial q$, then $\mathbf{s}_{t+1} = \mathbf{J}_t \mathbf{s}_t + \partial F\left(\mathbf{x}_t\right) / \partial q$.

The S-map coefficients (from either a univariate or multivariate embedding) provide estimates of the Jacobian elements at each point in the time series. In light of this, it should be possible to characterize state-dependent variation in stability using EDM. Specifically, given a sequence of Jacobians, a corresponding sequence of eigenvalues (e.g. Ushio et al., 2018), singular values, trace (Cenci & Saavedra, 2019) or other metrics can be calculated to provide a state-dependent characterization of stability. These estimates describe local contraction or divergence of trajectories in state space and can provide information on local predictability (Guégan & Leroux, 2009) and local susceptibility to perturbations (Abarbanel et al., 1992).

An interesting implementation of this idea is provided by Rypdal and Sugihara (2019), who modelled dengue fever outbreaks. The attractor was found to collapse seasonally to a fixed point during inter-outbreak periods, and was modelled in two parts (an outbreak period with $E = 9$ and inter-outbreak period with $E = 3$). A local eigenvalue could identify the initiation of an ensuing outbreak and estimate its magnitude, thus providing an early warning signal. The stability of the inter-outbreak period was shown to be a proxy for the size of the otherwise unmeasurable susceptible population. Subsequent models using this susceptibility measure and incorporating CCM-derived climatic drivers (Nova et al., 2021) have surpassed the best models in the Dengue Forecasting Challenge.

Local Lyapunov exponents, which evaluate stability within finite time segments (Abarbanel et al., 1992; Benincà et al., 2015; Ellner & Turchin, 1995), are a natural extension of stability at individual time points. The Jacobian for a segment is the product of step-specific Jacobians. In the long-time limit, the eigenvalues converge to the global Lyapunov exponent, which is an indicator of chaotic dynamics. Lyapunov exponents computed over long but finite empirical time series have been called 'effective Lyapunov exponents' (Grassberger et al., 1988), and give an indication of stable/unstable dynamics over the period of observation. Significantly, Rogers et al. (2022) found that effective Lyapunov exponents computed from S-map Jacobians can identify chaotic dynamics in simulated data from a variety of models, and when applied to ecological time series from the Global Population Dynamics Database chaotic dynamics were found to be common.

Local Lyapunov exponents are straightforward to calculate, whereas precise estimates of global Lyapunov exponents are challenging to obtain, particularly for systems with strong intermittency and noise. We note however that unlike global Lyapunov exponents, which are typically independent of the initial state and are 'metric invariant' (i.e. they are unchanged by a change in coordinates), the values of local Lyapunov exponents depend, often heavily, on the initial state, direction of the perturbation and coordinate system in which they are measured (see e.g. Chang et al., 2021). Thus, when using lags as surrogate coordinates, care should be taken in the interpretation of local Lyapunov exponents. As with the interpretation of interaction strength, we suggest evaluating the robustness of the exponent to variation in choice of embedding.

Cenci and Saavedra (2019) and Cenci et al. (2020) note that the sensitivity of estimated S-map coefficients to noise leads to negative bias in the estimation of the dominant eigenvalue, analogous to earlier results on estimating Lyapunov exponents (Kendall, 2001). They show that the volume contraction rate (VCR), given by the trace of the Jacobian in continuous time (or the log of the determinant over a discrete time step), is more robust to noise for a range of simulation models and propose the VCR as a data-driven index of structural stability.

3.8 | Scenario exploration

Because EDM produces an empirically driven discrete time model for the system, the estimated map can be used as we might any other parameterized model. For instance, it can be used to make predictions, identify equilibria, evaluate parameter sensitivity, evaluate various 'what-if' scenarios or test and optimize policies. For instance, when we build EDM models that include lags of an environmental driver and abundance, for example, $x_t = f\left(x_{t-1}, T_{t-1}\right)$, we can determine the functional dependence on x (or T) by perturbing the time series and evaluating the response (Deyle et al., 2013, 2022; Deyle, Maher, et al., 2016; McGowan et al., 2017). Doing so more generally with multiple drivers, for example, $x_t = f\left(x_{t-1}, T_{t-1}, H_{t-1}\right)$ can elucidate mechanisms that explain seemingly contradictory effects when viewing drivers one at a time; for example, the effect of absolute humidity on global flu incidence flips sign at a threshold temperature (Deyle, Maher, et al., 2016). In fisheries, scenario exploration could be used to advise ecosystem-based management with EDM 'experiments' that account for changing climatic drivers and different harvest policies. Taking this a step further, EDMs can be used to explore how spatial autocorrelation changes under different (hypothetical) environmental regimes to test hypotheses about sources of asynchrony (Rogers & Munch, 2020). By numerically evaluating $f(x, T)$ over a grid of x and T (Rogers & Munch, 2020), EDM can also suggest reasonable shapes for parametric approximations.

The starting point for many theoretical studies is to find the steady states and evaluate their stability. Though somewhat at odds with the dynamical philosophy underlying EDM, it is possible to do so, and the result is an equation-free estimate of the steady-state or feasible limit cycles. Specifically, say we have used EDM to extract a model g. $x_t = f\left(x_{t-1}, x_{t-2}\right)$ from the available time series. As with any parametric model, the steady state, x^*, satisfies $x^* = f(x^*, x^*)$. Plausible values of x^* can be obtained numerically (Munch et al., 2017). In cases where there is no steady state within the range of the data, the posterior probability for x^* will be nearly flat. This can be extended to models with external drivers, for example, temperature or harvesting. For example, if catch of species x is C_t and $x_t = f\left(x_{t-1}, C_{t-1}\right)$, the estimated steady state, $x^*(C) = f(x^*, C)$ is a function of C, which can be used to highlight average impacts of harvesting on the focal population (Giron-Nava et al., 2021).

3.9 | Hybrid models for predicting and managing non-analogue futures

In cases where there are acceptable physical models for extrinsic forcing variables, but where the underlying biology is unresolved, it may be possible to construct a hybrid model that provides significantly better predictions and mechanistic insights for environmental management (McGowan et al., 2017). For example, Deyle et al. (2022) applied this idea to understand the apparent irreversibility of eutrophication in Lake Geneva, where deep oxygen levels remain low despite having achieved fully remediated phosphorus levels. They used a physical model (Simstrat; Gaudard et al., 2019; Schwefel et al., 2016) based on climatic inputs (air temperature) to predict lake turnover, and included this as an additional driver to an EDM for the biological component. The resulting hybrid gives substantially better predictions and provides an actionable description of the emergent processes (biogeochemical, ecological, etc.) that drive water quality. For example, the hybrid model warns that an increase in air temperature of 3°C will have the same effect on water quality as eutrophication in the previous century, and that because of nonlinearities, effective management controls can change as the lake state changes so that reducing phosphorus inputs alone may no longer be sufficient. Models with demonstrated skill in out-of-sample prediction, that allow us to explore non-stationary, non-analogue futures, address a signature challenge for 21st-century environmental management.

3.10 | Optimal control for conservation and management

Optimal control theory and Markov decision processes are applied in many branches of ecology to derive conservation and management plans. Doing so requires us to divide the inputs to our dynamical model into 'state' and 'control' variables and to define a 'reward,' that is, the target to be optimized. State variables are typically some measure of the population or ecosystem, and the controls are indices measuring human interventions, such as harvest rates, protected areas, nutrient loads, etc. In a conservation context, the reward can be the size of the breeding population, or some measure of extinction risk. In harvesting problems, the reward is often the catch or profit extracted from the system. Given these, and a model for the system dynamics, an optimal policy can be derived using dynamic programming (Mangel & Clark, 1989), Pontryagin's maximum principle (Schaffer, 1983), reinforcement learning or other tools (e.g. linear programming; Hernández-Hernández et al., 1996).

In light of this, the first step to using EDM to derive an optimal policy is to estimate a delay embedding map that includes both the state x_t and control u_t from the available time series, that is, $x_t = f\left(x_{t-1}, u_{t-1}\right)$. The second step is to characterize the cumulative reward (or penalty) that is to be maximized (or minimized), say $R\left(x_t, u_t\right)$. An optimal policy is a recipe for determining u_t given x_t that maximizes $V = \sum_{t=0}^{T} \gamma^t R\left(x_t, u_t\right)$ where $\gamma(\leq 1)$ is the discount rate.

Boettiger et al. (2015) introduced the idea of coupling optimal control to a nonparametric model of population dynamics, restricted to one-dimensional state spaces. More recently, Brias and Munch (2021) expanded this framework to constructing policies for larger state spaces as well as several competing objectives. They found that the approximate policies obtained using reinforcement learning (Sutton & Barto, 2018) were similar to those obtained with dynamic programming, but scale much better with the input dimension. In both cases, performance is generally close to optimal and typically much better than the policy obtained using incorrectly specified parametric models.

3.11 | Software extensions for implementing EDM

Until recently, the greatest barrier to implementing EDM other than data constraints has been the accessibility of reliable software. This is quickly changing, and a variety software packages have emerged, most with tutorials, that include many of the tools and recent extensions described here. For example, C++ EDM (Park & Sugihara, 2020) is the computational core behind the popular R package rEDM (Park et al., 2021; Ye et al., 2019) and the even more popular Python package pyEDM (SugiharaLab, 2021), which implement Simplex, S-map, CCM and GP-EDM with a single length scale parameter. Implementations of GP-EDM with separable length scale parameters, ARD priors, hierarchical structures and variable step sizes (VS-EDM) are available in the R package GPEDM (Munch & Rogers, 2022). There are also EDM resources in Stata (Li et al., 2021) and an efficient KEDM algorithm in Kokkos (Takahashi et al., 2021) optimized to run massive factorial CCM calculations on very large databases (100k time series each with 10 k time points). Specialized visualization tools for S-map coefficient exploration are also becoming available (Natsukawa et al., 2021).

4 | CONCLUSIONS AND FUTURE DIRECTIONS

The number of datasets, tools for analysis and range of questions that can be addressed with EDM have expanded dramatically since Takens introduced time-delay embedding in 1981. These range from methods for handling short time-series and missing data, to the causal analysis of multivariate time series, to management applications. We have seen that the multivariate embedding theorems provide a formal justification for expanding EDM to multiple data streams. This results in an explosion of plausible input vectors that can be leveraged to improve prediction using multiview embedding. However, since many combinations are theoretically equivalent, extracting a unique mechanistic representation by finding the best fit becomes meaningless. CCM can help address this by using univariate embeddings to identify causally coupled variables. Given a set of causally coupled variables, we can then apply EDM to address many of the same tasks we typically reserve for parametric models, for example, calculating interaction strengths, computing local Lyapunov exponents, evaluating what-if scenarios, finding equilibria and evaluating their stability, and computing optimal control policies.

Nevertheless, there are still many interesting areas for further method development and applications of EDM. For instance, long-term changes in dynamics, or 'non-stationarity,' are increasingly important in light of global climate change. Although attractor reconstruction applies formally only to stationary systems, there are extensions of EDM that can deal with non-stationarity. At its simplest, EDM can provide an equation-free test for non-stationarity by evaluating prediction accuracy across libraries constructed from different time periods (Schreiber, 1997) or different modes of dynamic behaviour (Lorimer et al., 2021)—in effect, for anomaly detection. When the relevant environmental drivers are known, incorporating these into the embedding may be sufficient for prediction in non-stationary systems (e.g. Deyle et al., 2013), and longer-term ensemble forecasts may be possible by combining physical models with EDM (Deyle et al., 2022). When the relevant drivers are unknown, slow changes can be accommodated by overembedding (Hegger et al., 2000). When the unknown drivers change more rapidly, forgetting the past may be more appropriate (Munch et al., 2017). Ecological applications of EDM to non-stationary systems (or explicitly test for non-stationarity in systems) are rare so far, but are clearly an important avenue for future work.

Over the past decade, several methods for anticipating regime shifts have been developed, most based on the idea of critical slowing down (Scheffer et al., 2012). Although these tools provide robust indicators of impending bifurcations in a range of scenarios, they do not, as yet, provide information on what to expect following a critical transition. On the other hand, Säterberg and McCann (2021) have recently shown that applying EDM to the same system in different regimes produced measurably different attractors. These results suggest EDM could be used to expand the early warning signals toolbox and allow us to determine more precisely how dynamics will differ following a critical transition (e.g. Dakos et al., 2017).

One of the great strengths mechanistic models is that auxiliary information on parameters, sub-models, etc. can be readily incorporated into model formulation. Extending the EDM toolbox to make use of mechanistic information (beyond that contained in the available time series) is an important area for future development. When there is a clear separation in state, say between environmental forcing and biological responses, hybrid models can be used to incorporate mechanism (Deyle et al., 2022). Residual delay maps (Sugihara et al., 1999) and the use of mechanistic models as a priors in Bayesian EDM (Thorson et al., 2014) also provide routes to building in mechanism. More generally, the recursive structure for accounting for missing state variables proposed by Bhat and Munch (2022) provides another path to constructing partially specified EDM models. However, much more work is needed both in determining new ways to incorporate static observations and in evaluating the relative value of this information.

Finally, a completely novel application of EDM is in the realm of simulation, as opposed to prediction. For instance, generative manifold network (GMN) algorithms (Pao et al., 2021), originally developed for neuroscience, are hypothesized to be useful for simulating almost any complex dynamical network. Each node of the GMN is an empirical embedding built using a variant of CCM. It uses concepts from manifold learning, dimensionality reduction, and reservoir computing to generate one-step-ahead predictions recursively—each prediction becomes the added last value, that is then used to predict the next value. Using neural activity data, Pao et al. (2021) showed that the GMN is capable of producing novel behaviours not included in the training set but observed in out-of-sample data, suggesting that in this case, GMN can be used to explore emergent properties of the system not explicitly present in the training set. Whether this setup generalizes to ecological data simulation remains to be seen.

In conclusion, we see that EDM, its extensions and future developments can be useful for the exploration and prediction of increasingly available ecological data. By embracing the complexity and high dimensionality of natural ecosystems and making minimal assumptions, EDM offers a valuable alternative to parametric modelling approaches with potential for many management applications that are only just beginning to be explored.

AUTHOR CONTRIBUTIONS

Stephan B. Munch and George Sugihara led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

ACKNOWLEDGEMENTS

This work was supported by funding to SBM from the Lenfest Oceans Program and NOAA's High-Performance Computing Incubator; and to GS from the DoD-Strategic Environmental Research and Development Program 15 RC-2509, NSF DEB-1655203, NSF ABI-1667584, Department of Interior-NPS-P20AC00527, the McQuown Fund and the McQuown Chair in Natural Sciences, University of California, San Diego.

CONFLICT OF INTEREST

The authors declare no conflict of interest.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13983.

DATA AVAILABILITY STATEMENT

There are no data.

ORCID

Stephan B. Munch https://orcid.org/0000-0001-7471-5429
Tanya L. Rogers https://orcid.org/0000-0003-1253-9903

REFERENCES

Abarbanel, H. D. I., Brown, R., & Kennel, M. B. (1992). Local Lyapunov exponents computed from observed data. *Journal of Nonlinear Science*, *2*(3), 343–365.

Abarbanel, H. D. I., & Kennel, M. B. (1993). Local false nearest neighbors and dynamical dimensions from observed chaotic data. *Physical Review E*, *47*(5), 3057–3068. https://doi.org/10.1103/PhysRevE.47.3057

Allesina, S., & Tang, S. (2015). The stability–complexity relationship at age 40: A random matrix perspective. *Population Ecology*, *57*(1), 63–75. https://doi.org/10.1007/s10144-014-0471-0

Anderson, C. N., Hsieh, C. H., Sandin, S. A., Hewitt, R., Hollowed, A., Beddington, J., May, R. M., & Sugihara, G. (2008). Why fishing magnifies fluctuations in fish abundance. *Nature*, *452*, 835–839. https://doi.org/10.1038/nature06851

Benincà, E., Ballantine, B., Ellner, S. P., & Huisman, J. (2015). Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(20), 6389–6394. https://doi.org/10.1073/pnas.1421968112

Bhat, U., & Munch, S. B. (2022). Recurrent neural networks for partially observed dynamical systems. *Physical Review E*, *044205*, 1–9. https://doi.org/10.1103/physreve.105.044205

Boettiger, C., Mangel, M., & Munch, S. (2015). Avoiding tipping points in fisheries management through Gaussian process dynamic programming. *Proceedings of the Royal Society*, *282*, 20141631. https://doi.org/10.1098/rspb.2014.1631

Brias, A., & Munch, S. B. (2021). Ecosystem based multi-species management using empirical dynamic programming. *Ecological Modelling*, *441*, 109423. https://doi.org/10.1016/j.ecolmodel.2020.109423

Caswell, H., & Neubert, M. G. (2005). Reactivity and transient dynamics of discrete-time ecological systems. *Journal of Difference Equations and Applications*, *11*(4–5), 295–310. https://doi.org/10.1080/10236190412331335382

Cenci, S., & Saavedra, S. (2018). Uncertainty quantification of the effects of biotic interactions on community dynamics from nonlinear time-series data. *Journal of the Royal Society Interface*, *15*(147), 20180695. https://doi.org/10.1098/rsif.2018.0695

Cenci, S., & Saavedra, S. (2019). Non-parametric estimation of the structural stability of non-equilibrium community dynamics. *Nature Ecology and Evolution*, *3*(6), 912–918. https://doi.org/10.1038/s41559-019-0879-1

Cenci, S., Medeiros, L. P., Sugihara, G., & Saavedra, S. (2020). Assessing the predictability of nonlinear dynamics under smooth parameter changes. *Journal of the Royal Society Interface*, *17*(162), 20190627.

Cenci, S., Sugihara, G., & Saavedra, S. (2019). Regularized S-map for inference and forecasting with noisy ecological time series. *Methods in Ecology and Evolution*, *10*(5), 650–660. https://doi.org/10.1111/2041-210X.13150

Chan, K.-S., & Tong, H. (2001). *Chaos: A statistical perspective*. Springer Science & Business Media.

Chang, C. W., Miki, T., Ushio, M., Ke, P. J., Lu, H. P., Shiah, F. K., & Hsieh, C. H. (2021). Reconstructing large interaction networks from empirical time series data. *Ecology Letters*, *24*(12), 2763–2774. https://doi.org/10.1111/ele.13897

Chang, C. W., Ushio, M., & Hsieh, C. H. (2017). Empirical dynamic modeling for beginners. *Ecological Research*, *32*(6), 785–796. https://doi.org/10.1007/s11284-017-1469-9

Clark, A. T., Ye, H., Isbell, F., Deyle, E. R., Cowles, J., Tilman, G. D., Sugihara, G., & Inouye, B. D. (2015). Spatial convergent cross mapping to detect causal relationships from short time series. *Ecology*, *96*(5), 1174–1181. https://doi.org/10.1890/14-1479.1.sm

Clark, T. J., & Luis, A. D. (2020). Nonlinear population dynamics are ubiquitous in animals. *Nature Ecology and Evolution*, *4*(1), 75–81. https://doi.org/10.1038/s41559-019-1052-6

Cummins, B., Gedeon, T., & Spendlove, K. (2015). On the efficacy of state space reconstruction methods in determining causality. *SIAM SIAM Journal on Applied Dynamical Systems*, *14*(1), 335–381.

Dakos, V., Glaser, S. M., Hsieh, C. H., & Sugihara, G. (2017). Elevated nonlinearity as an indicator of shifts in the dynamics of populations under stress. *Journal of the Royal Society Interface*, *14*(128), 20160845. https://doi.org/10.1098/rsif.2016.0845

De Valpine, P. (2002). Review of methods for fitting time-series models with process and observation error and likelihood calculations for nonlinear, non-gaussian state-space models. *Bulletin of Marine Science*, *70*(2), 455–471.

Deyle, E. R., Bouffard, D., Frossard, V., Schwefel, R., Melack, J., & Sugihara, G. (2022). A hybrid empirical and parametric approach for managing ecosystem complexity: Water quality in Lake Geneva under nonstationary futures. *Proceedings in the National Academy of Science of the United States of America*, *119*(26), e2102466119. https://doi.org/10.1073/pnas.210246611

Deyle, E. R., Fogarty, M., Hsieh, C., Kaufman, L., Maccall, A. D., & Munch, S. B. (2013). Predicting climate effects on Pacific sardine. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(16), 6430–6435. https://doi.org/10.1073/pnas.1215506110

Deyle, E. R., Maher, M. C., Hernandez, R. D., Basu, S., & Sugihara, G. (2016). Global environmental drivers of influenza. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(46), 13081–13086. https://doi.org/10.1073/pnas.1607747113

Deyle, E. R., May, R. M., Munch, S. B., & Sugihara, G. (2016). Tracking and forecasting ecosystem interactions in real time. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1822), 20152258. https://doi.org/10.1098/rspb.2015.2258

Deyle, E. R., & Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS ONE*, *6*(3), e18295. https://doi.org/10.1371/journal.pone.0018295

Dixon, P. A., Milicich, M. J., & Sugihara, G. (1999). Episodic fluctuations in larval supply. *Science*, *283*(5407), 1528–1530. https://doi.org/10.1126/science.283.5407.1528

Ellner, S. P., & Turchin, P. (1995). Chaos in a Noisy world: New methods and evidence from time-series analysis. *The American Naturalist*, *145*(3), 343–375.

FAO. (2020). *The state of world fisheries and aquaculture 2020*. Sustainability in Action.

Gaudard, A., Råman Vinnå, L., Bärenbold, F., Schmid, M., & Bouffard, D. (2019). Toward an open access to high-frequency lake modeling and statistics data for scientists and practitioners -- the case of swiss lakes using Simstrat v2.1. *Geoscientific Model Development*, *12*(9), 3955–3974. https://doi.org/10.5194/gmd-12-3955-2019

Giron-Nava, A., CC, J., AF, J., Dannecker, D., Kolody, B., Lee, A., Nagarkar, M., GM, P., Ye, H., DG, J., & Sugihara, G. (2017). Quantitative argument for long-term ecological monitoring. *Marine Ecology Progress Series*, *572*, 269–274. https://doi.org/10.3354/meps12149

Giron-Nava, A., Ezcurra, E., Brias, A., Velarde, E., Deyle, E., Cisneros-Montemayor, A. M., Munch, S. B., Sugihara, G., & Aburto-Oropeza, O. (2021). Environmental variability and fishing effects on the Pacific sardine fisheries in the Gulf of California. *Canadian Journal of Fisheries and Aquatic Sciences*, *78*(5), 623–630. https://doi.org/10.1139/cjfas-2020-0010

Glaser, S. M., Fogarty, M. J., Liu, H., Altman, I., Hsieh, C.-H. H., Kaufman, L., MacCall, A. D., Rosenberg, A. A., Ye, H., & Sugihara, G. (2014). Complex dynamics may limit prediction in marine fisheries. *Fish and Fisheries*, *15*(4), 616–633. https://doi.org/10.1111/faf.12037

Glaser, S. M., Ye, H., & Sugihara, G. (2014). A nonlinear, low data requirement model for producing spatially explicit fishery forecasts. *Fisheries Oceanography*, *23*(1), 45–53. https://doi.org/10.1111/fog.12042

Granger, C. W. J. (1986). Developments in the study of cointegrated economic variables. *Oxford Bulletin of Economics and Statistics*, *48*(3), 213–228.

Grassberger, P., Badii, R., & Politi, A. (1988). Scaling laws for invariant measures on hyperbolic and nonhyperbolic atractors. *Journal of Statistical Physics*, *51*(1), 135–178. https://doi.org/10.1007/BF01015324

Guégan, D., & Leroux, J. (2009). Forecasting chaotic systems: The role of local Lyapunov exponents. *Chaos, Solitons and Fractals*, *41*(5), 2401–2404. https://doi.org/10.1016/j.chaos.2008.09.017

Hegger, R., Kantz, H., Matassini, L., & Schreiber, T. (2000). Coping with nonstationarity by Overembedding. *Physical Review Letters*, *84*(18), 4092–4095. https://doi.org/10.1103/PhysRevLett.84.4092

Hernández-Hernández, D., Hernández-Lerma, O., & Taksar, M. (1996). The linear programming approach to deterministic optimal control problems. *Applicationes Mathematicae*, *24*(1), 17–33. https://doi.org/10.4064/am-24-1-17-33

Holmes, E. E., Ward, E. J., & Wills, K. (2012). MARSS: Multivariate autoregressive state-space models for analyzing time-series data. *The R Journal*, *4*(1), 11–19. https://journal.r-project.org/archive/2012-1/RJournal_2012-1_Holmes~et~al.pdf

Hooker, G., & Ellner, S. P. (2015). Goodness of fit in nonlinear dynamics: Misspecified rates or misspecified states? *Annals of Applied Statistics*, *9*(2), 754–776. https://doi.org/10.1214/15-AOAS828

Hsieh, C., Anderson, C., & Sugihara, G. (2008). Extending nonlinear analysis to short ecological time series. *The American Naturalist*, *171*(1), 71–80. https://doi.org/10.1086/524202

Hsieh, C. H., Glaser, S. M., Lucas, A. J., & Sugihara, G. (2005). Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. *Nature*, *435*(7040), 336–340.

IUCN. (2020). *The IUCN Red List of Threatened Species. Version 2020–2*. https://www.iucnredlist.org.

Ives, A. R., Dennis, B., Cottingham, K. L., & Carpenter, S. R. (2003). Estimating community stability and ecological interactions from time-series data. *Ecological Monographs*, *73*(2), 301–330. https://doi.org/10.1890/0012-9615(2003)073[0301:ECSAEI]2.0.CO;2

Johnson, B., Gomez, M., & Munch, S. B. (2021). Leveraging spatial information to forecast nonlinear ecological dynamics. *Methods in Ecology and Evolution*, *12*(2), 266–279. https://doi.org/10.1111/2041-210X.13511

Johnson, B., & Munch, S. B. (2022). An empirical dynamic modeling framework for missing or irregular samples. *Ecological Modelling*, *468*(March), 109948. https://doi.org/10.1016/j.ecolmodel.2022.109948

Judd, K., & Mees, A. (1998). Embedding as a modeling problem. *Physica D: Nonlinear Phenomena*, *120*(3), 273–286. https://doi.org/10.1016/S0167-2789(98)00089-X

Judd, K., Reynolds, C. A., Rosmond, T. E., & Smith, L. A. (2008). The geometry of model error. *Journal of the Atmospheric Sciences*, *65*(6), 1749–1772. https://doi.org/10.1175/2007JAS2327.1

Kantz, H., & Schreiber, T. (2003). *Nonlinear time series analysis* (2nd ed.). Cambridge University Press.

Kendall, B. E. (2001). Cycles, chaos, and noise in predator–prey dynamics. *Chaos, Solitons & Fractals*, *12*(2), 321–332. https://doi.org/10.1016/S0960-0779(00)00180-6

Kennel, M. B., Brown, R., & Abarbanel, H. D. I. (1992). Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Physical Review A*, *45*(6), 3403–3411. https://doi.org/10.1103/PhysRevA.45.3403

Klein, E., Glaser, S., Jordaan, A., Kaufman, L., & Rosenberg, A. A. (2016). A complex past: Historical and contemporary fisheries demonstrate nonlinear dynamics and a loss of determinism. *Marine Ecology Progress Series*, *557*, 237–246. https://www.int-res.com/abstracts/meps/v557/p237-246/

Kuriyama, P. T., Sugihara, G., Thompson, A. R., & Semmens, B. X. (2020). Identification of shared spatial dynamics in temperature, salinity, and ichthyoplankton community diversity in the California current system with empirical dynamic modeling. *Frontiers in Marine Science*, *7*(October), 1–17. https://doi.org/10.3389/fmars.2020.557940

Lamon, E. C., III, Carpenter, S. R., & Stow, C. A. (1998). Forecasting PCB concentrations in Lake MICHIGAN salmonids: A dynamic linear model approach. *Ecological Applications*, *8*(3), 659–668. https://doi.org/10.1890/1051-0761(1998)008[0659:FPCILM]2.0.CO;2

Leng, S., Ma, H., Kurths, J., Lai, Y. C., Lin, W., Aihara, K., & Chen, L. (2020). Partial cross mapping eliminates indirect causal influences. *Nature*

Communications, *11*(1), 1–9. https://doi.org/10.1038/s41467-020-16238-0

Li, J., Zyphur, M. J., Sugihara, G., & Laub, P. J. (2021). Beyond linearity, stability, and equilibrium: The edm package for empirical dynamic modeling and convergent cross-mapping in Stata. *Stata Journal*, *21*(1), 220–258. https://doi.org/10.1177/1536867X211000030

Liu, H., Fogarty, M., Glaser, S., Altman, I., Hsieh, C., Kaufman, L., Rosenberg, A. A., & Sugihara, G. (2012). Nonlinear dynamic features and co-predictability of the Georges Bank fish community. *Marine Ecology Progress Series*, *464*, 195–207. https://www.int-res.com/abstracts/meps/v464/p195-207/

Lorimer, T., Goodridge, R., Bock, A. K., Agarwal, V., Saberski, E., Sugihara, G., & Rifkin, S. A. (2021). Tracking changes in behavioural dynamics using prediction error. *PLoS ONE*, *16*(5), e0251053. https://doi.org/10.1371/journal.pone.0251053

Ma, H., Leng, S., Aihara, K., Lin, W., & Chen, L. (2018). Randomly distributed embedding making short-term high-dimensional data predictable. *Proceedings of the National Academy of Sciences of the United States of America*, *115*(43), E9994–E10002. https://doi.org/10.1073/pnas.1802987115

Mangel, M., & Clark, C. W. (1989). *Dynamic modeling in behavioral ecology*. Princeton University Press.

May, R. M. (1972). Will a large complex system be stable? *Nature*, *238*(5364), 413–414. https://doi.org/10.1038/238413a0

May, R. M. (1974). *Stability and complexity in model ecosystems*. Princeton University Press.

May, R. M. (2020). What determines population density? In A. Dobson, D. Tilman, & R. D. Holt (Eds.), *Unsolved problems in ecology* (pp. 67–76). Princeton University Press. https://doi.org/10.1515/9780691195322-008

McGowan, J. A., Deyle, E. R., Ye, H., Carter, M. L., Perretti, C. T., Seger, K. D., de Verneil, A., & Sugihara, G. (2017). Predicting coastal algal blooms in southern California. *Ecology*, *98*(5), 1419–1433. https://doi.org/10.1002/ecy.1804

Millar, R. B., & Meyer, R. (2000). Non-linear state space modelling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, *49*(3), 327–342. https://doi.org/10.1111/1467-9876.00195

Munch, S. B., Brias, A., Sugihara, G., & Rogers, T. L. (2020). Frequently asked questions about nonlinear dynamics and empirical dynamic modelling. *ICES Journal of Marine Science*, *77*(4), 1463–1479. https://doi.org/10.1093/icesjms/fsz209

Munch, S. B., Giron-Nava, A., & Sugihara, G. (2018). Nonlinear dynamics and noise in fisheries recruitment: A global meta-analysis. *Fish and Fisheries*, *19*(6), 964–973. https://doi.org/10.1111/faf.12304

Munch, S. B., Poynor, V., & Arriaza, J. L. (2017). Circumventing structural uncertainty: A Bayesian perspective on nonlinear forecasting for ecology. *Ecological Complexity*, *32*, 134–143. https://doi.org/10.1016/j.ecocom.2016.08.006

Munch, S. B., & Rogers, T. L. (2022). *GPEDM: Gaussian Process regression for Empirical Dynamic Modeling*. https://github.com/tanyalrogers/GPEDM

Natsukawa, H., Deyle, E. R., Pao, G. M., Koyamada, K., & Sugihara, G. (2021). A visual analytics approach for ecosystem dynamics based on empirical dynamic modeling. *IEEE Transactions on Visualization and Computer Graphics*, *27*(2), 506–516. https://doi.org/10.1109/TVCG.2020.3028956

Neal, R. M. (1996). *Bayesian learning for neural networks*. Springer-Verlag.

Nova, N., Deyle, E. R., Shocket, M. S., MacDonald, A. J., Childs, M. L., Rypdal, M., Sugihara, G., & Mordecai, E. A. (2021). Susceptible host availability modulates climate effects on dengue dynamics. *Ecology Letters*, *24*(3), 415–425. https://doi.org/10.1111/ele.13652

Nychka, D., Ellner, S., Gallant, A. R., & McCaffrey, D. (1992). Finding chaos in Noisy systems. *Journal of the Royal Statistical Society: Series B (Methodological)*, *54*(2), 399–426. https://doi.org/10.1111/j.2517-6161.1992.tb01889.x

Okuno, S., Aihara, K., & Hirata, Y. (2019). Combining multiple forecasts for multivariate time series via state-dependent weighting. *Chaos*, *29*, 033128. https://doi.org/10.1063/1.5057379

Pao, G. M., Smith, C., Park, J., Takahashi, K., Watanakeesuntorn, W., Natsukawa, H., Chalasani, S. H., Lorimer, T., Takano, R., Rungratsameetaweemana, N., & Sugihara, G. (2021). Experimentally testable whole brain manifolds that recapitulate behavior, arXiv preprint. *ArXiv*, 2106.10627. https://doi.org/10.48550/ARXIV.2106.10627

Park, J., Smith, C., Sugihara, G., & Deyle, E. (2021). *rEDM: Empirical Dynamic Modeling ('EDM')*. https://cran.rproject.org/web/packages/rEDM/index.html

Park, J., & Sugihara, G. (2020). *C++ EDM implementation of EDM tools*. Reviewed by RunTime Analytics.

Rogers, T. L., Johnson, B. J., & Munch, S. B. (2022). Chaos is not rare in natural ecosystems. *Nature Ecology & Evolution*, *6*, 1105–1111. https://doi.org/10.1038/s41559-022-01787-y

Rogers, T. L., & Munch, S. B. (2020). Hidden similarities in the dynamics of a weakly synchronous marine metapopulation. *Proceedings in the National Academy of Sciences of the United States of America*, *117*(1), 479–485. https://doi.org/10.1073/pnas.1910964116

Rogers, T. L., Munch, S. B., Stewart, S. D., Palkovacs, E. P., Giron-Nava, A., Matsuzaki, S., Ichiro, S., & Symons, C. C. (2020). Trophic control changes with season and nutrient loading in lakes. *Ecology Letters*, *23*, 1287–1297. https://doi.org/10.1111/ele.13532

Runge, J., Bathiany, S., Bollt, E., Camps-Valls, G., Coumou, D., Deyle, E., Glymour, C., Kretschmer, M., Mahecha, M. D., Muñoz-Marí, J., van Nes, E. H., Peters, J., Quax, R., Reichstein, M., Scheffer, M., Schölkopf, B., Spirtes, P., Sugihara, G., Sun, J., … Zscheischler, J. (2019). Inferring causation from time series in earth system sciences. *Nature Communications*, *10*(1), 1–13. https://doi.org/10.1038/s41467-019-10105-3

Rypdal, M., & Sugihara, G. (2019). Inter-outbreak stability reflects the size of the susceptible pool and forecasts magnitudes of seasonal epidemics. *Nature Communications*, *10*, 2374. https://doi.org/10.1038/s41467-019-10099-y

Saberski, E., Bock, A. K., Goodridge, R., Agarwal, V., Lorimer, T., Rifkin, S. A., & Sugihara, G. (2021). Networks of causal linkage between eigenmodes characterize behavioral dynamics of Caenorhabditis elegans. *PLoS Computational Biology*, *17*(9), e1009329. https://doi.org/10.1371/journal.pcbi.1009329

Säterberg, T., & McCann, K. (2021). Detecting alternative attractors in ecosystem dynamics. *Communications Biology*, *4*(1), 975. https://doi.org/10.1038/s42003-021-02471-w

Sauer, T., Yorke, J. A., & Casdagli, M. (1991). Embedology. *Journal of Statistical Physics*, *65*(3), 579–616. https://doi.org/10.1007/BF01053745

Schaffer, W. M. (1983). The application of optimal control theory to the general life history problem. *The American Naturalist*, *121*(3), 418–431. https://doi.org/10.1086/284070

Schaffer, W. M. (1984). Stretching and folding in lynx fur returns: Evidence for a strange attractor in nature? *The American Naturalist*, *124*(6), 798–820. https://doi.org/10.1086/284318

Scheffer, M., Carpenter, S. R., Lenton, T. M., Bascompte, J., Brock, W., Dakos, V., Van De Koppel, J., Van De Leemput, I. A., Levin, S. A., Van Nes, E. H., Pascual, M., & Vandermeer, J. (2012). Anticipating critical transitions. *Science*, *338*(6105), 344–348. https://doi.org/10.1126/science.1225244

Schreiber, T. (1997). Detecting and analyzing nonstationarity in a time series using nonlinear cross predictions. *Physical Review Letters*, *78*(5), 843–846. https://doi.org/10.1103/PhysRevLett.78.843

Schwefel, R., Gaudard, A., Wüest, A., & Bouffard, D. (2016). Effects of climate change on Deepwater oxygen and winter mixing in a deep lake (Lake Geneva): Comparing observational findings and modeling. *Water Resources Research*, *52*(11), 8811–8826. https://doi.org/10.1002/2016WR019194

Shertzer, K. W., Ellner, S. P., Fussmann, G. F., & Hairston, N. G. (2002). Predator-prey cycles in an aquatic microcosm: Testing hypotheses of mechanism. *Journal of Animal Ecology*, *71*(5), 802–815. https://doi.org/10.1046/j.1365-2656.2002.00645.x

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

Stark, J. (1999). Delay embeddings for forced systems. I. Deterministic forcing. *Journal of Nonlinear Science*, *9*(3), 255–332. https://doi.org/10.1007/s003329900072

Stark, J., Broomhead, D. S., Davies, M. E., & Huke, J. (1997). Takens embedding theorems for forced and stochastic systems. *Nonlinear Analysis: Theory, Methods & Applications*, *30*(8), 5303–5314. https://doi.org/10.1016/S0362-546X(96)00149-6

Stark, J., Broomhead, D. S., Davies, M. E., & Huke, J. (2003). Delay embeddings for forced systems. II. Stochastic forcing. *Journal of Nonlinear Science*, *13*(6), 519–577. https://doi.org/10.1007/s00332-003-0534-4

Sugihara, G. (1994). Nonlinear forecasting for the classification of natural time series. *Philosophical transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, *348*(1688), 477–495. https://doi.org/10.1098/rsta.1994.0106

Sugihara, G., Casdagli, M., Habjan, E., Hess, D., Dixon, P., & Holland, G. (1999). Residual delay maps unveil global patterns of atmospheric nonlinearity and produce improved local forecasts. *Proceedings of the National Academy of Sciences of the United States of America*, *96*(25), 14210–14215. https://doi.org/10.1073/pnas.96.25.14210

Sugihara, G., Deyle, E. R., & Ye, H. (2017). Reply to Baskerville and Cobey: Misconceptions about causation with synchrony and seasonal drivers. *Proceedings of the National Academy of Sciences of the United States of America*, *114*(12), E2272–E2274. https://doi.org/10.1073/pnas.1700998114

Sugihara, G., May, R., Ye, H., Hsieh, C. H., Deyle, E., Fogarty, M., & Munch, S. (2012). Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500. https://doi.org/10.1126/science.1227079

Sugihara, G., & May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, *344*(April), 734–741. https://doi.org/10.1038/344734a0

SugiharaLab. (2021). *pyEDM*. https://github.com/SugiharaLab/pyEDM. https://pepy.tech/project/pyEDM

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Tajima, S., Yanagawa, T., Fujii, N., & Toyoizumi, T. (2015). Untangling Brain-Wide Dynamics in Consciousness by Cross-Embedding. *PLOS Computational Biology*, *11*(11), e1004537. https://doi.org/10.1371/journal.pcbi.1004537

Takahashi, K., Watanakeesuntorn, W., Ichikawa, K., Park, J., Takano, R., Haga, J., Sugihara, G., & Pao, G. M. (2021). KEDM: A performance-portable implementation of empirical dynamic modeling using Kokkos. In *Practice and Experience in Advanced Research Computing*. Association for Computing Machinery. https://doi.org/10.1145/3437359.3465571

Takens, F. (1981). Detecting strange attractors in turbulence. In D. A. Rand & L. S. Young (Eds.), *Dynamical systems and turbulence* (pp. 366–381). Springer.

Thorson, J. T., Ono, K., Munch, S. B., Thorson, J. T., Ono, K., & Munch, S. E. (2014). A Bayesian approach to identifying and compensating for model misspecification in population models. *Ecology*, *95*(2), 329–341.

Turchin, P., & Taylor, A. D. (1992). Complex dynamics in ecological time series. *Ecology*, *73*(1), 289–305.

Ushio, M., Hsieh, C. H., Masuda, R., Deyle, E. R., Ye, H., Chang, C. W., Sugihara, G., & Kondoh, M. (2018). Fluctuating interaction network and time-varying stability of a natural fish community. *Nature*, *554*(7692), 360–363. https://doi.org/10.1038/nature25504

Viglione, G. (2020). How COVID-19 could ruin weather forecasts and climate records. *Nature*, *580*(7804), 440–441. https://doi.org/10.1038/d41586-020-00924-6

Ye, H., Clark, A., Deyle, E., & Munch, S. (2019). *rEDM: Applications of Empirical Dynamic Modeling from Time Series*. https://ha0ye.github.io/rEDM, https://github.com/ha0ye/rEDM.

Ye, H., Deyle, E. R., Gilarranz, L. J., & Sugihara, G. (2015). Distinguishing time-delayed causal interactions using convergent cross mapping. *Scientific Reports*, *5*, 14750. https://doi.org/10.1038/srep14750

Ye, H., & Sugihara, G. (2016). Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality. *Science*, *353*(6302), 922–925. https://doi.org/10.1126/science.aag0863

How to cite this article: Munch, S. B., Rogers, T. L., & Sugihara, G. (2022). Recent developments in empirical dynamic modelling. *Methods in Ecology and Evolution*, *14*, 732–745. https://doi.org/10.1111/2041-210X.13983