BETHANY JOHNSON (Orcid ID : 0000-0003-2463-4086)

Article type : Research Article

Handling editor: Professor Robert Freckleton

Leveraging spatial information to forecast nonlinear ecological dynamics

Bethany Johnson[1], Marcella Gomez[1], Stephan B. Munch[2]

[1] *Department of Applied Mathematics, University of California, Santa Cruz, CA, USA*

[2] *Southwest Fisheries Science Center, Fisheries Ecology Division, National Oceanic and Atmospheric Administration, Santa Cruz, CA USA*

Corresponding author: Bethany Johnson, 1156 High St, Santa Cruz, CA 95064,
+1 (559) 903 - 0497, bejajohn@ucsc.edu

Running head: Ecological forecasting with spatial information

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the *Version of Record*. Please cite this article as *doi: 10.1111/2041-210X.13511*

Abstract

1. There has been a recent demand for forecasting in ecology, particularly in the field of ecosystem management. Empirical dynamic modeling (EDM), an equation-free nonlinear forecasting method, is receiving growing attention, but it requires long time series to produce accurate predictions. Though most ecological time series are short, spatial replicates are often available. Here we explore how utilizing available spatial data can improve our ability to forecast ecological dynamics.
2. There are several ways to incorporate spatial information into EDM and not all have been applied in ecology. We compare spatial EDM approaches used in ecology and physics and introduce a flexible Bayesian model that makes use of prior movement information.
3. We test these methods on simulated data generated with three population dynamics models with varying levels of complexity, time series length, spatial symmetry, and heterogeneity. Adding spatial data generally improves accuracy, though the best method depends on the spatial process. We applied the methods to empirical fisheries data, highlighting the complexity of real population dynamics.
4. Leveraging spatial data is an effective way to overcome the problem of short ecological time series. Since the best forecasting method depends on the underlying dynamics, we suggest that users apply several in concert and that this may be useful in identifying spatial heterogeneity in dynamics.

Key-words: Bayesian prior, Gaussian process, empirical dynamic modeling, nonlinear forecasting, spatial heterogeneity, spatiotemporal forecasting, time-delay embedding

Introduction

Over the past three decades there have been numerous calls to make ecology a more predictive science (Clark, 2001; Dietze, 2017; Dietze et al., 2018; Evens et al., 2012). One of the most pressing demands for ecological forecasting is in the field of ecosystem management (Christensen et al. 1996; Panel, 1999; Pikitch, 2004), which accounts for interactions among competing ecosystem services when setting management policy. Unfortunately, in most

ecosystems, we do not have complete knowledge of all the system variables and their interactions. In such partially observed systems, nonlinear interactions between numerous species and exogenous variables make mechanistic predictions difficult (Wood and Thomas, 1999).

Alternatively, equation-free approaches, collectively called empirical dynamic modeling (EDM) (Sugihara and May, 1990; Takens, 1981; Ye et al., 2015), can be used to make predictions in partially observed systems. The state variables of a dynamical system evolve through time and tend to converge on a set of values (e.g. a fixed point, limit cycle, or complex shape) referred to as the attractor for the system (Chang et al., 2017). Takens' theorem (Takens, 1981) demonstrates that time lags of a single variable can be used as synthetic coordinates to reconstruct an image of the attractor. In an ecological context, this means that it is possible to use historical data from a single species to reconstruct the dynamics even when we do not have data for other species in the community. In practice, this is done by constructing a collection of short segments of the time series, $\{x_{t-E}, x_{t-E+1}, \ldots, x_t\}$ and using this 'library' to identify similar segments. If the segments are long enough and the dynamics are smooth, similar segments imply similar future states allowing robust, model-free predictions to be made. See Movie S1 of (Ye et al., 2015) for an explanation. EDM is increasingly used in ecology to forecast population size (Perretti et al., 2013), identify causal connections (Sugihara et al., 2012), quantify species interactions (Deyle et al., 2016; Rogers et al., 2020), and identify chaotic dynamics (Sugihara, 1994; Sugihara and May, 1990). However, EDM requires long time series and frequent sampling to fully resolve the attractor. This limits the accuracy of EDM in ecology where time series are often short (20-50 years).

There are several avenues to overcome the problem of short time series. For instance, libraries from several state variables (i.e. species) can be concatenated to increase the density of points on a common attractor (Banbrook et al., 1997; Hsieh et al., 2008; Sugihara, 1994). This method improves predictions, but the obvious trade-off is that it requires data for multiple species. In cases where data for only one species are available from multiple locations, it may be possible to use the available spatial information and improve forecasts.

The state of the art for incorporating spatial information into EDM varies by discipline. In ecology, it is common to concatenate data from multiple sites to fill in the attractor more completely (Clark et al., 2015; Glaser et al., 2014). Hierarchical approaches have also been used to combine information across sites (Munch et al., 2017; Rogers and Munch, 2020). Although

these methods often improve forecasts, they rely on two opposing requirements. First, the spatial replicates must be similar enough that their dynamics are mutually informative. Second, the replicates must be different enough to provide new information. For instance, highly synchronized replicates contribute very little. These methods use only lags of the local time series to construct an average attractor across multiple sites. However, nearby populations can exchange migrants or share common drivers. In this case, the dynamics at a focal site depend on the state of its neighbors, so lags from neighboring sites may provide additional information for forecasts.

Spatial EDM (sEDM), uses both spatial neighbors and temporal lags to leverage this spatial coupling. sEDM has been used in a variety of fields since its introduction (Ørstavik and Stark, 1998; Parlitz and Merkwirth, 2000) including forecasting coastline changes (Grimes et al., 2015) and sunspots (Covas, 2017; Covas and Benetos, 2019). Although sEDM seems like a promising avenue for forecasting in ecology, there is, to our knowledge, no ecological application of it, and it has not been compared to the method of concatenating data. Here we aim to determine which of these approaches more effectively utilizes spatial information to improve ecological forecasts.

Since sEDM is typically applied outside ecology, there are three limitations to address to make it suitable for ecological forecasting. First, sEDM applications usually construct predictions for each focal location independently. That is, although sEDM inputs include lags of the focal site and neighbors, the library of data used to make predictions is based on predictor-target pairs centered at the focal site. Including predictor-target pairs centered at other locations by concatenating libraries (Clark et al., 2015; Glaser et al., 2014) could possibly improve sEDM predictions. This would simultaneously account for dynamic influences of dispersal or gradients and also borrow information across locations on the shape of the attractor.

Second, standard sEDM implementation makes simplifying assumptions including symmetric coupling between sites and identical dynamics across sites. Since many ecological systems break these assumptions (e.g. with advection and spatial heterogeneity) (Barnett et al., 2019; Kolasa et al., 1991; Largier, 2003), these assumptions should be relaxed when evaluating these methods.

Finally, there is no consensus on how to choose which neighbors and lags to include in attractor reconstructions with sEDM. Mutual information and false nearest neighbors can be used (Abarbanel, 1997; Covas, 2017; Covas and Benetos, 2019; Kantz and Schreiber, 2004), or combinations of neighbors and lags may be selected to minimize prediction error on some subset

of training data (Bialonski et al., 2015). However, these methods force the spatial and temporal lags to be equally spaced, and a more flexible lag spacing may be preferable (Judd and Mees, 1998).

Automatic relevance determination (ARD) (MacKay and Neal, 1994; Neal, 1996), a technique used in machine learning, is useful for selecting relevant lags in Bayesian approaches to EDM (Munch et al., 2018, 2017). However, the multivariate embedding theorem (Deyle and Sugihara, 2011) asserts that, from a theoretical perspective, sEDM is inherently non-identifiable. Here, it is important to distinguish between model identifiability, a well-known statistical problem, and dynamical identifiability which arises because there are multiple ways to *exactly* reconstruct dynamics (Deyle and Sugihara, 2011). For instance, in a predator-prey system, we could write a model using either the current densities of both species, lags of the predator, or lags of the prey. All three formulations are equally valid and produce identical dynamics. In contrast, statistical identifiability occurs whenever multiple parameter combinations produce identical values for the likelihood and can often be eliminated by reparameterization. The lack of dynamical identifiability in sEDM implies that different lag selections are equally valid and will produce nearly equivalent fits, making it highly likely that a standard implementation of ARD would settle at a local maximum. This is an issue because settling at a local maximum in-sample could reduce out-of-sample predictive accuracy. Moreover, lack of identifiability implies that ARD may select different lags for two realizations of the same dynamics, reducing mechanistic interpretability of lag selections. Ideally, lag selections should be consistent across repetitions of equivalent dynamics.

To overcome this problem, we introduce a modification of ARD intended to resolve this identifiability issue. We use the physical network topology of a spatiotemporal system to derive a Bayesian prior for the expected relevance of each lag. By using this prior, ARD is more likely to converge to a collection of lags that are biologically and physically plausible, without restricting the inference to a fixed set.

Here we examine how to leverage spatial information with EDM and address how our modifications of sEDM compare to the standard implementation of sEDM and other EDM methods that are used in ecology. Specifically, we compare the performance of several combinations incorporating local lags (EDM) vs. spatial lags (sEDM) and local vs. concatenated libraries. For models with spatial lags, we compare performance with and without the informative

prior. We present results on simulated data generated with population dynamics models and also on empirical data for several widely distributed marine species.

Materials and Methods

Empirical dynamic modeling and spatial extensions

EDM is grounded in Takens' theorem (Takens, 1981), which states that an image of the attractor of a multi-dimensional deterministic dynamical system can be reconstructed using the time series of a single state variable, $x_0,...,x_T$. From the single-variable time series, *embedding vectors* of the form $[x_t,x_{t-1},...,x_{t-E}]^\top$ for $t \in (E,...,T)$ are used to reconstruct the attractor. In other words, the present value of a state variable can be written as a function of its past values. That is, $x_t = f(\mathbf{x_{t-1}})$ where $\mathbf{x_{t-1}} = [x_{t-1},...,x_{t-E}]^\top$ for some unknown function f and embedding dimension E.

As an intuitive example of why EDM should work, consider a predator-prey system in which the dynamics fall on a limit cycle. Given a single observation of the prey abundance, it is impossible to predict the next step because the observation can come from either an increasing or decreasing part of the cycle. This ambiguity disappears if we augment the prey abundance with information on which side of the cycle the observation is. The current predator abundance is one obvious possibility, but the current direction of change in prey abundance (e.g. increasing or decreasing) would also be sufficient. Prey abundance a short time earlier provides the same information. Thus, time lags allow us to reconstruct dynamics in partially observed systems. See (Deyle et al., 2016; Perretti et al., 2013; Sugihara, 1994; Ye et al., 2015) for more details on EDM in an ecological context.

For spatial extensions of EDM, consider a single-variable time series at N locations, $x_0^1,...,$ $x_T^1,...,x_0^N,...x_T^N$. To make use of these spatial replicates in EDM, the common approach in ecology is construct embedding vectors for each location and concatenate libraries across sites. This relies on the assumption that all replicates come from the same attractor (Banbrook et al., 1997), so data from any site can inform dynamics of the others.

Alternatively, sEDM embedding vectors include lags of both the focal location as and its nearest spatial neighbors (Ørstavik and Stark, 1998; Parlitz and Merkwirth, 2000). Essentially, this

says that the state at time t and site i is a function of the past values of the state in sites $i-S,...,i,...,i+S$, given by

$$x_t^i = \quad f(\mathbf{x_{t-1}^{i-S}},...,\mathbf{x_{t-1}^{i}},...,\mathbf{x_{t-1}^{i+S}}), \tag{1}$$

for some unknown function f and embedding dimensions E and S. Here $\mathbf{x_{t-1}^{i}} = \left[x_{t-1}^i,...,x_{t-E}^i\right]^\top$.

Regardless of how spatial data are used, our goal is to approximate f from data so that it can be iterated forward to make predictions. There is a vast literature on function approximation techniques (Iatan, 2016; Judd, 1999; Stalph, 2014), but the most commonly used in ecological EDM are piece-wise constant models (e.g. Simplex (Sugihara and May, 1990)), local linear regression (e.g. S-map (Sugihara, 1994)), splines (e.g. (Ellner and Turchin, 1995)), and Gaussian process (GP) regression (Munch et al., 2017). We implement EDM and sEDM via GP regression to facilitate incorporating auxiliary biological and physical information through prior specification (Thorson et al., 2014). Additionally, the GP framework admits a hierarchical structure to share information across different locations (Rogers and Munch, 2020), though we do not pursue this here.

Here we provide a brief introduction to GP regression, define a naive ARD prior, and then develop a novel ARD prior for sEDM. For additional background, Rasmussen & Williams provides an overview of modeling with GP regression (Rasmussen and Williams, 2006), and Munch et. al. provides examples of using GP regression specifically for EDM (Munch et al., 2017).

In general, we want to fit a model for

$$\boldsymbol{x_{t+1}^i} = \boldsymbol{f}(\mathbf{x}) + \boldsymbol{\epsilon_t}\,, \tag{2}$$

where the approximation error, ϵ_t, has mean 0 and variance V. Note that the input $\mathbf{x}$ has a different structure for EDM and sEDM as described above. To infer f, we assume a GP prior and update using the observed time series. The GP is a continuous generalization of the multivariate normal distribution, completely specified by a mean function, $m(\mathbf{x}) = E[f(\mathbf{x})]$, and a covariance function, $C(\mathbf{x},\mathbf{x}') = E[(f(\mathbf{x}) - m(\mathbf{x}))(f(\mathbf{x}') - m(\mathbf{x}'))]$ (Rasmussen and Williams, 2006). Without information on the characteristics of f *a priori*, we use a constant prior mean function $m(\mathbf{x}) = 0$. Consistent with previous GP applications, we use a squared-exponential covariance function

$$\boldsymbol{C}(\mathbf{x},\mathbf{x}') = \boldsymbol{\sigma_f^2}\mathbf{exp}\Big(-\frac{\mathbf{1}}{\mathbf{2}}(\mathbf{x}-\mathbf{x}')^{\top}\boldsymbol{M}(\mathbf{x}-\mathbf{x}')\Big), \tag{3}$$

with

$$M = \begin{bmatrix} \phi_1 & 0 & \dots & 0 \\ 0 & \phi_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \phi_n \end{bmatrix}.$$

Here, ϕ_i is a characteristic length scale and governs how much f varies in the direction of the i^{th} input (Rasmussen and Williams, 2006). Note that equation (3) has $n+1$ free parameters, ϕ_1 ,...,ϕ_n,σ_f^2. For σ_f^2, we use a weakly informed prior distribution, $\mathrm{Beta}(1.1,1.1)$. We use the same prior distribution for V, the variance in approximation error, which limits uncertainty in the next state to less than twice the observed total variance (assuming that the data have been scaled to unit standard deviation prior to analysis). Priors for ϕ are described below. We estimate these hyperparameters by maximizing the marginal likelihood using the resilient backpropagation (Rprop) algorithm (Riedmiller and Braun, 1993). Subsequent predictions are made conditional on these maximum *a posteriori* estimates (MAP) (Rasmussen and Williams, 2006).

Note that when $\phi_i = 0$, f is constant in the direction of the i^{th} input. To encourage sparsity in f, previous GP-EDM applications (Munch et al., 2017; Rogers and Munch, 2020) have used identical half-normal priors for each ϕ, i.e.

$$\boldsymbol{p}(\boldsymbol{\phi_j}) = \sqrt{\frac{\mathbf{2}}{\boldsymbol{\pi\gamma}}}\mathbf{exp}\Big(-\frac{\boldsymbol{\phi_j^2}}{\mathbf{2\gamma}}\Big), \tag{4}$$

for all $1 \leq j \leq n$. A value of $\gamma = \frac{\pi}{2}$ is typically chosen so f will have a single extremum over a unit interval of the input on average (Sacks and Ylvisaker, 1966), which regularizes f and avoids over-fitting (Munch et al., 2017). While this prior has been successful for GP-EDM, setting identical priors does nothing to ameliorate identifiability issues in sEDM (Fig. 1) and we refer to this model as sEDM with a naive prior. Next, we derive an informed prior for ϕ that makes use of the physical network to promote identifiability.

Physically-Informed Prior Specification

In order to facilitate identifiability, we take advantage of the fact that some combinations of sites are more likely to have coupled dynamics than others. If sites are near each other, it is likely for them to be coupled through similar environmental conditions and dispersal. This coupling decreases for sites that are far apart, so we expect the length scale parameters to decay with distance. We modify the ARD framework to account for this by setting independent but informative half-normal priors for the length scale parameters allowing the mean for each to decrease with distance to the target. For diffusive coupling or, equivalently, nearest neighbor dispersal, the prior expected value of each ϕ is given by

$$\boldsymbol{E(\phi_{\delta,\tau}) \approx \frac{a_{\delta,\tau}}{\sqrt{2\sigma^2}}} \tag{5}$$

where $a_{\delta,\tau} = \frac{1}{\sqrt{2\pi v\tau}} e^{-\frac{\delta^2}{2v\tau}}$. Here, δ and τ are the spatial and temporal distance from the target respectively, v is the mean-square dispersal distance, measured in whatever units are used to measure distance between sites, and σ^2 is the variance of the data. See Supplementary Information SI.2 for a derivation of (5). Since the expected value of the half-normal distribution is $E(\phi) = \sqrt{\frac{2\gamma}{\pi}}$, we set $\gamma = \frac{\pi}{2}(E(\phi))^2$ in (4).

Fig. 1(a) demonstrates the identifiability problem in naive sEDM by displaying length scale estimates for 20 independent but identical simulations of a two-species competition model with nearest-neighbor dispersal (see SI.1 for model and parameter values). In each simulation, we generated 20 time points from random initial conditions on a one-dimensional lattice with 7 sites. Using data from only one species, we included 4 time lags from each site in embedding vectors. ARD identified which of those 28 inputs were relevant for making predictions.

Since the dynamics for each simulation were generated by the same model, ARD should select the same inputs across simulations. However, consistent with the multivariate embedding theorem (Deyle and Sugihara, 2011) and earlier work (Kantz and Schreiber, 2004), the selected lags varied widely, indicating an identifiability problem. Although forecasting may or may not suffer, this lack of identifiability makes interpretation of 'relevance' based on the estimated length

scales impossible. Fig. 1(b) shows analogous outputs when the physically-informed prior specification is used and highlights its ability to improve identifiability.

Local vs. spatial methods

Here we describe the distinction between the methods used. Fig. 2 shows the relationship between input embedding vectors and their outputs for EDM vs. sEDM with and without an informed prior. We can construct multiple embedding vector/output pairs through time so that all N locations have a library of data available to train the GP. Fig. 3 shows different uses of these libraries.

Forecasting Population Dynamics

Here we compare six forecasting methods: EDM (local), EDM (concatenated), naive sEDM (local), naive sEDM (concatenated), informed sEDM (local), and informed sEDM (concatenated). We evaluate these on both simulated and empirical data.

Simulated Data

Forecast performance of EDM methods are sensitive to a variety of factors including the length of time series, system dimension (e.g. number of coupled species), complexity of dynamics (e.g. periodic, chaotic, etc.), and spatial characteristics (e.g. dispersal, heterogeneity, etc.). To understand these factors, we varied them in three ecological models (Table 1) Importantly, these models cover a range of dimensions and ecological interactions to increase the generality of our results.

We used the Table 1 models to simulate spatiotemporal data on a one-dimensional lattice with periodic boundary conditions. For each simulation, we assigned random initial conditions to every variable in every site and removed the first 100 time points to avoid transients. We split the time series into 'training' data for fitting the GP hyperparameters, and 'testing' data for estimating the one-step-ahead forecast accuracy, out of sample. For all models, we used a lattice size of $N = 75$, test set of length $T_{test} = 20$. We used the abundance data for only one species, x, to train and test the methods and ignored the other species abundances.

The training time series length, T_{train}, varied by simulation. We repeated every model 100 times to produce summary statistics for forecast errors. GP regression has $\mathcal{O}(n^3)$ complexity stemming from a matrix inversion step that is required to compute the posterior predictive distribution. Hence, to ease the computational burden, we restricted our forecasting to a subset of locations (i.e. 10 randomly chosen sites used for both training and testing). Importantly, all methods used the same locations within each simulation, but locations varied across the repetitions.

Note that the informed sEDM prior (5) has one additional parameter v for the mean-square dispersal distance. This parameter is not easily estimated from data. To avoid choosing v arbitrarily, we fit the model with v fixed at either: $v = 10$ (i.e. short dispersal distance and steep decay in prior ϕ) or $v = 2^8 \cdot 10 = 2560$ (i.e. long dispersal and nearly identical prior). Note that $v = 2560$ was chosen arbitrarily but is large enough such that all lags are assigned similar priors, making the naive prior a special case of the informed prior. Importantly, users applying this method should consider the units of distance in their data and may need to adjust the candidate values of v. With both sets of hyperparameters corresponding to the different v values, we evaluated leave-one-out predictions on the training data. The v associated with the lower prediction error on training data was used on testing data.

We measured forecast accuracy with the root mean squared error (RMSE) given by

$$\text{RMSE} = \sqrt{\frac{1}{NT_{test}}\frac{1}{}\sum_{i=1}^{N}\sum_{j=1}^{T_{test}}\left(x^{i,j}_{\text{predicted}} - x^{i,j}_{\text{observed}}\right)^2}.$$

where the sum is over the testing data for each site. We chose to use the RMSE instead of Pearson's correlation coefficient – the criterion commonly used in ecological EDM studies – because the RMSE is more widely used in other disciplines and measures whether the absolute magnitude of the predictions and observations match. As a performance benchmark we calculated the RMSE obtained when using the mean of the training data to predict all points in the test data. We refer to this as the "mean predictor."

For both EDM and sEDM we used a temporal embedding dimension up to $E = \sqrt{T_{train}}$ (Cheng and Tong, 1992). In sEDM, we set the spatial embedding dimension to $S = 2$ (i.e. embedding vectors contain five locations) and used ARD to prune irrelevant inputs.

We used this setup in four simulation experiments. The first two addressed how dynamical complexity and time series length affect the forecast performance. These assumed symmetric dynamics (i.e. equal dispersal to nearest neighbors) and spatial homogeneity (i.e. identical parameter values for all sites). Since this may not apply for ecological dynamics (Barnett et al., 2019; Kolasaet al., 1991; Largier, 2003), two additional experiments addressed the impact of asymmetry and heterogeneity.

Effects of Dynamical Regime

We determined how dynamical complexity influences forecast performance by simulating each model in Table 1 with three parameter sets generating periodic dynamics, chaotic dynamics with high spatial synchrony, and chaotic dynamics with low synchrony for $T_{train} = 25$. See Table 1 for parameter values and Table S1 in SI.3for synchrony measures.

Effects of Time Series Length

To examine the effect of time series length on forecasts, we simulated each model using the parameters for chaotic (high synchrony) dynamics for time series lengths $T_{train} = 25, 50,$ and 75.

Effects of Asymmetric Coupling

To test whether advection influences forecasts, we simulated dynamics with unidirectional dispersal. That is, updates were given by

$$x_{t+1}^{i} = \left[(1-\mu) f(\bullet_t^{i}) + \frac{\mu}{3}\left(f(\bullet_t^{i+1}) + f(\bullet_t^{i+2}) + f(\bullet_t^{i+3})\right)\right] e^{\xi_t^{i}}$$

Dynamics were simulated with $T_{train} = 25$ and chaotic (high synchrony) parameters. Fig. 4 illustrates the difference between the two types of coupling compared.

Effects of Spatial Heterogeneity

To generate heterogeneity, we set a spatial gradient in the growth rate of x for each model (Model 1: r, Model 2: a, Model 3: r). Since simulations had periodic boundary conditions, the gradient was given by a sinusoidal function with period N. For example, for Model 1, we defined

$r_i = H_r \sin\left(\frac{2\pi}{N} i\right) + r_0$ for each location i. The amplitude H_r scales the heterogeneity in growth rate, which we varied from 0 to 1. This measure of heterogeneity requires model-specific knowledge that would not be accessible in real applications. As a practical empirical measure, we used the spatial variance in mean abundance given by $H_m = \frac{\max \bar{x}_i - \min \bar{x}_i}{\min \bar{x}_i}$, where $\bar{x}_i$ is the average abundance of the species in site i over the time series. Models in this analysis were simulated with $T_{train} = 25$ and chaotic (high synchrony) parameters.

Empirical Data

To further evaluate the forecasting utility of these methods, we compared them on data for several marine species from the northeast United States continental shelf. The NOAA Northeast Fisheries Science Center (NEFSC) fall bottom trawl survey has sampled from Cape Hatteras, North Carolina to the Gulf of Maine since 1963 (Politis et al., 2014). To avoid any effects of survey design changes, we restricted our analysis to offshore strata from 1973-2008. We predicted dynamics of three different species: longfin squid (*Loligo pealeii*), silver hake (*Merluccius bilinearis*), and butterfish (*Peprilus triacanthus*). These species were chosen because they are widely distributed and have short generation times.

We compared forecast performance for data aggregated at both coarse and fine spatial resolutions. The coarse resolution aggregated the data into four major regions: Mid-Atlantic Bight, Southern New England, Georges Bank, and Gulf of Maine. These regions are grouped to have similar biophysical characteristics and have been used in previous studies (Lucey and Nye, 2010; Nye et al., 2009; Walsh et al., 2015). The fine resolution used the survey strata as spatial sites. Fig. 5 shows the midpoint of all strata and their corresponding major regions. At both resolutions, we used the Euclidean distance between midpoints as δ in equation (5). We selected a temporal embedding dimension between $E = 5$ and $E = 7$ depending on which minimized the average forecast errors. At the coarse resolution, lags from the nearest site to the focal site were included in the sEDM embedding vectors and libraries from all regions were used in the concatenated methods. In the fine resolution analysis, the embedding vectors used lags from the two strata closest to the focal site, and concatenated methods included libraries from sites within the same major region as the focal site.

At both resolutions, population abundance was estimated by averaging the number of individuals in the catch over all tows in the site each year. We computed sequential forecasts (i.e. forecasts that only used data earlier in the time series to predict the next point) through the entire time series and reported errors on the last 10 time series points. We predicted log population abundance, $\ln(x_{t+1})$, for all species.

Results

Results on Simulated Data

Overall, simulations show that utilizing spatial information in EDM is advantageous. However, the characteristics of the dynamics influence which method most effectively leverages spatial information.

Effects of Dynamical Regime

Fig. 6 shows the mean and standard deviation RMSE of one-step-ahead forecasts of the Table 1 models with three levels of complexity. Even with a short training time series of $T_{train} = 25$, concatenating spatial replicates provides accurate predictions ($\sim 75\%$ average error reduction from the mean predictor), and these methods consistently outperform local EDM, which ignores spatial data. Alternatively, including spatial lags without concatenating libraries (local sEDM) does not improve prediction over local EDM, and these methods are more sensitive to dynamical complexity; forecast errors increase as dynamics become more complex and less synchronized. See Table S2 for RMSE and variance explained (R^2) for greater interpretability.

Dynamical complexity plays a role in the forecasting ability of these methods but does not qualitatively change their performance compared to one another. Complexity primarily influences how much improvement is made by using spatial data. For instance, when dynamics are simple (e.g. periodic) all methods perform well and concatenated libraries only reduce forecast errors from local libraries by about 20% on average, but concatenated libraries reduce errors by much more ($\sim$ 65%) when the dynamics are complex (i.e. chaotic and asynchronous).

Effects of Time Series Length

Longer time series yield better predictions for all methods (Fig. 7 and Table S3). As the length of the training time series increases from 25 to 75, the RMSE decreases from about 50% to at least 70% average error reduction from mean predictor. As with dynamical complexity, the performance ranking of the methods does not change with time series length; concatenated methods are consistently best and local sEDM methods are consistently worst. For short time series ($T_{train} = 25$), concatenation reduces forecast error compared to local libraries by about 50%, but all methods produce similar accuracy for long time series ($T_{train} = 75$), so the benefit of concatenation decreases with time series length..

Effects of Asymmetry

Interestingly, results are qualitatively different when dispersal is unidirectional compared to symmetric (Fig. 8 and Table S4). When advection is present, local sEDM outperforms local EDM ($\sim$ 10% average reduction in RMSE) even on the short time series of 25 points. Concatenating time series still produces a substantial improvement. Hence, any method of incorporating spatial data is better than implementing the standard local EDM method in this case.

Effects of Spatial Heterogeneity

As we increase spatial heterogeneity, the RMSEs for all concatenated methods increase until they become worse than their corresponding local methods (Fig. 9). In contrast, local methods are relatively insensitive to spatial heterogeneity. This demonstrates that when spatial replicates become too different, they can no longer be concatenated.

One final result that pertains to all simulations is that sEDM with a physically-informed prior provides minimal predictive advantage over the naive implementation, typically providing no more than 10% reduction in forecast error.

Results on Empirical Data

Results from simulations suggest that utilizing spatial information is likely to provide the strongest advantage over standard local EDM when time series are short and dynamics are complex and asynchronous. This is closely analogous to the data from NEFSC bottom trawl

survey. The time series is only 36 years long, the dynamics vary irregularly, and they are not highly synchronized (Table S1 in SI.3). In general, these data are ideal for evaluating EDM methods because they come from a system that is estimated to support over 5000 species (Fautin et al., 2010) and we do not have complete understanding of all relevant species and drivers.

Concatenated methods always provide more accurate predictions than their corresponding local methods, but the improvement is more substantial at the fine spatial resolution where local methods lose accuracy (Fig. 10). Specifically, concatenating data from all regions at the coarse resolution produces forecast errors that are approximately 12% lower than corresponding local methods on average. Concatenating data from sites within the local region at the fine resolution reduces error by $\sim 20\%$ on average. In most cases, there are only slight differences between EDM and sEDM performance, and the best-performing method is not consistent across species. See Tables S5 and S6 for exact RMSE and R^2.

Discussion

Here we explored various ways to leverage spatial data when forecasting ecological dynamics in partially observed systems. Specifically, we evaluated the forecast performance of the ecologists' approach (concatenated EDM), the physicists' approach (local naive sEDM), and our dispersal-motivated prior to combine information from multiple sites. We compared these methods to classic local EDM, which does not use spatial information. Results indicate that when dynamics are primarily homogeneous, any concatenated method produces substantially better forecasts than local EDM. If dynamics differ significantly among spatial replicates, however, concatenated methods produce poorer predictions since they rely on the assumption that spatial replicates have comparable dynamics. Applying both approaches to a given data set may provide a direct way of identifying heterogeneous dynamics.

Dynamic complexity and time series length play a role in the forecasting ability of all methods. Generally, as complexity decreases and time series length increases, forecasts improve (Figs 6 and 7). Intuitively, when dynamics are regular (e.g. periodic), the attractor can be filled with a short time series and all methods predict dynamics well. When dynamics are chaotic, more data are required to reconstruct the attractor.

Regardless of complexity and time series length in Figs 6 and 7, local sEDM methods consistently have higher errors than local EDM, implying that it is not advantageous to incorporate spatial data through embedding vectors alone. This result is somewhat counterintuitive since EDM is a special case of sEDM. But we can understand this as a trade-off between in-sample fitting and out-of-sample forecasting: sEDM has more length-scale parameters and better in-sample fits on training data, but local EDM is more robust out-of-sample. ARD is insufficient to eliminate the unnecessary spatial lags when dynamics are complex and time series are short.

In the first two simulation experiments, concatenated methods have lower prediction errors than local methods. This is likely because concatenated methods fit a single GP on data from 10 locations while local methods fit a separate GP on data from each location. Therefore, concatenation uses 10 times more data than local methods. Fig. S2 shows that the forecast errors of local and concatenated methods converge as the time series length for local methods increases. Thus, under homogeneous conditions, concatenating data from 10 locations is essentially equivalent to having a time series that is 10 times longer. Therefore, concatenation is an effective way to overcome limitations of short time series provided that the spatial replicates contribute independent information, which is easily determined by measuring the improvement in forecast performance.

With symmetric dispersal, local EDM outperforms local sEDM, but the opposite occurs with unidirectional dispersal (Fig. 8) as the length scale parameters in sEDM can accommodate asymmetry. Spatial heterogeneity is also important because it hinders the utility of concatenated methods (Fig. 9) since they assume stationarity across replicates (i.e. all replicates must come from the same attractor) (Banbrook et al., 1997). This assumption is reasonable if the underlying dynamical equations are identical for all spatial replicates. However, if they vary across sites, information is not easily shared between them.

Given the simulation results, it is reasonable to imagine cases that could yield different qualitative outcomes. Fig. S3 shows that when strong heterogeneity and asymmetry are both present, local sEDM methods are most accurate and concatenated methods are least accurate, contradicting the results in Figs 6 and 7 in nearly every way. Importantly, we cannot identify precise thresholds of asymmetry or heterogeneity at which these changes occur. We present this simply as an important area for future work and a proof-of-concept that these aspects of the dynamics are instrumental in the utility of spatial extensions of EDM.

Although simulations reveal that none of these methods perform uniformly well in every scenario, the fact that their performance depends on characteristics of the dynamics offers interesting potential for diagnosing ecological conditions. For example, if we did not know anything about the dynamics of a system, comparing forecast performance across all of these methods may provide insights into whether advection and heterogeneity are influencing the dynamics. This knowledge of ecological conditions could be useful when making management decisions.

In our empirical analysis on coarse resolution data from the NEFSC survey, local sEDM outperformed local EDM on silver hake and butterfish data but did not on longfin squid data. In light of simulations, this is consistent with advective dispersal of silver hake and butterfish. Interestingly, there is empirical evidence that dynamics in this region are advection-dominated from particle tracking oceanographic models (Lynch et al., 2014), stable isotope studies (Clarke et al., 2009), and genetic data (Mach et al., 2011). Furthermore, concatenated methods produced slightly lower forecast errors than local methods at the coarse resolution. Given the complexity and asymmetry in the empirical data (Table S1), we might expect concatenation to provide a stronger improvement (e.g. right panels of Fig. 6). However, it is possible that a moderate level of heterogeneity hinders the utility of concatenation in these cases. There is ample spatial heterogeneity in environmental drivers (e.g. temperature, salinity, dissolved oxygen) that are often linked to differences in population growth (Hofmann and Powell, 1998), and the ecological communities (e.g. prey, predators, competitors) also differ substantially between these regions . The empirical measures of heterogeneity on these for all species at the coarse resolution further support this deduction: $H_m = 1.54$ for longfin squid, $H_m = 1.23$ for silver hake, and $H_m = 1.82$ for butterfish.

In contrast, at fine resolution, concatenated methods provided greater predictive improvement. This is likely due to the geographic ranges spanned: At the coarse resolution, concatenated methods used data from along the entire coast whereas at the fine resolution, concatenation was restricted within biophysical regions. As a result, there is lower heterogeneity in concatenated data at the fine resolution ($H_m = 1.10$ for longfin squid, $H_m = 0.97$ for silver hake, and $H_m = 1.06$ for butterfish) allowing for information to be more easily shared across sites.

Note that local methods performed more accurately at the coarse spatial resolution data than they did at the fine resolution. One explanation for this is that averaging abundance measures

over several strata at the coarse resolution distorts nonlinear signals that appear at the fine resolution (Glaser et al., 2014; Sugihara et al., 1999, 1990; Ye, 2015), making it easier to predict coarse dynamics accurately.

All analyses reveal that informed sEDM usually produces more accurate predictions than naive sEDM, but the improvement is typically no more than 1-5% (Tables S2, S3, S4, S5, S6). This suggests that although our informed prior helps resolve the lack of identifiability (Fig. 1) and improves interpretability, it does not substantially improve predictive accuracy. This is consistent with the multivariate embedding theorem's notion that different combinations of lags can produce equivalent predictions (Deyle and Sugihara, 2011). Thus, using the informed prior is not necessarily helpful if the goal is to improve predictions, though it may facilitate inference about ecological mechanisms based on the ARD selection of relevant lags (Browne et al., 2008; Marwala, 2015; Munch et al., 2018).

Regardless of whether EDM is used for inference or prediction, it is crucial for EDM users to understand the importance of temporal and spatial scale. If the temporal scale of the dynamics is too large for EDM to capture with time lags or if samples are not taken with high frequency, none of these methods will give accurate forecasts (e.g. SI.9) and a linear or other parametric modeling approach may be more suitable. Similarly, the spatial scale at which data are collected may influence the heterogeneity across sites. For example, if data are collected over a large geographic range at distant sites, dynamics likely vary across locations due to environmental gradients. In this case, concatenated methods are unlikely to perform well. Generally, given the ubiquity of spatial heterogeneity in ecological systems at the scale we typically observe (Aksnes et al., 1989; Juanes and Conover, 1995; Tommasi et al., 2014), we suggest that concatenating libraries should be done in conjunction with the sEDM approach presented here.

Although we have identified asymmetry and heterogeneity as factors that influence the efficacy of each method there are many other factors that could also play a role but were not addressed in our simulations. For example, variation in dispersal strength and distance, and differences in movement between species and age classes are likely to affect which method performs best. In light of this, it would be helpful to develop specific diagnostic tests that provide guidelines for choosing among methods (e.g. a measure of information flow between sites). In early attempts, we considered embedding dimension and spatial synchrony as potential indices, but these proved inconsistent. More work is needed on this topic.

Our main simulations also did not address how these methods perform as we predict multiple steps into the future. However, Fig. S4 shows that all methods produce better predictions than the mean up to 10 steps into the future. Changing the forecast horizon generally does not influence relative performance across methods and we expect our conclusions about complexity, asymmetry, and heterogeneity extend to multi-step predictions.

We compared the six methods here because they are commonly used, they are easy to implement, and they are good starting points for integrating spatial information into EDM. In the future, it would be valuable to evaluate hierarchical approaches (Munch et al., 2017; Rogers and Munch, 2020), approaches that average over several sub-optimal embeddings (Okuno et al., 2020), diffusion maps (Coifman et al., 2005), or deep-learning (Chattopadhyay et al., 2020) to determine which methods are more robust.

All of our analyses made inferences from single time series. This was intended to provide proof-of-concept and determine general conditions that influence our ability to integrate spatial information. However, time series for several species are often available at multiple locations, and studies on multiview (Ye and Sugihara, 2016) and multivariate (Deyle and Sugihara, 2011; Dixon et al., 1999) embedding suggest that predictions may improve if we include information from other species and drivers. We could do this by simply including lags of other species as additional inputs in embedding vectors, we could construct a hierarchical version to share information across species, or we could combine the method of multiview embedding with sEDM. Comparing the benefits of these approaches is potential territory for future work. Additionally, many species are influenced by harvesting or other anthropogenic effects. In these cases, incorporating the past history of exploitation or anthropogenic inputs is likely to improve prediction accuracy. This is an important next step that opens up the possibility of using EDM as a tool for quantitative ecosystem management (Boettiger et al., 2015; Giron-Nava et al., 2017).

The flexibility of EDM to describe complex dynamics with incomplete data makes this a promising avenue for making strong inferences of ecological dynamics and developing multi-species management policies that are robust to structural uncertainty. However, these methods require long time series for convergence. Although typical ecological time series are fairly short, aggregating information across multiple sites - either through concatenation or sEDM - can substantially improve the utility of empirical approaches to ecological dynamics.

Acknowledgments

This work was supported by funding from the Lenfest Oceans program, NOAA's HPCC program, and NOAA Grant #NA19OAR4170353 through the Fisheries-Sea Grant Joint Graduate Fellowship Program in Population and Ecosystem Dynamics. The statements, findings, conclusions and recommendations are those of the authors and do not necessarily reflect the views of Sea Grant. We thank Charles Perretti and Michael Fogarty for providing access to the NEFSC trawl data, and George Sugihara, Ethan Deyle, Marc Mangel, Hongyun Wang, and Hao Ye for stimulating conversations about these ideas. We are grateful to Uttam Bhat and Tanya Rogers, and Adam T. Clark for many helpful comments on this manuscript.

Authors' contributions

BJ and SBM conceived the ideas and designed methodology; BJ analysed the data and led the writing of the manuscript. All authors contributed critically to the drafts and gave final approval for publication.

Data availability

The empirical data from the Northeast Fisheries Science Center fall bottom trawl survey is publicly available at https://catalog.data.gov/dataset/fall-bottom-trawl-survey. Matlab code for forecasting with the six methods used in this paper is available at: https://doi.org/10.5281/zenodo.4041565.

References

Abarbanel, H. D. I. (1997). *Analysis of observed chaotic data*. Springer.

Aksnes, D., Aure, J., Kaartvedt, S., Magnesen, T., and Richard, J. (1989). Significance of advection for the carrying capacities of fjord populations. *Marine Ecology Progress Series*, 50:263–274.

Banbrook, M., Ushaw, G., and Mclaughlin, S. (1997). How to extract lyapunov exponents from short and noisy time series. *IEEE Transactions on Signal Processing*, 45(5):1378–1382.

Barnett, L. A., Ward, E. J., Jannot, J. E., and Shelton, A. O. (2019). Dynamic spatial heterogeneity reveals interdependence of marine faunal density and fishery removals. *Ecological Indicators*, 107, 105585.

Bialonski, S., Ansmann, G., and Kantz, H. (2015). Data-driven prediction and prevention of extreme events in a spatially extended excitable system. *Physical Review E*, 92(4).

Boettiger, C., Mangel, M., and Munch, S. (2015). Avoiding tipping points in fisheries management through gaussian process dynamic programming. *Proceedings of the Royal Society B: Biological Sciences*, 282(1801):20141631.

Browne, A., Jakary, A., Vinogradov, S., Fu, Y., and Deicken, R. (2008). Automatic relevance determination for identifying thalamic regions implicated in schizophrenia. *IEEE Transactionson Neural Networks*, 19(6):1101–1107.

Chang, C.-W., Ushio, M., and Hsieh, C.-H. (2017). Empirical dynamic modeling for beginners. *Ecological Research*, 32(6):785–796.

Chattopadhyay, A., Hassanzadeh, P., and Pasha, S. (2020). Predicting clustered weather patterns: A test case for applications of convolutional neural networks to spatio-temporal climate data. *Scientific Reports*, 10(1).

Cheng, B. and Tong, H. (1992). On consistent nonparametric order determination and chaos. *Journal of the Royal Statistical Society: Series B (Methodological)*, 54(2):427–449.

Christensen, N. L., Bartuska, A. M., Brown, J. H., Carpenter, S., Dantonio, C., Francis, R., Franklin, J. F., Macmahon, J. A., Noss, R. F., Parsons, D. J., and et al. (1996). The report of the ecological society of america committee on the scientific basis for ecosystem management. *Ecological Applications*, 6(3):665–691.

Clark, A. T., Ye, H., Deyle, E. R., Cowles, J., Tilman, G. D., Sugihara, G., and Isbell, F. (2015). Spatial convergent cross mapping to detect causal relationships from short time series. *The Ecological Society of America*.

Clark, J. S. (2001). Ecological forecasts: An emerging imperative. *Science*, 293(5530):657–660.

Clarke, L., Walther, B., Munch, S., Thorrold, S., and Conover, D. (2009). Chemical signatures in the otoliths of a coastal marine fish, menidia menidia, from the northeastern united states: spatial and temporal differences. *Marine Ecology Progress Series*, 384:261–271.

Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F., and Zucker, S. W. (2005). Geometric diffusions as a tool for harmonic analysis and structure definition of data: Multiscale methods. *Proceedings of the National Academy of Sciences*, 102(21):7432–7437.

Covas, E. (2017). Spatial-temporal forecasting the sunspot diagram. *Astronomy and Astrophysics*,605.

Covas, E. and Benetos, E. (2019). Optimal neural network feature selection for spatial-temporal forecasting. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 29(6).

Crutchfield, J. P. and Kaneko, K. (1987). Phenomenology of spatio-temporal chaos. *Series on Directions in Condensed Matter Physics Directions in Chaos* — Volume 1, page 272–353.

Deyle, E. R., May, R. M., Munch, S. B., and Sugihara, G. (2016). Tracking and forecasting ecosystem interactions in real time. *Proceedings of the Royal Society B: Biological Sciences*,283(1822):2015–2258.

Deyle, E. R. and Sugihara, G. (2011). Generalized theorems for nonlinear state space reconstruction. *PLoS ONE*, 6(3).

Dietze, M. C. (2017). *Ecological forecasting*. Princeton University Press.

Dietze, M. C., Fox, A., Beck-Johnson, L. M., Betancourt, J. L., Hooten, M. B., Jarnevich, C. S.,Keitt, T. H., Kenney, M. A., Laney, C. M., Larsen, L. G., Loescher, H. W., Lunch, C. K.,Pijanowski, B. C., Randerson, J. T., Read, E. K., Tredennick, A. T., Vargas, R., Weathers, K. C.,and White, E. P. (2018). Iterative near-term ecological forecasting: Needs, opportunities, and challenges. *Proceedings of the National Academy of Sciences*, 115(7):1424–1432.

Dixon, P. A., Milicich, M., and Sugihara, G. (1999). Episodic fluctuations in larval supply. *Science*, 283(5407):1528–1530.

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

Evans, M. R., Norris, K. J., and Benton, T. G. (2012). Predictive ecology: systems approaches. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1586):163–169.

Fautin, D., Dalton, P., Incze, L. S., Leong, J.-A. C., Pautzke, C., Rosenberg, A., Sedberry, P. S. G., Jr, J. W. T., Abbott, I., Brainard, R. E., Brodeur, M., Eldredge, L. G., Feldman, M.,

Moretzsohn, F., Vroom, P. S., Wainstein, M., and Wolff, N. (2010). An overview of marine biodiversity in united states waters. *PLoS ONE*, 5(8):427–449.

Giron-Nava, A., James, C., Johnson, A., Dannecker, D., Kolody, B., Lee, A., Nagarkar, M., Pao, G., Ye, H., Johns, D., and et al. (2017). Quantitative argument for long-term ecological monitoring. *Marine Ecology Progress Series*, 572:269–274.

Glaser, S. M., Ye, H., and Sugihara, G. (2014). A nonlinear, low data requirement model for producing spatially explicit fishery forecasts. *Fisheries Oceanography*, 23(1):45–53.

Grimes, D. J., Cortale, N., Baker, K., and Mcnamara, D. E. (2015). Nonlinear forecasting of intertidal shoreface evolution. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, 25(10):103–116.

Hofmann, E. E. and Powell, T. M. (1998). Environmental variability effects on marine fisheries: Four case histories. *Ecological Applications*, 8(1).

Hsieh, C., Anderson, C., and Sugihara, G. (2008). Extending nonlinear analysis to short ecological time series. *The American Naturalist*, 171(1):71–80.

Iatan, I. F. (2016). *Modern Neural Methods for Function Approximation*. Springer.

Johnson, B., Gomez, M., and Munch, S.B. (2020). bejajohn/Johnson_etal_2020: Spatial EDM extension methods from Johnson et al. 2020 (Version v0) [Computer software]. Zenodo. https://doi.org/10.5281/ZENODO.4041565

Juanes, F. and Conover, D. (1995). Size-structured piscivory: advection and the linkage between predator and prey recruitment in young-of-the-year bluefish. *Marine Ecology Progress Series*,128:287–304.

Judd, K. and Mees, A. (1998). Embedding as a modeling problem. *Physica D: Nonlinear Phenomena*, 120(3-4):273–286.

Judd, K. L. (1999). *Numerical methods in economics*. MIT Press.

Kantz, H. and Schreiber, T. (2004). *Nonlinear time series analysis*. Cambridge University Press.

Kolasa, J., Pickett, S. T., and Allen, T. F. H. (1991). *Ecological heterogeneity*. Springer-Verlag.

Kot, M. and Schaffer, W. M. (1986). Discrete-time growth-dispersal models. *Mathematical Biosciences*, 80(1):109–136.

Kuramoto, Y. (1984). *Chemical oscillations, waves, and turbulence*. Springer.

Largier, J. L. (2003). Considerations in estimating larval dispersal distances from oceanographic data. *Ecological Applications*, 13(sp1):71–89.

Lucey, S. and Nye, J. (2010). Shifting species assemblages in the northeast us continental shelf large marine ecosystem. *Marine Ecology Progress Series*, 415:23–33.

Lynch, D. R., Greenberg, D. A., Bilgili, A., McGillicuddy, D. J., Manning, J. P., and Aretxabaleta, A. L. (2014). *Particles in the coastal ocean: theory and applications*. Cambridge University Press.

Mach, M. E., Sbrocco, E. J., Hice, L. A., Duffy, T. A., Conover, D. O., and Barber, P. H. (2011).Regional differentiation and post-glacial expansion of the atlantic silverside, menidia menidia,an annual fish with high dispersal potential.Marine Biology, 158(3):515–530.

MacKay, D. and Neal, R. (1994). Automatic relevance determination for neural networks. Technical report, Cambridge University.

Marwala, T. (2015). *Economic modeling using artificial intelligence methods*. Springer.

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

Munch, S. B., Poynor, V., and Arriaza, J. L. (2017). Circumventing structural uncertainty: A bayesian perspective on nonlinear forecasting for ecology. *Ecological Complexity*, 32:134–143.

Nathan, R., Klein, E., Robledo-Arnuncio, J. J., and Revilla, E. (2012). *Dispersal kernels: review*. Oxford University Press.

Neal, R. M. (1996). *Bayesian Learning for Neural Networks*. Springer-Verlag New York.

Neubert, M., Kot, M., and Lewis, M. (1995). Dispersal and pattern formation in a discrete-time predator-prey model. *Theoretical Population Biology*, 48(1):7–43.

Nye, J., Link, J., Hare, J., and Overholtz, W. (2009). Changing spatial distribution of fish stocks in relation to climate and population size on the northeast united states continental shelf. *Marine Ecology Progress Series*, 393:111–129.

Okuno, S., Aihara, K., and Hirata, Y. (2020). Forecasting high-dimensional dynamics exploiting suboptimal embeddings. *Scientific Reports*, 10(1).

Ørstavik, S. and Stark, J. (1998). Reconstruction and cross-prediction in coupled map lattices using spatio-temporal embedding techniques. *Physics Letters A*, 247(1-2):145–160.

Panel, U. S. E. P. A. (1999). *Ecosystem-based fishery management: a report to Congress by the Ecosystem Principles Advisory Panel*. U.S. Dept. of Commerce, National Oceanic and Atmospheric Administration, National Marine Fisheries Service.

Parlitz, U. and Merkwirth, C. (2000). Prediction of spatiotemporal time series based on reconstructed local states. *Physical Review Letters*, 84(9):1890–1893.

Perretti, C. T., Munch, S. B., and Sugihara, G. (2013). Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. *Proceedings of the National Academy of Sciences*, 110(13):5253–5257.

Pikitch, E. K. (2004). Ecology: Ecosystem-based fishery management. *Science*, 305(5682):346–347.

Pinto, S. E. D. S. and Viana, R. L. (2000). Synchronization plateaus in a lattice of coupled sine-circle maps. *Physical Review E*, 61(5):5154–5161.

Politis, P. J., Galbraith, J. K., Kostovick, P., and Brown, R. W. (2014). *Northeast Fisheries Science Center bottom trawl survey protocols for the NOAA Ship Henry B. Bigelow.* US Dept Commer, Northeast Fish Sci Cent Ref Doc. 14-06; 138 p. Available from: National Marine FisheriesS ervice, 166 Water Street, Woods Hole, MA 02543-1026.

Rasmussen, C. E. and Williams, C. K. I. (2006). *Gaussian process for machine learning*. The MIT Press.

Ricker, W. E. (1954). Stock and recruitment. *Journal of the Fisheries Research Board of Canada*, 11(5):559–623.

Riedmiller, M. and Braun, H. (1993). A direct adaptive method for faster backpropagation learning: the RPROP algorithm. *IEEE International Conference on Neural Networks.*

Rogers, T. L. and Munch, S. B. (2020). Hidden similarities in the dynamics of a weakly synchronous marine metapopulation. *PNAS*, 117(1):479–485.

Rogers, T. L., Munch, S. B., Stewart, S. D., Palkovacs, E. P., Giron-Nava, A., Matsuzaki, S. S., and Symons, C. C. (2020). Trophic control changes with season and nutrient loading in lakes. *Ecology Letters*.

Sacks, J. and Ylvisaker, D. (1966). Designs for regression problems with correlated errors. *The Annals of Mathematical Statistics*, 37(1):66–89.

Stalph, P. (2014). *Introduction to Function Approximation and Regression*. Springer Fachmedien Wiesbaden GmbH.

Sugihara, G. (1994). Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions of the Royal Society of London: Mathematical, Physical and Engineering Sciences,* 348.

Sugihara, G., Casdagli, M., Habjan, E., Hess, D., Dixon, P., and Holland, G. (1999). Residual delay maps unveil global patterns of atmospheric nonlinearity and produce improved local forecasts. *Proceedings of the National Academy of Sciences*, 96(25):14210–14215.

Sugihara, G., Grenfell, B., May, R. M., Chesson, P., Platt, H. M., and Williamson, M. (1990). Distinguishing error from chaos in ecological time series. *Philosophical Transactions of The Royal Society B Biological Sciences*, 330(1257):235–251.

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

Sugihara, G. and May, R. M. (1990). Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature*, 344(6268):734–741.

Takens, F. (1981). Detecting strange attractors in turbulence. *Lecture Notes in Mathematics Dynamical Systems and Turbulence, Warwick 1980*, pages 366–381.

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

Tommasi, D., Hunt, B. P. V., Allen, S. E., Routledge, R., and Pakhomov, E. A. (2014). Variability in the vertical distribution and advective transport of eight mesozooplankton taxa in spring in rivers inlet, british columbia, canada. *Journal of Plankton Research*, 36(3):743–756.

Vasconcelos, D., Viana, R., Lopes, S., Batista, A., and Pinto, S. D. S. (2004). Spatial correlations and synchronization in coupled map lattices with long-range interactions. *Physica A: Statistical Mechanics and its Applications*, 343:201–218.

Walsh, H. J., Richardson, D. E., Marancik, K. E., and Hare, J. A. (2015). Long-term changes inthe distributions of larval and adult fish in the northeast u.s. shelf ecosystem. *PLoS One*, 10(9).

Wood, S. N. and Thomas, M. B. (1999). Super sensitivity to structure in biological models. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, 266(1419):565–570.

Ye, H. (2015). *Nonlinear Tools for a Nonlinear World: Applications of Empirical Dynamic Modeling to Marine Ecosystems*. PhD thesis.

Ye, H., Beamish, R. J., Glaser, S. M., Grant, S. C. H., Hsieh, C.-H., Richards, L. J., Schnute, J. T., and Sugihara, G. (2015). Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proceedings of the National Academy of Sciences*, 112(13).

Ye, H. and Sugihara, G. (2016). Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality. *Science*, 353(6302):922–925.

Yu, H., Zhao, M., Lv, S., and Zhu, L. (2009). Dynamic complexities in a parasitoid-host-parasitoid ecological model. *Chaos, Solitons Fractals*, 39(1):39–48.

Zhang, Ma, Huang, Xuebing, Gao, Zichun, Zhang, Feifan, and Cong (2018). Complex dynamics on the routes to chaos in a discrete predator-prey system with crowley-martin type functional response. *Discrete Dynamics in Nature and Society*.

#	Name	Model	Parameters
1	Ricker (1 species: x)	$f(x) = xe^{r(1-x)}$	$r = \{3,3,3.5\}$
2	Predator-Prey (2 species: x,y)	$f(x,y) = xe^{a - x - \frac{by}{(1+\alpha x)(1+\beta y)}}$ $g(x,y) = ye^{-c + \frac{dx}{(1+\alpha x)(1+\beta y)}}$	$a = \{2.4,2.6,2.8\}$ $b = 2,\ c = 2,\ d = 1.75$ $\alpha = 0.1\ ,\ \beta = 0.1$
3	Host-Parasitoid-Parasitoid (3 species: x,y,z)	$f(x,y,z) = xe^{r(1-x/K) - ay^{-m+1} - bz^{-n+1}}$ $g(x,y,z) = x\left(1 - e^{-ay^{-m+1} - bz^{-n+1}}\right)\frac{ay^{-m+1}}{ay^{-m+1} + bz^{-n+1}}$ $h(x,y,z) = x\left(1 - e^{-ay^{-m+1} - bz^{-n+1}}\right)\frac{bz^{-m+1}}{ay^{-m+1} + bz^{-n+1}}$	$a = 0.4,\ K = 50$ $m = 0.7,\ n = 0.4$ $r = \{3.2,3.2,3.6\}$ $b = \{0.75,0.84,0.84\}$

Table 1: Model structure and parameter values for {periodic, chaotic (high spatial synchrony), and chaotic (low synchrony)} dynamics of a single-species Ricker model (Ricker, 1954), a two-species predator prey model (Zhang et al., 2018), and a three-species host-parasitoid-parasitoid model (Yu et al., 2009). We simulated spatiotemporal data on a one-dimensional lattice with dynamics given by

$$x_{t+1}^{i} = \left[(1-\mu) f(\bullet_t^{i}) + \frac{\mu}{2}\left(f(\bullet_t^{i-1}) + f(\bullet_t^{i+1})\right)\right] e^{\xi_t^{i}}$$

$$y_{t+1}^{i} = \left[(1-\mu) g(\bullet_t^{i}) + \frac{\mu}{2}\left(g(\bullet_t^{i-1}) + g(\bullet_t^{i+1})\right)\right] e^{\xi_t^{i}}$$

$$z_{t+1}^{i} = \left[(1-\mu) h(\bullet_t^{i}) + \frac{\mu}{2}\left(h(\bullet_t^{i-1}) + h(\bullet_t^{i+1})\right)\right] e^{\xi_t^{i}}$$

Here $\bullet \equiv x$ (Model 1), $\bullet \equiv x,y$ (Model 2), $\bullet \equiv x,y,z$ (Model 3), μ is the dispersal rate given by $\mu = \{0.1,0.25,0.25\}$ (Model 1), $\mu = \{0.25,0.25,0.25\}$ (Model 2), and $\mu = \{0.2,0.3,0.1\}$ (Model 3). The noise term is drawn independently from a normal distribution, i.e. $\xi_t^{i} \sim N(-\frac{s^2}{2}, s^2)$ with $s = 0.1$ (Model 1), $s = 0.05$ (Model 2), and $s = 0.1$ (Model 3).

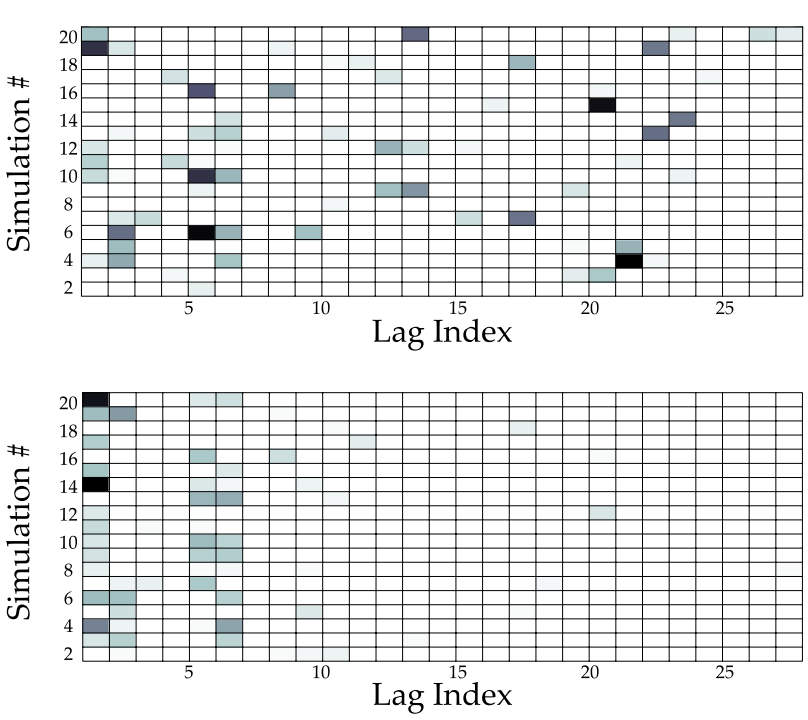


Figure 1: ARD outputs for (a) naive sEDM and (b) informed sEDM. Each row corresponds to a single simulation of a two-species competition model (SI.1). sEDM framework included 28 'lags' in attractor reconstructions. Lag indices are ordered $x_{t-1}^{i},\ldots,x_{t-4}^{i},x_{t-1}^{i-1},\ldots,x_{t-4}^{i-1},x_{t-1}^{i+1},\ldots,x_{t-4}^{i+1},\ldots,x_{t-1}^{i+3},\ldots,x_{t-4}^{i+3}$. Shaded lag indices were determined by ARD to be relevant for making predictions

(darker colors indicate stronger relevance) and white indices were determined to be irrelevant. The strength and index of relevant lags is not consistent in (a), highlighting an identifiability issue. More agreement across simulations in (b) demonstrates the utility of the informed prior.

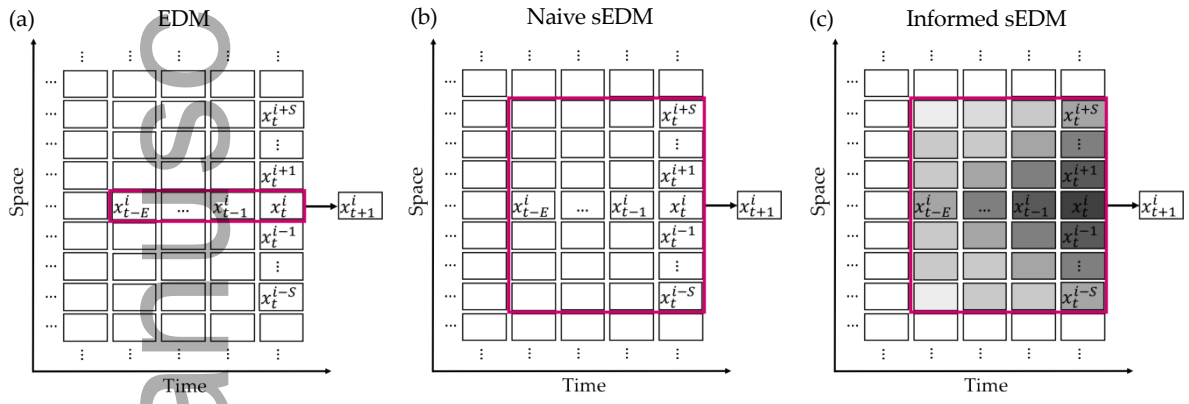


Figure 2: Embedding vectors that map to the one-step-ahead output at location i. The box shows embedding vector entries for EDM (a), sEDM with a naive prior (b), and sEDM with an informed prior (c). The shading of entries represents the prior specification for length scales. Darker colors represent higher values of $\mathbb{E}(\phi)$.

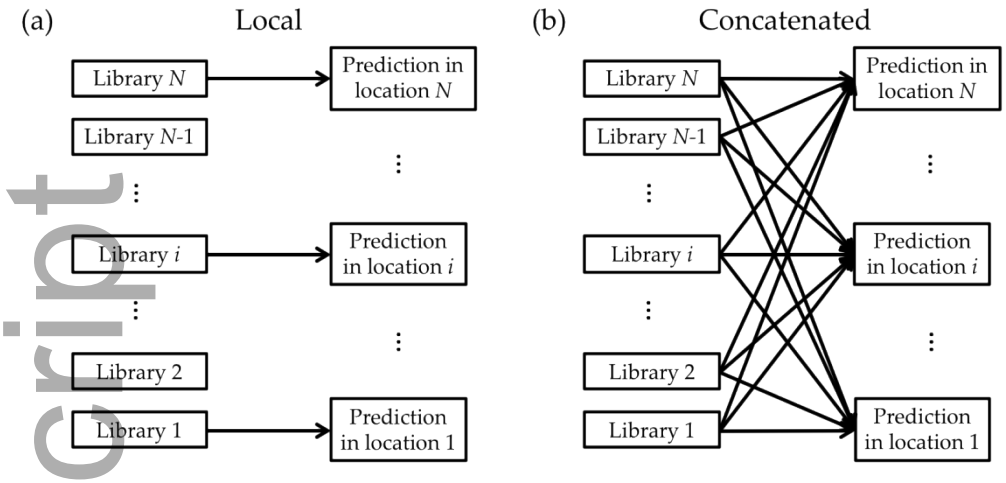


Figure 3: Libraries of embedding vectors used to train the GP and make predictions at location i. A local library uses embedding vectors centered at site i (a). A concatenated library combines libraries across multiple sites (b). A concatenated library may combine data from all locations or a subset of locations.

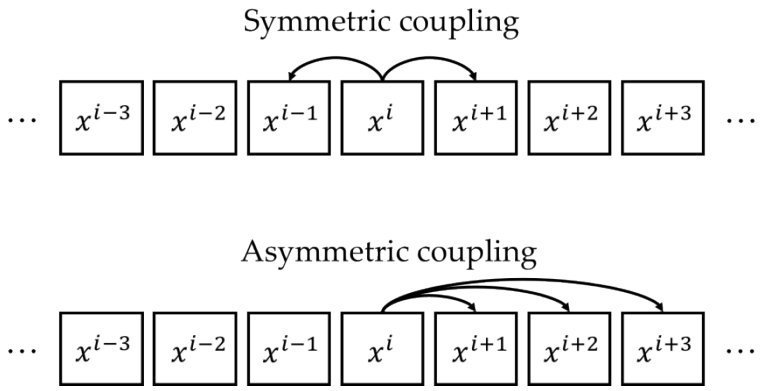


Figure 4: Illustration of dispersal for symmetric and asymmetric coupling schemes.

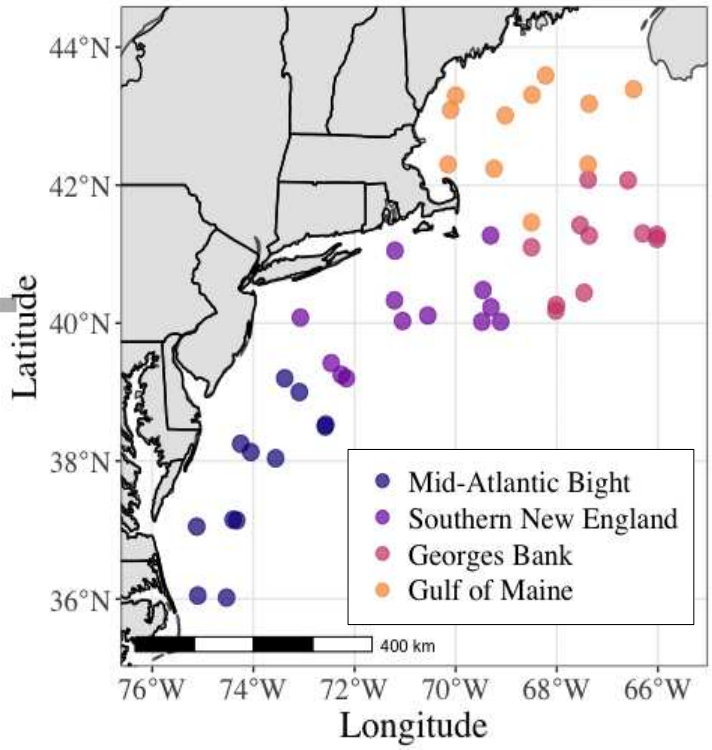


Figure 5: Strata locations from the NEFSC fall bottom trawl survey. Dots are midpoints of offshore strata, and colors indicate which region each stratum is in.

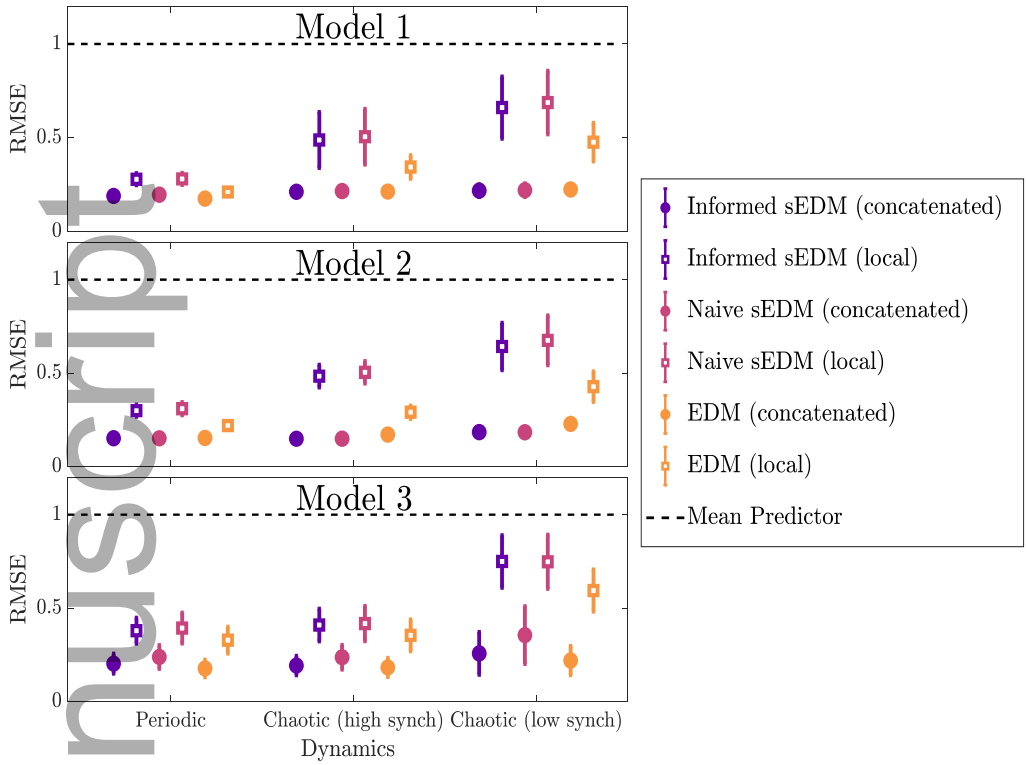


Figure 6: Average RMSE (points) and standard deviation (error bar) on testing data over 100 simulations of each model from Table 1. Results are shown for local (open boxes) and concatenated (closed circles) informed sEDM (purple), naive sEDM (pink), and EDM (yellow). Each model was simulated with 3 parameter sets generating periodic dynamics (left), highly synchronous chaotic dynamics (middle), and asynchronous chaotic dynamics (right).

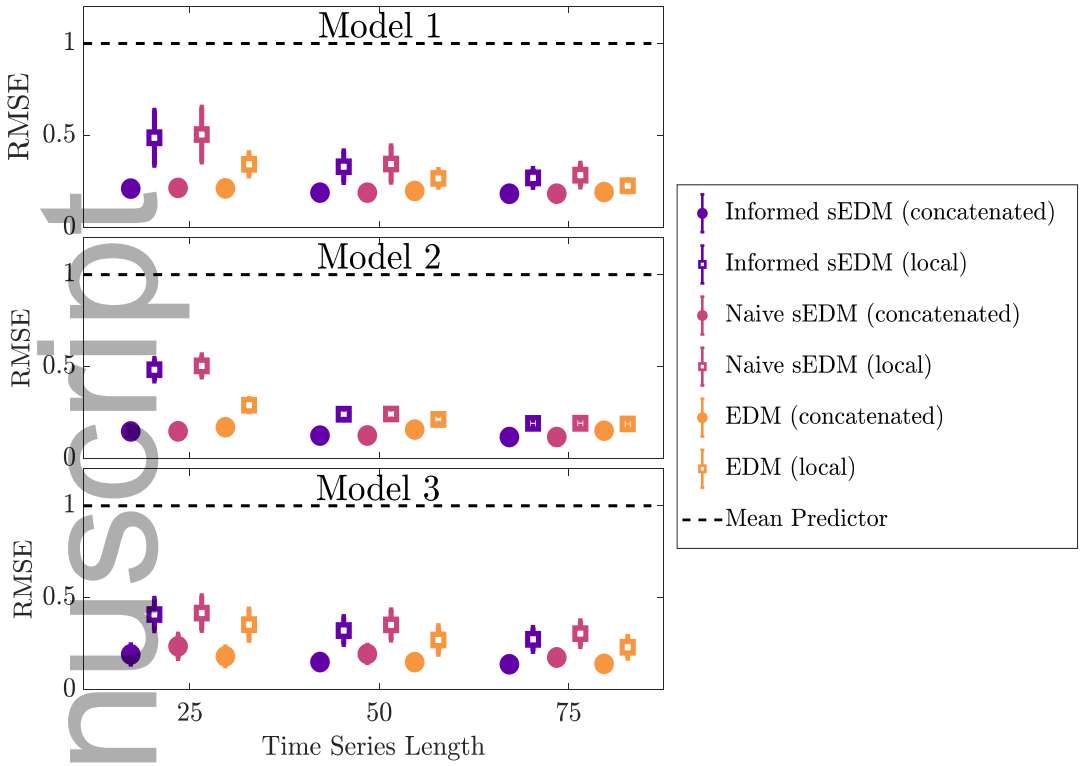


Figure 7: Average RMSE (points) and standard deviation (error bar) on testing data over 100 simulations of each model from Table 1. Results are shown for local (open boxes) and concatenated (closed circles) informed sEDM (purple), naive sEDM (pink), and EDM (yellow). Each model was simulated with 3 time series lengths: $T_{train} = 25$ (left), $T_{train} = 50$ (middle), and $T_{train} = 75$ (right).

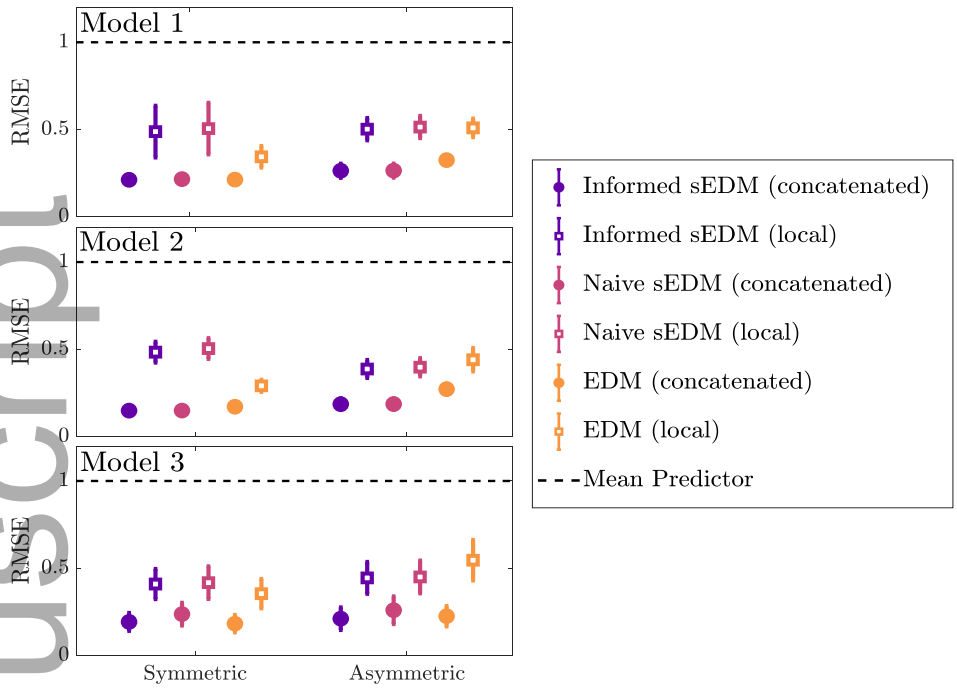


Figure 8: Average RMSE (points) and standard deviation (error bar) on testing data over 100 simulations of each model from Table 1. Results are shown for local (open boxes) and concatenated (closed circles) informed sEDM (purple), naive sEDM (pink), and EDM (yellow). Each model was simulated with symmetric nearest neighbor dispersal (left) and asymmetric unidirectional dispersal (right).

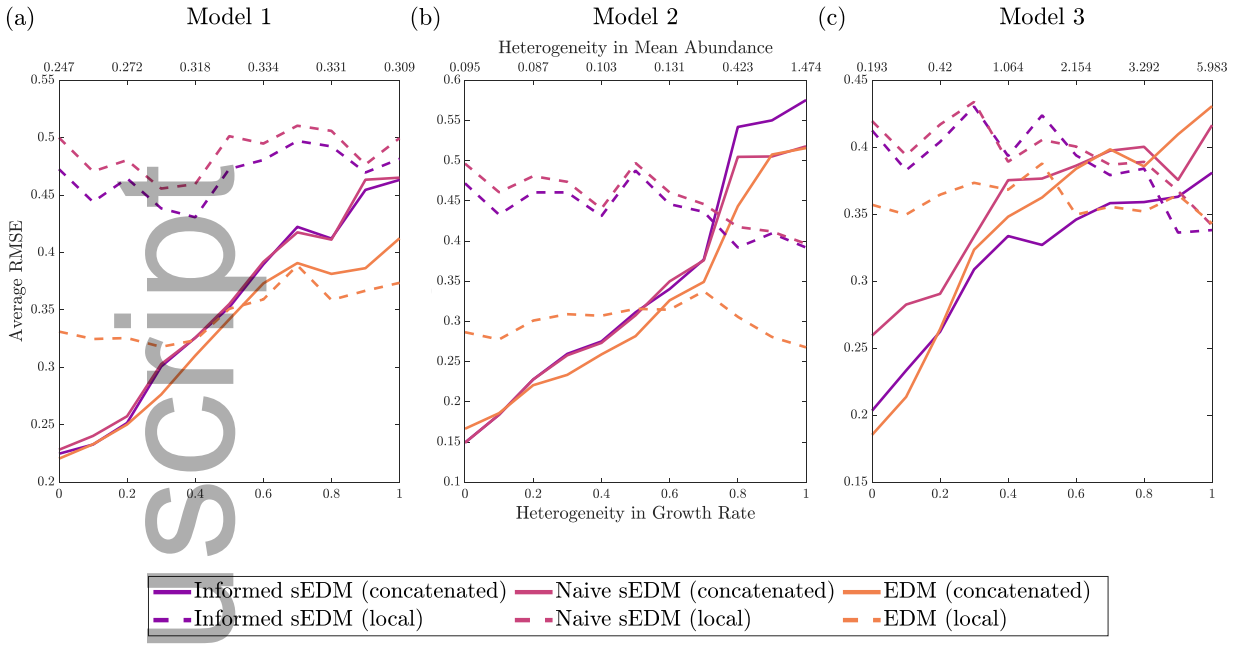


Figure 9: Average RMSE on testing data over 50 simulations of each model from Table 1. for various degrees of heterogeneity. Results are shown for local (dotted lines) and concatenated (solid lines) informed sEDM (purple), naive sEDM (pink), and EDM (yellow). The measure of heterogeneity in growth rate is given by the amplitude of the sinusoidal function used to vary the growth rate parameter. The auxiliary upper x-axis is the empirical measure of heterogeneity in mean abundance (see Methods). Note that this empirical measure is not linear nor monotonic due to complexity and noise in dynamics.

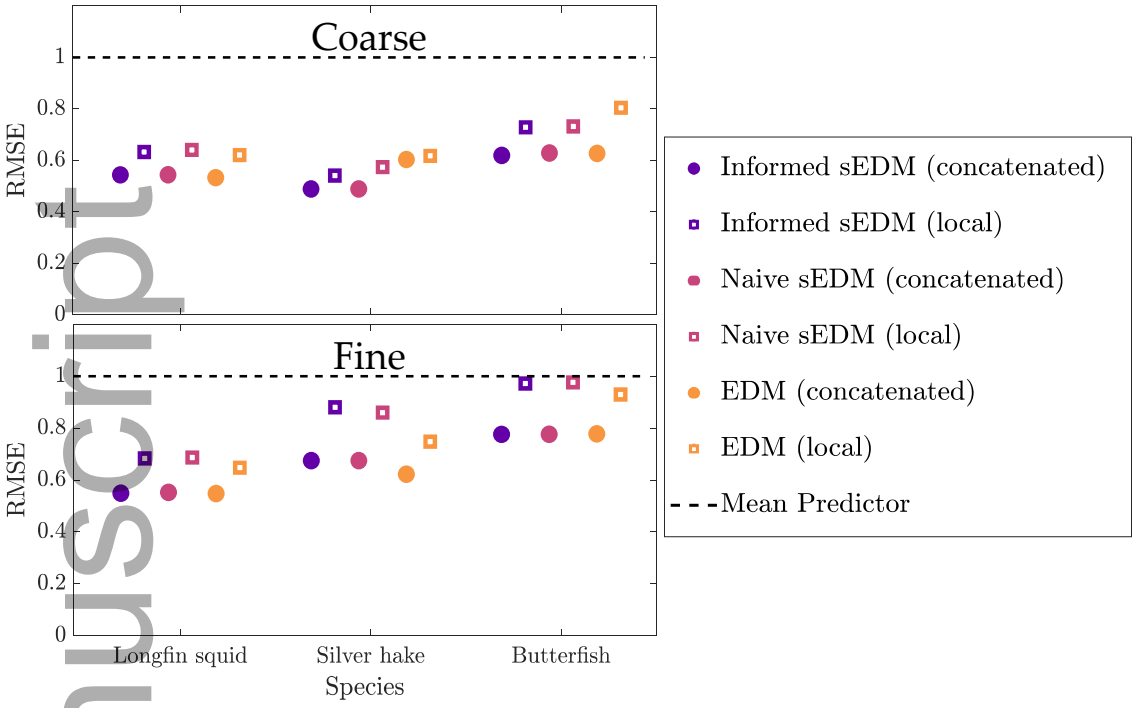
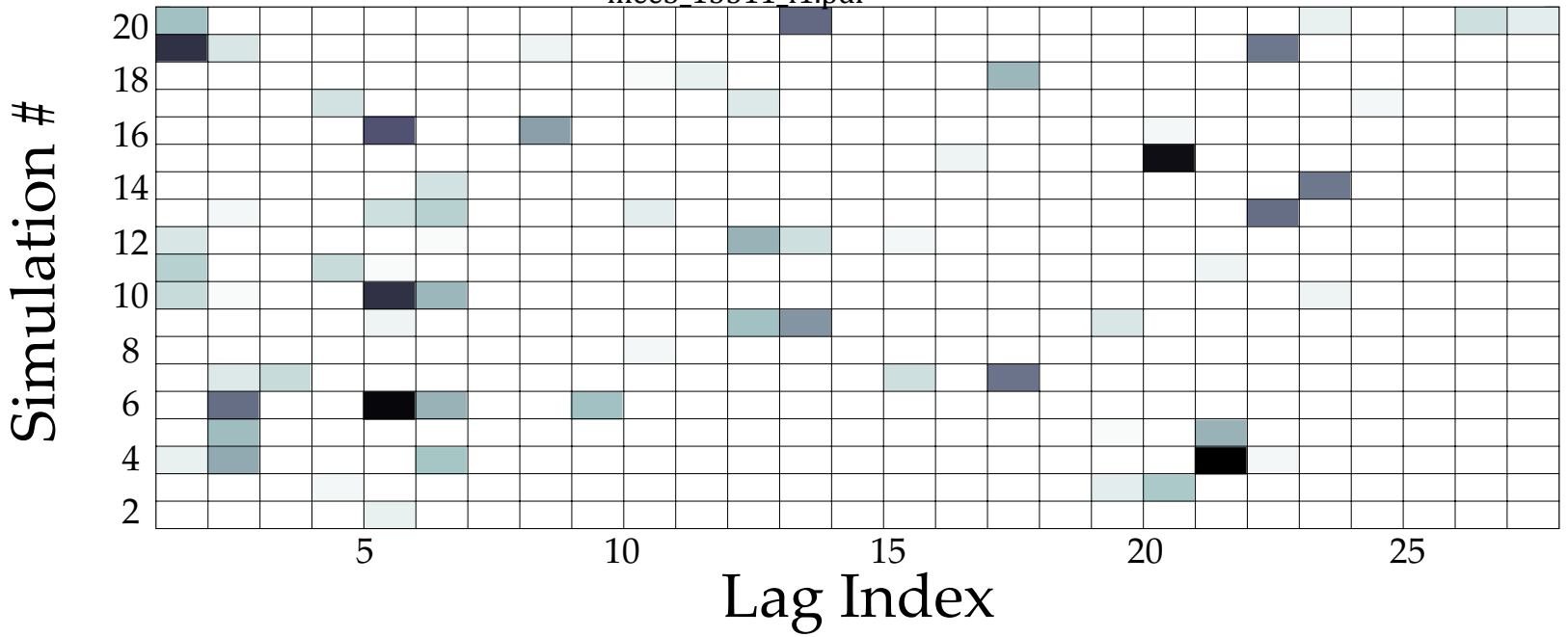


Figure 10: Sequential forecast RMSEs on NEFSC bottom trawl survey data for three species: longfin squid, silver hake, and butterfish. Results are shown for local (open boxes) and concatenated (closed circles) informed sEDM (purple), naive sEDM (pink), and EDM (yellow). Top (bottom) panels show results for coarse (fine) spatial resolution data separated by major biophysical region (survey strata).

(a)



(b)