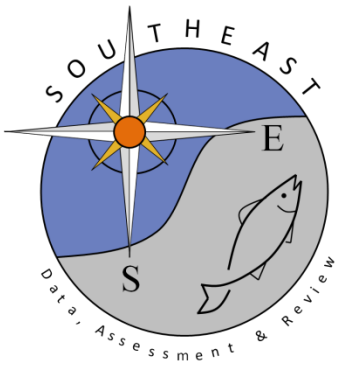
Empirical dynamic modeling for sustainable benchmarks of short-lived species

Cheng-Han Tsai, Stephan B. Munch, Michelle D. Masi, and Molly H. Stevens

SEDAR87-RD-13

November 2024



ICES Journal of Marine Science, 2024, Vol. 81, Issue 7, 1209–1220
https://doi.org/10.1093/icesjms/fsae080
Received: 20 October 2023; **revised:** 25 May 2024; **accepted:** 27 May 2024
Advance access publication date: 21 June 2024
Original Article

Empirical dynamic modeling for sustainable benchmarks of short-lived species

Cheng-Han Tsai [1,*], Stephan B. Munch [2,3,*], Michelle D. Masi [4], Molly H. Stevens [5]

[1]Department of Life Science, National Cheng Kung University, Tainan 701, Taiwan
[2]Department of Applied Mathematics, University of California, Santa Cruz, CA 95064, USA
[3]Southwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Santa Cruz, CA 95060, USA
[4]Southeast Regional Office, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Saint Petersburg, FL 33701, USA
[5]Southeast Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Saint Petersburg, FL 33701, USA

*Corresponding authors. Department of Life Science, National Cheng Kung University, No.1 University Road, Tainan 701, Taiwan.
E-mail: tsaich@gs.ncku.edu.tw; Department of Applied Mathematics, Baskin School of Engineering, 1156 High Street Santa Cruz, CA 95064, USA.
E-mail: smunch@ucsc.edu

Abstract

The abundance dynamics of short-lived marine species often exhibit large-amplitude fluctuations, potentially driven by unknown but important species interactions and environmental effects. These complex dynamics pose challenges in forecasting and establishing robust reference points. Here, we introduce an empirical dynamic modeling (EDM) framework using time-delay embeddings to recover unspecified species interactions and environmental effects, and use walk-forward simulations with varying harvest rates to estimate maximum sustainable yield (MSY). Firstly, we apply our framework to simulated data under various dynamics scenarios and demonstrate the statistical robustness of EDM-based MSY. Secondly, we apply our framework to abundance and catch time series (>30 years) of federally managed brown shrimp stocks in the US Gulf of Mexico. We identify nonlinear signals and achieve high prediction accuracy in the empirical dynamics of brown shrimp. Lastly, based on the EDM of brown shrimp dynamics, we obtain MSY for timely and effective management. Our results highlight the utility of EDM in deriving reference points for short-lived species, particularly in situations where stock abundance and catch dynamics are influenced by unobserved species interactions and environmental effects in a complex ecosystem.

Keywords: short-lived species fishery; nonlinear dynamics; maximum sustainable yield; unobserved species interactions; empirical dynamic modelling

Introduction

Short-lived forage species play a crucial role within a marine ecosystem, transferring energy up the food chain through trophic interactions (Engelhard et al. 2014, Siple et al. 2021). Their population dynamics often exhibit substantial fluctuations, spanning multiple orders of magnitude, rendering them hard to predict (Lindegren et al. 2013, Arkhipkin et al. 2021). These dynamics are potentially nonlinear and chaotic (Deyle et al. 2013, Munch et al. 2018, Tsai et al. 2023), and highly responsive to environmental variability and/or fishing pressure (Hsieh et al. 2006, Essington et al. 2015, Pinsky and Byler 2015). These fundamental characteristics of short-lived species are often overlooked by conventional stock assessments models designed to model the dynamics of long-lived stocks and commonly used to establish fishery reference points. This approach ignores the potential influence of species interactions (Fujiwara et al. 2016, Masi et al. 2018, Munch et al. 2020), and environmental changes, possibly leading to underestimating of the impact of nonlinear feedbacks between competing ecosystem uses (Munch et al. 2018, 2022). Moreover, it may lead to inaccurate estimation of fishery reference points (Glaser et al. 2014, Fogarty et al. 2016) and suboptimal harvesting practices (Brias and Munch 2021). Crucially, for short-lived stocks, the time required for data processing and the development of traditional assessment models can exceed the species' generation time (Arkhipkin et al. 2021). Such delays hinder effective management advice, making it difficult to implement precautionary policy actions in a timely manner (Peterson and Walter 2023).

As an alternative approach, we propose utilizing empirical dynamic modeling (hereafter, EDM) (Chang et al. 2017, Munch et al. 2020, 2023) to establish fishery reference points tailored for short-lived species. EDM uses time-delayed observations to account for unobserved state variables (e.g. species traits, interactions with other species, and environmental drivers). Given a sufficient collection of lags, EDM then uses non-parametric function approximation to address the state dependency of nonlinear dynamics prevalent in complex ecosystems (Chang et al. 2017, Munch et al. 2020, 2023). Takens (1981) theorem and its generalizations shows that under a wide range of conditions, there is a one-to-one correspondence between the dynamics inferred using delays of the observables and the dynamics in the native state space. (For further details, readers may refer to the works of Chang et al. 2017, Munch *et al.* 2018, Munch et al. 2023.) By extending EDM to include the history of exploitation, we can predict

the impact of harvest control rates on fisheries dynamics while implicitly accounting for unobserved interactions with other components of the ecosystem (Brias and Munch 2021, Giron-Nava et al. 2021, Munch et al. 2023). The EDM approach may prove particularly advantageous for modeling short-lived fish species, as demonstrated in the cases of California sardine fisheries (Giron-Nava et al. 2021) and Gulf of Mexico penaeid shrimp fisheries (Tsai et al. 2023).

Application of EDM has proven successful across a range of marine species. Notable examples of abundance forecasting include North Pacific albacore (Glaser et al. 2011), Pacific sardine (Deyle et al. 2013), Fraser River salmon (Ye et al. 2015), Gulf of Mexico and Atlantic menhaden (Deyle et al. 2018), Atlantic blue crab (Rogers and Munch 2020), Atlantic white shrimp (Garcia et al. 2007), as well as Gulf of Mexico brown and white shrimp (Tsai et al. 2023). EDM has also been used for inferring causal relationships and dynamical stability in marine systems, including Georges Bank fish communities (Liu et al. 2012), Maizuru reef fish communities (Ushio et al. 2018), and Gulf of Mexico estuary fish communities (Li and Liu 2023). Despite these successes, there is a paucity of cases applying EDM in fishery management. Harvest control policies based on EDM have only been empirically established for California sardine fisheries (Giron-Nava et al. 2021), whereas Brias and Munch (2021, 2024) showed that it is possible to obtain robust harvest policies for several multi-species fishery models. Perhaps the limited use of EDM in fisheries management is due to a misconception that EDM is suitable only for forecasting rather than for establishing fishery reference points, which is more typically the domain of parametric, mechanistically motivated models (Munch et al. 2020, 2022, 2023).

In this study, we aim to provide a framework to showcase the utility of EDM in deriving fishery reference points, particularly for short-lived species (schematic Fig. 1). Importantly, our framework aims to integrate the existing simulation methods, data transformation methods and model selection methods into a single workflow, to improve the statistical robustness of EDM-based reference points. As a proof of concept, we apply the framework to the brown shrimp fishery in the US Gulf of Mexico (GOM). Our objectives are fourfold. First, we investigate whether EDM exhibits out-of-sample prediction capability by combining both fishery-dependent and fishery-independent data. Second, we assess the effectiveness of incorporating data transformations and ecological assumptions regarding density dependence and catchability in improving the prediction skill of EDM. Third, we explore the feasibility of using EDM to estimate steady-state policy and reference points, such as maximum sustainable yield (MSY). To evaluate the robustness of EDM-derived MSY, we compare them with theoretical MSY using simulated data and model selection methods. Fourth, we apply this framework to the brown shrimp fishery in the GOM, thus providing practical insights into its real-world application.

Materials and methods

Gaussian process empirical dynamic modelling (GP-EDM)

We employ Gaussian process (GP) regression as a function approximation technique to capture the nonlinear dynamics of the stock using EDM (hereafter GP-EDM) (Munch et al. 2017, Johnson and Munch 2022). Below, we outline the GP-EDM framework used to predict the univariate fishery-independent catch-per-unit-effort (CPUE) time series, i.e. CPUE or abundance index from surveys. Additionally, we introduce a method to estimate the relative catchability, which establishes the connection between the fishery-dependent catch/landings series and the fishery-independent CPUE data (also refer to schematic Fig. 1).

Let x_t represent the CPUE at time t and consider it as a random function dependent on lag embeddings of CPUE (i.e. $x_{t\text{-}1,\ldots,}\ x_{t\text{-}m}$) for up to m years, where m represents the maximum number of time lags utilized. That is, $x_t = f(x_{t-1}, \ldots, x_{t-m}) + \varepsilon_t$, where f is a random function that approximates the dynamical process and ε_t is the process error. According to Takens theorem (Takens 1981), these lags are a proxy for the true state space. Here, $E = m + 1$ serves as the maximum embedding dimension, where $E \geq 2D$ with D representing the dimension of the true dynamical system. As a practical guideline, empirical observations suggest that the statistically feasible E is less than or equal to $\sqrt{N}$, where N represents the length of the time series data (Munch et al. 2020, 2023).

We employ GP regression to approximate the delay-embedding map f (Munch et al. 2017). The inputs, gathered in a vector $X_{t\text{-}m} = \{x_{t\text{-}1}, \ldots, x_{t\text{-}m}\}$, are incorporated into a Bayesian GP model formulated as follows:

$$\begin{aligned} P\left(x_t | f, X_{t-m}, \phi, \tau, V_e\right) &\sim Normal\left(f\left(X_{t-m}\right), V_e\right) \\ P\left(f | \phi, \tau, V_e\right) &\sim GP\left(0, \Sigma\right) \\ P\left(\phi\right) &\sim Half normal\left(\frac{\pi}{\sqrt{12}}\right) \\ P\left(\tau\right) &\sim Beta\left(1.1, 1.1\right) \\ P\left(V_e\right) &\sim Beta\left(1.1, 1.1\right) \end{aligned} \tag{1}$$

In (1), the first layer represents the probability density of observing the CPUE at time t given the function approximation f, the input data of time lags $X_{t\text{-}m}$, and the parameters ϕ, τ, and V_e, where V_e represents the process noise. The second layer represents the probability density of the function approximation f given the parameters. The unknown function f is assigned a Gaussian process prior which generalizes the multivariate normal distribution with mean zero and covariance function Σ. The covariance function Σ is a tensor product of squared exponential covariances for each input with pointwise variance τ^2 and inverse length scales $\phi = \{\phi_1, \ldots, \phi_m\}$ where the index i is from 1 to m (the maximum number of time lags). Specifically, the covariance function is $\Sigma(X_t, X_s) = \tau^2 \prod_{i=1}^{m} exp[-\phi_i(x_{t-i} - x_{s-i})^2]$ where X_t and X_s are delay coordinate vectors for years t and s. The final layer of the Bayesian model specifies priors for the hyperparameters $\{\phi_1, \ldots, \phi_m, \tau, V_e\}$. We encourage sparsity in the fitted model by assigning a half-normal prior with variance $\frac{\pi}{\sqrt{12}}$ for ϕ_i. This prior choice asserts that f has, on average, one local extremum within the data range, and the mode at zero encourages uninformative lags drop out of the model (i.e. ϕ approaches 0). This prior is widely used in GP regression under the name "automatic relevance determination" (ARD, Neal 1996). To further encourage sparsity, ϕ is determined as "zero" by rounding to the second decimal place. To improve the interpretability of the inverse-length-scale parameter ϕ, time series data are centered and standardized to unit variance prior to analyses. Note that under this prior specification, the variance of a prediction

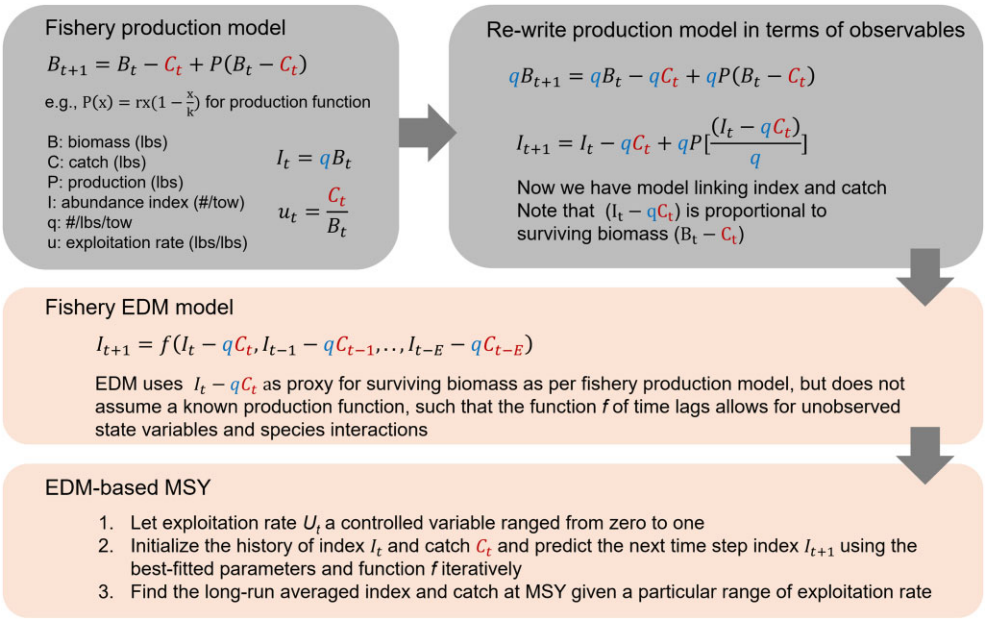


Figure 1. A schematic representation of the GP-EDM framework for deriving the maximum sustainable yield (MSY). This MSY is comparable to that derived from a conventional production model. Notably, when the dynamical system's maximum embedding dimension is set to one, the GP-EDM framework, given the right function approximation, simplifies to a traditional production model. Refer to the Materials and methods section for detailed explanations of how Gaussian process regression is employed to identify the production function and parameters in the EDM framework, as well as for simulation methods used to estimate GP-EDM MSY.

at any input is $V_e + \tau^2$. Since the data are standardized to unit variance, $V_e + \tau^2 \approx 1$ should be sufficient. Nearly flat, independent beta distributions are therefore used as priors for V_e and τ^2 allowing up to twice the variance in the data. (For further details on prior specification and data standardization, refer to Munch et al., 2017.)

Fishery-centric GP-EDM

We expand the delay embedding map to include catch/landings (schematic Fig. 1). One way to do so would be to include CPUE and catch as an independent coordinate, leading to models of the form $x_t = f(x_{t-1}, c_{t-1}, \ldots, x_{t-m}, c_{t-m})$ for the delay embedding map. However, doing so doubles the dimension of the input space. Moreover, it ignores the fact that CPUE and catch both depend on the biomass of the stock. To address this issue, we incorporate a catchability parameter denoted by "q" within the GP-EDM framework to establish a connection between CPUE and fishery landings. We assume that spawning occurs after the fishing season and that spawning biomass is proportional to escapement. Hence, if CPUE and landings are proportional to biomass, then $x_{t-1} - qc_{t-1}$ is proportional to spawning biomass. With these assumptions, we set $x_t = f(x_{t-1} - qc_{t-1}, \ldots, x_{t-m} - qc_{t-m})$ which asserts that the stock abundance one year ahead is determined by the escapement following the previous year fishing (i.e. $x_{t-1} - qc_{t-1}$), amended with lags to account for missing state variables. To simplify the notation, we set $C_{t\text{-}m} = \{c_{t\text{-}1}, \ldots, c_{t\text{-}m}\}$ and update the first layer of (1) as:

$$P\left(x_t \mid f, X_{t-m} - qC_{t-m}, \phi, \tau, V_e\right) \sim Normal\left(f\left(X_{t-m} - qC_{t-m}\right), V_e\right) \quad (2)$$

To complete our model specification, we used a nearly flat beta prior on q with the additional constraint on q that escapement is always positive, i.e. $q < \min_t(\frac{x_t}{c_t})$.

Note that, after some algebraic simplification, this is equivalent to the catchability assumption employed in traditional production models (schematic Fig. 1). Hence, GP-EDM with $m = 1$ can be thought of as a nonparametric production model (Thorson et al. 2014). Importantly, incorporating these standard assumptions regarding the relationships among total catch, catchability, and stock abundance enables us to use EDM for the purposes of evaluating harvest controls (Thorson et al. 2012).

Parameter estimation of GP-EDM

We employ the maximum *a posteriori* (MAP) approach to obtain parameter estimates. By collecting the next time steps (i.e. year ahead fishery-independent CPUE) into a column vector $\boldsymbol{y} = \{x_{t+m}, \ldots, x_T\}^T$ and let $\boldsymbol{X} = \{X_{t-m} - qC_{t-m}\}$ represents the matrix of embedding vectors (i.e. time lags of CPUE reduced by fishing), the logarithm of the marginal posterior probability can be expressed as follows:

$$lnP\left(\phi, \tau, V_e \mid \boldsymbol{y}, \boldsymbol{X}\right) = const - \frac{1}{2} ln\left|\boldsymbol{\Sigma} + V_e \boldsymbol{I}\right| - \frac{1}{2} \boldsymbol{y}^T \left|\boldsymbol{\Sigma} + V_e \boldsymbol{I}\right|^{-1} \boldsymbol{y} - \frac{1}{2} lnP\left(\phi, \tau, V_e, q\right) \quad (3)$$

In this equation (3), $\boldsymbol{I}$ represents the identity matrix and $\boldsymbol{\Sigma}$ represents the covariance matrix created by applying the covariance function $\Sigma(X_t, X_s)$ (i.e. the exponential decay kernel) to all lagged observations (i.e. $\boldsymbol{X} = \{X_{t-m} - qC_{t-m}\}$). The term "*const*" represents an arbitrary constant that does not affect

the optimization procedure. The last term on the right-hand side corresponds to the hyperparameter priors.

Prediction skill and model selection of GP-EDM

We utilize the GP-EDM MAP estimates to make predictions for the year-ahead brown shrimp stock abundance (CPUE) time series. The conditional posterior mean and posterior variance for these predictions, based on the GP-EDM estimates and available data, follow a multivariate normal distribution with the mean and the covariance defined as follows:

$$\begin{aligned} M\left(\boldsymbol{X}'\right) &= \Sigma\left(\boldsymbol{X}', \boldsymbol{X}\right)\left[\Sigma\left(\boldsymbol{X}, \boldsymbol{X}\right) + V_e \boldsymbol{I}\right]^{-1} \boldsymbol{y} \\ V\left(\boldsymbol{X}'\right) &= \Sigma\left(\boldsymbol{X}', \boldsymbol{X}'\right) - \Sigma\left(\boldsymbol{X}', \boldsymbol{X}\right)\left[\Sigma\left(\boldsymbol{X}, \boldsymbol{X}\right) + V_e \boldsymbol{I}\right]^{-1} \\ &\quad \times \Sigma\left(\boldsymbol{X}', \boldsymbol{X}\right)^T \end{aligned} \tag{4}$$

In this equation (4), the notations used are consistent with those in (1–3). In (4), M and V represents the posterior mean and posterior variance, respectively. $\boldsymbol{X}$ and $\boldsymbol{X}'$ represents the in-sample and out-of-sample delay coordinate vectors, respectively.

Because both the number of time lags (the embedding dimension) and the number of inverse length scale parameters are unknown, it is necessary to estimate variants with several time lags and inverse length scale parameters, leading to the need to establish model selection procedures. We propose using out-of-sample prediction skill and information-based model selection to determine the best EDM candidate. Many metrics have been proposed to evaluate prediction accuracy, including correlation coefficient (ρ), coefficient of determination (R^2), root mean square error (RMSE), and mean absolute error (MAE). In keeping with other applications of EDM, e.g. Dolan et al. (2023) and Tsai et al. (2023), we assess the performance of GP-EDM using a leave-one-out (LOO) prediction approach. LOO prediction skill was evaluated using the Pearson correlation coefficient (ρ) between the observed and predicted year-ahead CPUE. LOO prediction accuracy is also evaluated with the predictive R^2 computed as follows:

$$R^2 = 1 - \frac{\sum\left(y_{obs} - y_{pred}\right)^2}{\sum\left(y_{obs} - \bar{y}_{obs}\right)^2}, \tag{5}$$

where y_{obs} represents the observed stock abundance one year ahead, y_{pred} represents the predicted stock abundance one year ahead, and $\bar{y}_{obs}$ represents the averaged observed stock abundance. Additionally, using the same notation in (5), we define LOO RMSE and MAE as follows:

$$\begin{aligned} RMSE &= \sqrt{\frac{\sum\left(y_{obs} - y_{pred}\right)^2}{n}} \\ MAE &= \frac{\sum\left|\left(y_{obs} - y_{pred}\right|\right.}{n} \end{aligned} \tag{6}$$

In addition to evaluating out-of-sample (LOO) prediction skill, we introduce two additional model selection procedures for GP-EDM that explicitly account for the effective degrees of freedom of the nonparametric model in order to avoid overfitting. The complexity of a GP model is determined by both the number of inputs and the length scale parameters. A single input with a large φ would have many degrees of freedom, while a model with many inputs but all φ's close to 0 would have relatively few. Hence, to approximate the effective degrees of freedom, df_{approx}, we use the trace of the hat matrix ($\boldsymbol{H}$, Cleveland and Grosse 1991; Krämer and Sugiyama 2011) computed as:

$$\begin{aligned} \boldsymbol{H} &= \Sigma\left(\boldsymbol{X}', \boldsymbol{X}\right)\left[\Sigma\left(\boldsymbol{X}, \boldsymbol{X}\right) + V_e \boldsymbol{I}\right]^{-1} \\ df_{approx} &= tr\left[\boldsymbol{H}\right] \end{aligned} \tag{7}$$

In a linear model setting, df_{approx} reduces to the number of parameters.

Using (7), we obtain an adjusted variance estimator $\tilde{\sigma}^2$ as:

$$\tilde{\sigma}^2 = \frac{\sum\left(y - M\left(X'\right)\right)^2}{n - df_{approx}}, \tag{8}$$

where n represents the length of vector $\mathbf{y}$, i.e. the year-ahead CPUE time series. Based on the calculations in (7) and (8), we derive approximate versions of the Akaike information criteria (AIC) [commonly known as conditional AIC (Vaida and Blanchard 2005)] and Bayesian information criteria (BIC) as follows:

$$\begin{aligned} AIC &= n \cdot ln\left(\tilde{\sigma}^2\right) + 2 \cdot df_{approx} \\ BIC &= n \cdot ln\left(\tilde{\sigma}^2\right) + ln\left(n\right) \cdot df_{approx} \end{aligned} \tag{9}$$

Data transformation of GP-EDM

Because the mechanistic relationship between the stock abundance and escapement is left un-specified, we explore four data transformations for fishery-centric GP-EDM analyses: *Untransformed, Log-transform, Type-I log-difference transform,* and *Type-II log-difference transform*. Using the notation from (1–8), the first type is the original untransformed CPUE at time t (i.e. x_t) as the response variable. The second type utilizes natural logarithm transformed CPUE at time t [i.e. $ln(x_t)$] as response variable. The third type utilizes the first-order difference between the log-transformed CPUE at year-ahead time t and the log-transformed CPUE at time *t-1* [i.e. $ln(\frac{x_t}{x_{t-1}})$] as the response variable. Lastly, the fourth type utilizes the first-order difference between the log-transformed CPUE at year-ahead time t and the log-transformed CPUE subtracted from the annual fishery landings at time *t-1* [i.e. $ln(\frac{x_t}{x_{t-1}-qC_{t-1}})$]. Because these four data transformations produce different sets of model performance metrics at various scales, including LOO predictions, AIC, and BIC, they are not directly comparable. To ensure a consistent basis for model comparison, we transform the GP-EDM performance metrics to the same (logarithmic) scale. This approach facilitates the assessment and comparison of model performance across the different data transformation types.

Deriving MSY from GP-EDM

Having determined the transformation that yields the greatest prediction accuracy, we use the resulting GP-EDM model to estimate MSY (see schematic Fig. 1). To do so, we iterate the GP-EDM predictions over the next 40 years to estimate the long-run catch/landings given a particular constant harvest rate. The projection length of 40 years depends on computational power and convergence toward the steady state, we recommend >30 data points as a rule of thumb for computing the "long-run" average. Specifically, at each time step, we randomly simulate GP-EDM forecasts by sampling from a normal distribution with the mean and variance functions derived from (4) and use these forecasts as inputs for the next forecast. The harvest rate (represented by u_t at time t, bounded between 0 and 1) for each time step is kept constant

throughout the simulation. To account for uncertainty associated with the model, we repeat the procedure 1000 times to obtain the mean and standard error of the long-run yield estimate. We also repeat these steps for a range of harvest rates (u_t from 0 to 1) to determine the rate that maximizes the average long-run yield as our proxy for MSY.

Simulation testing for the robustness of MSY derived from GP-EDM

To assess the robustness of GP-EDM-based MSY to variation in underlying dynamics and exploitation history, we simulate different ecological scenarios under two different harvest histories. The ecological scenarios are a single-species system using a Pella–Tomlinson model and a predator–prey system (prey harvested) of Ricker type. Additionally, we use single-species Ricker models where dynamics are chaotic to assess the statistical robustness of EDM-based MSY under various assumptions about relationships between abundance index (CPUE) and true abundance. The harvest histories include increasing harvest rate (i.e. uni-directional one-way trip) and oscillating harvest rate. More details of the ecological scenarios and harvest histories are provided below.

Single-species dynamics

As the first example, we utilize parameter values derived from an empirical cross-stock meta-analysis (Thorson et al. 2012) to establish the Pella–Tomlinson model as follows:

$$B_{t+1} = (B_t + P_t - s_t B_t)\, e^{\varepsilon}$$
$$P_t = \left(\frac{a^{\frac{a}{a-1}}}{a-1}\right) r \left[\frac{B_t}{K} - \left(\frac{B_t}{K}\right)^a\right] \quad (10)$$

In the above equation (10), B_t represents the biomass of the species at time t, and P_t corresponds to the production at time t. The scalar s_t accounts for the impact of harvesting on species biomass, with $s_t B_t$ indicating the catch/landings quantity at time t. The growth parameter of the production function is denoted by r. The carrying capacity is represented by K. The parameter a, r, and K control the MSY. We assume that the process noise ε follows a normal distribution with a mean of zero and a standard deviation of 0.1. The parameters a and r are fixed at 1.478 and 0.404, respectively, which are the empirical average estimates derived from multiple stocks (refer to Table 1 and Table 2 in Thorson et al. 2012). Note that the (10) is valid when $a > 1$. For simplicity, we assume $K = 1$. Assuming constant harvest, we can express the steady-state relative biomass (B^*) and catch ($Yield^*$) as follows:

$$B^* = \sqrt[1-a]{\frac{ra^{a/a-1}}{-sa + s + ra^{a/a-1}}}$$
$$Yield^* = sB^* \quad (11)$$

Predator–prey dynamics

As a more challenging example, we consider a Ricker-type prey–predator model that exhibits reasonable fluctuations in biomass and catch/landings. The dynamics of the prey−predator system and harvesting are given by:

$$B_{1,t+1} = B_{1,t} e^{\left[r_1\left(1-B_{1,t}\right) - dB_{2,t} - s_t\right]} e^{\varepsilon_1}$$
$$B_{2,t+1} = B_{2,t} e^{\left[r_2\left(1-B_{2,t}\right) + dB_{1,t}\right]} e^{\varepsilon_2} \quad (12)$$

In the above equation (12), $B_{1,t}$ represents the biomass of the prey species (or forage species) and $B_{2,t}$ represents the biomass of the predator species at time t. The parameters r_1 and r_2 correspond to the intrinsic growth rates of the prey and predator, respectively. The scalar d represents the biomass transfer efficiency from prey to predator, and s_t is a scalar reflecting the harvesting effect on the prey at time t. We assume that the process noise ε_1 and ε_2 follow normal distributions with mean zero and standard deviations of 0.1, respectively. Heuristically, we set $r_1 = 2.2$ for the prey, which is higher than $r_2 = 1.8$ for the predator, and we set $d = 0.1$. Note that we use only prey biomass and catch/landings as our observational time series data for GP-EDM, meaning that we do not observe the predator. Further assuming constant harvest, we can express the steady-state relative biomass (B^*) and catch ($Yield^*$) for the prey species as follows:

$$B_1^* = \frac{r_2\left(r_1 - d - s\right)}{d^2 + r_1 r_2}$$
$$Yield^* \approx e^s B_1^* \quad (13)$$

Note that, in this model, the fixed point, B_1^*, and long-run average, $\overline{B_1} = \lim_{t\to\infty} \frac{1}{t}\sum_t B_{1,t}$, are equivalent under a constant harvest policy.

Chaotic Ricker-type dynamics

Because short-lived stocks may have higher intrinsic growth rates leading to quasi-cyclic or chaotic dynamics, we consider chaotic single-species Ricker models as simulation examples. Additionally, we consider three types of the relationship between abundance index (CPUE) and true abundance/biomass, including proportionality, hyper-stability, and hyper-depletion (Harley et al. 2001). The dynamics of the chaotic Ricker system and harvesting are given by

$$B_{t+1} = B_t e^{\left[r\left(1-\frac{B_t}{K}\right) - s_t\right]} e^{\varepsilon}$$
$$CPUE_t = qB_t^{\beta} \quad (14)$$

In the above equation (14), B_t represents the biomass of the species at time t, and r and K are intrinsic growth rate and carrying capacity, respectively. We assume $r = 4$ and $K = 1000$ such that higher growth rate produces chaotic dynamics. s_t is a scalar reflecting the harvesting effect. $CPUE_t$ represents an abundance index as a function of true abundance/biomass, such that $CPUE_t$ is proportional to B_t when q is a constant and $\beta = 1$. To simulate proportionality, we set $q = 0.9$ and $\beta = 1$. Additionally, we set $\beta = 0.8$ and $\beta = 1.2$ to represent the hyper-stability and hyper-depletion relationships (Harley et al. 2001). We assume that the process noise ε follows a normal distribution with a mean of zero and a standard deviation of 0.05. Further assuming constant harvest, we can express the steady-state (or long-run average) biomass (B^*) and catch ($Yield^*$) as follows:

$$B^* = K\left(1 - \frac{s}{r}\right)$$
$$Yield^* \approx e^s B^* \quad (15)$$

For all the single-species and predator–prey dynamics models, simulations are repeated 150 times for a duration of 100 years (time steps). For each 100-years simulated series, a randomized subsample is conducted to extract a continued 32-years window so that the simulated time series length is

comparable with GOM brown shrimp data. For each simulated series, we apply GP-EDM and MSY estimation. The simulated biomass (B_t) or abundance index ($CPUE_t$) and catch/landings time-series data are subjected to GP-EDM analysis, where model selection and data transformation procedures are performed to determine the best fitted GP-EDM, with varying embedding dimensions (E). The initial value of E ranges from 2 to 5. This E is before the ARD regularization of GP-EDM reduces the posterior dimension to match the number of effective time lags, as described by Munch et al. (2017), in each simulation. The best fitted model is subsequently employed to determine the MSY.

For comparison with the GP-EDM results, we use the known model parameters to compute the "true" analytical MSY using (11) and (13). Specifically, the "true" analytical MSY, given the model parameters, can be obtained by taking the first derivative of the steady-state yield function with respect to the harvest parameter "s" [specifically, MSY = 0.404 using (11), MSY = 1.361 using (13), and MSY = 5020.83 using (15)]. To test for the statistical robustness, the estimates of GP-EDM-based MSY is compared with the analytical MSY for both population dynamics models under two harvesting scenarios (i.e. unidirectional one-way trip harvest rate and fluctuating harvest rate over time) (Fig. S1).

Shrimp fishery management history and data in the US Gulf of Mexico

The National Marine Fisheries Service is mandated to provide annual, stock determination criteria (SDC) to the Gulf of Mexico Fishery Management Council (GMFMC) for Brown, White, and Pink Shrimp (*Farfantepenaeus aztecus*, *Litopenaeus setiferus*, and *Farfantepenaeus duorarum*). Since 2012, integrated stock synthesis models (i.e. age-structured assessment models), have been used to provide the annual, SDC for all three shrimp stocks. In 2019, a model review of these three, shrimp SS models revealed technical concerns (e.g. conflicting indices, convergence issues, residual patterns). Consequently, the GMFMC moved all three shrimp stocks into a Southeast Data, Assessment, and Review (SEDAR) research track process, which allows for the consideration of new data inputs and modeling approaches. A primary objective of the SEDAR research track process is to identify all available data inputs, limitations, and assumptions in order to select the most robust tool, for assessing these shrimp stocks. Preliminary findings from this SEDAR research track process have indicated that existing Brown, White, and Pink shrimp data limitations, such as the lack of recruitment and age-structured information, make advanced catch-at-age models inappropriate for assessing these stocks. Further, GOM shrimp stocks are considered annual crops, and so data processing and age-structured, model development timelines exceed their longevity—hindering the ability to provide timely management advice. Thus, it is necessary to consider next generation assessment models (such as index-based and reference point approaches) for these short-lived, shrimp stocks (Peterson and Walter 2023).

As a proof of concept tailored for a short-lived species fishery, our analyses focus on the Brown shrimp stock, which contributes to roughly half of the total annual Penaeid shrimp fishery landings each year (since 1984). To inform our study, we collated the most recent fishery landings and market size (i.e. pounds of shrimp, in 12 market size categories) data, as well as fishery-independent CPUE. The fishery-independent monitoring CPUE data were collected by the Southeast Area Monitoring and Assessment Program (SEAMAP) (1987–2019) (hereafter, SEAMAP CPUE). Traditionally, this collection of data was also the input into the shrimp stock synthesis models, along with life history parameters.

We use annual time series data of both Brown shrimp fishery landings and fishery-independent SEAMAP CPUE to examine reference points derived from EDM. Our analysis focuses on the GOM region, as it allows for comparisons to historical brown shrimp management practices. It is worth noting that the fishery landings data is only available at the GOM scale due to federal trip reporting limitations (i.e. no depth information is collected, only landings by statistical grids). We use the annual Brown Shrimp landings data, reported to National Marine Fisheries Service via state trip tickets (i.e. dealer reported landings, collated by all 5 US Gulf states). Fishery landings data are collected with 1–21 statistical zones that run approximately latitudinal along Florida and Texas and longitudinal throughout the rest of the GOM. Depth information was historically collected via the federal port agent surveys but is not reliably collected on state trip ticket forms, which are considered the more accurate source of recent total landings. For consistency, and to ensure a representative analysis—at the GOM-wide scale, we use the "averaged" SEAMAP CPUE time series as a proxy for annual Brown shrimp abundance dynamics. It is worth noting that a design-based estimator, e.g. weighted by the standard errors of data, might be used. However, in the present case the raw averages we used were nearly perfectly correlated with the design-based estimate ($r > 0.9$). Details of SEAMAP survey methodology and data description, including depth, season, and spatial zones, can be referred to (SEDAR87-RD-01, The SEAMAP Trawl Shrimp Data and Index Estimation Working Group). Statistical zones 1–21 (Nance 1992) were considered and then reduced based on the realized spatial distribution of Brown shrimp from SEAMAP trawl survey data. The statistical zones (8–21) used in this application are consistent with Tsai *et al.* (2023). Brown shrimp fishery-independent monitoring CPUE data are then averaged across all depths, seasons, and statistical zones for each year, resulting in yearly stock abundance dynamics for our analyses. Note that, it is the aggregated, instead of size- or age-disaggregated, time-series data that are used in our analyses.

GP-EDM analysis and simulation platform

All analyses and simulations of dynamics were performed using R Statistical Software (v4.3.1). Algorithms of fishery-centric GP-EDM framework were developed under R version 4.3.1 using packages: TMB (v1.9.6), yardstick (v1.2.0), and CaDENCE (v1.2.5). R codes are available at https://github.com/TsaiCH/Fishery-GPEDM.

Results

We found that, in general, the GP-EDM framework produced statistically robust and accurate estimates of MSY under various scenarios of dynamics, although estimate were much more precise for scenarios with fluctuating harvesting (Fig. 2). Out-of-sample LOO prediction skill metrics (i.e. RMSE, MAE, predictive Pearson correlation (ρ), and

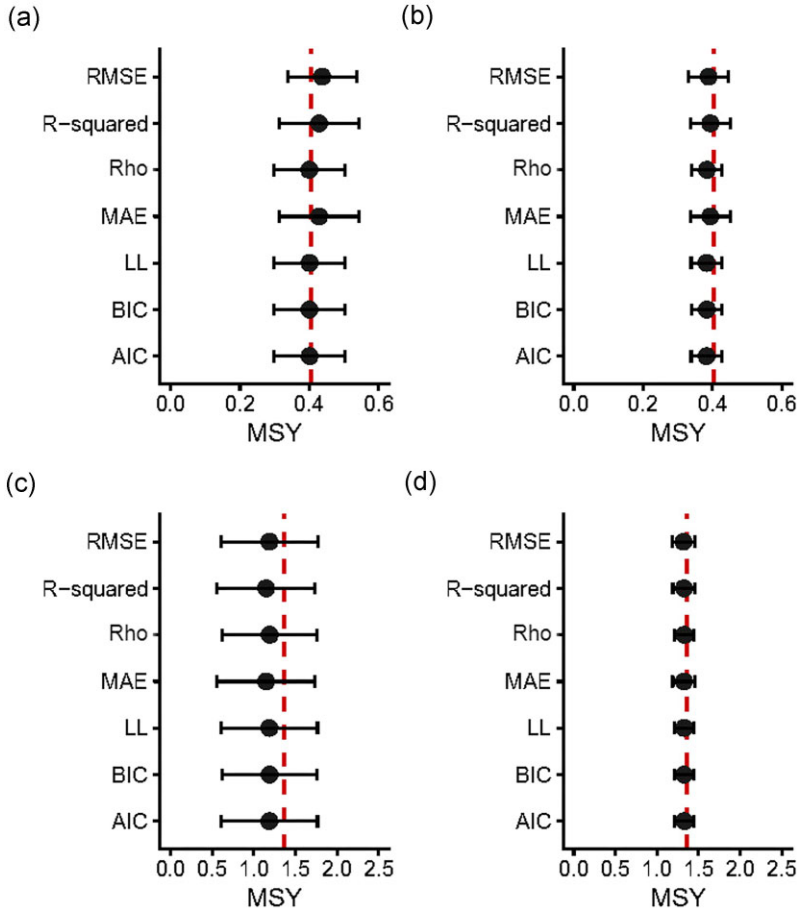


Figure 2. MSY estimates derived from the GP-EDM framework across various model selection methods and dynamics scenarios. (a) MSY estimates from the Pella–Tomlinson model with unidirectional harvesting. (b) MSY estimates from the Pella–Tomlinson model with fluctuating harvesting. (c) MSY estimates from the prey–predator dynamics with unidirectional harvesting. (d) MSY estimates from the prey–predator dynamics with fluctuating harvesting. The solid circle indicates the average MSY derived from 150 simulations, with the error bar denoting +/− 1 standard error of these estimates. The red dashed line represents the analytical "true" MSY for each dynamic scenario: 0.404 for the Pella–Tomlinson model and 1.361 for the prey–predator model. For model selection, out-of-sample prediction skill metrics (leave-one-out, LOO) include root mean square error (RMSE), mean absolute error (MAE), predictive Pearson correlation (ρ), and R^2. Information-related metrics include Akaike information criteria (AIC), Bayesian information criteria (BIC), and log-likelihood (LL). Notice the differences in the horizontal axis scale.

predictive R^2), information criteria (i.e. AIC and BIC), and maximum log-likelihood function (LL) all performed similarly to select the best fitted GP-EDM for estimating MSY. All model selection methods produced a robust averaged MSY with +/− 1 standard error intervals that covered the true MSY, regardless of embedding dimension and data transformation (Fig. 2).

Models chosen based on out-of-sample forecasting metrics (RMSE, MAE, ρ, and R^2) and information-based parsimony, on average, yielded similar accuracy for MSY (Fig. 2). However, GP-EDM tended to reduce the estimation error for MSY under fluctuating harvesting scenarios compared with unidirectional harvesting scenarios, i.e. the one-way trip harvest rate (cf. Fig. 2 left panels versus right panels). Additionally, EDM MSY estimates are robust to the assumptions about chaotic dynamics and assumptions about proportionality and hyper-stability (Fig. S2a, b, d, e), but potentially sensitive to hyper-depletion with mild conditions ($\beta = 1.2$) (Fig. S2c and f). EDM MSY tended to have lower standard errors under fluctuating harvesting scenarios than those under unidirectional harvesting scenarios, regardless of assumptions about chaotic dynamics or proportionality of abundance index (Fig. S2).

In applying the GP-EDM framework to the Gulf of Mexico Brown shrimp landings and fishery-independent CPUE data, we observed a consistency in model selection. Both prediction skill metrics and information criteria favored the same model, characterized by an optimal embedding dimension $E = 4$ and a Type-I log difference transformation (refer to Table 1). This specific model, which we refer to as "ModelE4T1," yielded the most accurate out-of-sample CPUE forecasts (LOO $R^2 = 0.71$) and exhibited the smallest out-of-sample prediction errors (LOO RMSE = 0.206 and LOO MAE = 0.163) as depicted in Fig. 3. Such metrics are crucial for generating statistically robust MSY estimates, as shown in Figs 2 and 3. Given the unanimous agreement across all model selection metrics, we designated ModelE4T1 as the primary candidate for MSY estimation.

The LOO prediction skill exhibited a noteworthy trend: it improved as the embedding dimension (E) increased. Notably, it achieved peak accuracy at $E = 4$ (LOO $R^2 = 0.71$), a stark contrast to its lowest at $E = 1$ (LOO $R^2 = 0.1$), which

Table 1. Model selection using GP-EDM for Brown Shrimp dynamics in the US Gulf of Mexico. The top-performing model, highlighted in bold, is designated as "ModelE4T1." It features an embedding dimension of $E = 4$ and utilizes a Type-I log-difference transformation

	LL	AIC	BIC	MAE	RMSE	ρ	R^2	MSY	SE(MSY)
$E = 5$									
Non-transform	−160.19	322.99	324.7	0.318	0.391	0.57	−0.06	n.a.	n.a.
Log-transform	−151.52	309.39	313.6	0.241	0.281	0.68	0.44	n.a.	n.a.
Type-I log-difference	74.62	−95.02	−58.89	0.175	0.212	0.83	0.68	1373.15	41.74
Type-II log-difference	26.45	−1.83	32.18	0.314	0.381	0.66	−0.01	79 895	17 152
$E = 4$									
Non-transform	−165.65	333.95	335.8	0.309	0.384	0.56	−0.05	n.a.	n.a.
Log-transform	−157.74	321.01	324.8	0.241	0.284	0.66	0.42	n.a.	n.a.
Type-I log-difference	**77.99**	**−100.21**	**−62.06**	**0.163**	**0.206**	**0.84**	**0.71**	**225.35**	**3.52**
Type-II log-difference	25.41	−6.76	23.34	0.203	0.241	0.77	0.58	54 233	26 734
$E = 3$									
Non-transform	−174.34	350.62	352	0.345	0.418	0.01	−0.29	n.a.	n.a.
Log-transform	−167.99	339.32	341.7	0.272	0.348	0.33	0.11	n.a.	n.a.
Type-I log-difference	19.95	2.29	31.86	0.254	0.322	0.59	0.23	789.04	22.18
Type-II log-difference	−7.93	28.06	36.61	0.257	0.339	0.46	0.15	251 993	15 364

E is embedding dimension. LL is log marginal likelihood as defined in (2). Out-of-sample (leave-one-out, LOO) prediction skill metrics include MAE, RMSE, ρ, and R^2, which are mean absolute error, root mean square error, and Pearson correlation, and predictive R^2, respectively defined in (5) and (6). AIC and BIC are Akaike information criteria and Bayesian information criteria as defined in (9). MSY and SE(MSY) are maximum sustainable yield and the corresponding standard error. "n.a." represents numerical instability w/o convergence.

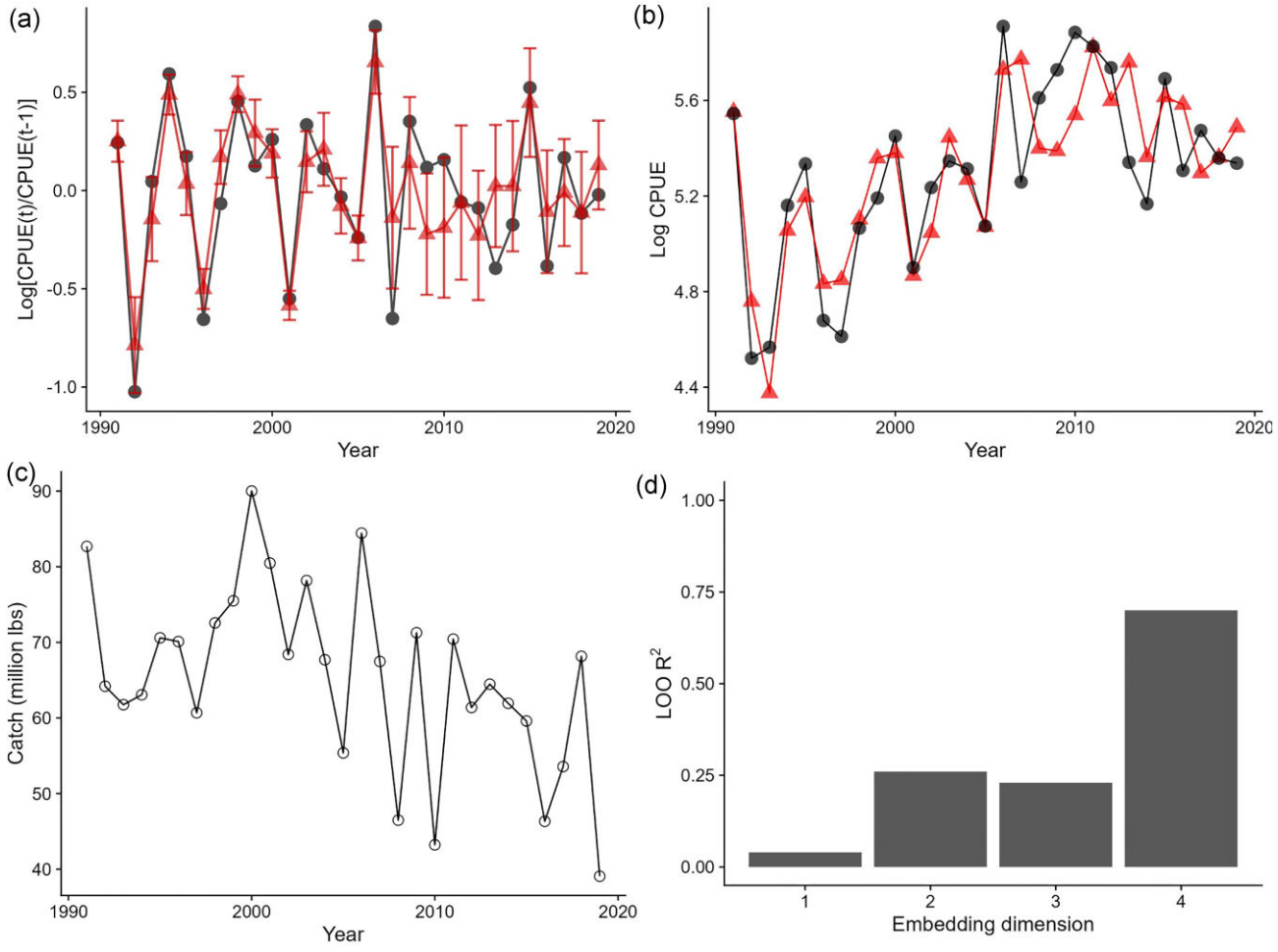


Figure 3. Analysis and predictions of Brown Shrimp stock abundance dynamics in the US Gulf of Mexico. (a) Prediction mean +/− 1 standard error (red triangle and error bar) versus observation (black circle) for Brown Shrimp. Predictions are derived from the same model favored by both out-of-sample (leave-one-out, LOO) prediction skill metrics and information criteria, characterized by an optimal embedding dimension $E = 4$ and a Type-I log difference transformation (Table 1). (b) A log-scaled comparison of observed versus GP-EDM predicted catch per unit effort (CPUE) for Brown shrimp (LOO $R^2 = 0.71$; Table 1). (c) Fishery landings for Brown Shrimp, represented in million pounds. This data, combined with CPUE, serves as input for the GP-EDM framework tailored to fisheries. (d) A depiction of how GP-EDM's out-of-sample predictive accuracy correlates with the number of time-delay vectors [referenced in (1) and (2)].

is similar to traditional single-species production models (refer to Fig. 3c and schematic Fig. 1). The catchability factor, denoted as q, which bridges CPUE and fishery landings [as per (2)], was estimated at 1.18, with a standard error ranging from 1.06 to 1.18. The parameter values estimated by ModelE4T1 for Brown shrimp can be found in Fig. S3.

Moreover, our analysis revealed marginal enhancements in the GP-EDM's out-of-sample predictive accuracy for Brown shrimp when pivotal environmental variables were incorporated (Tables S1 and S2). These variables were previously thought to be important drivers of annual Brown shrimp abundance. The inclusion of time lags for both the recruitment and rainfall indices within the GP-EDM framework did not offer any substantial boost to the prediction skill, especially when juxtaposed with ModelE4T1, which only considered lags of landings and CPUE data.

From the fits of ModelE4T1, we deduced the GP-EDM MSY for the Gulf of Mexico's Brown shrimp (Fig. 4). The long-term average CPUE at MSY stood roughly at 376.94 $\pm$ 5.64 tails per tow (Fig. 4a), with the MSY roughly at 225.35 $\pm$ 3.52 million pounds or 102.22 $\pm$ 1.60 million kilograms of tails (Fig. 4b). The harvest rate at MSY was approximated to be 0.72, such that the biomass at MSY (BMSY) amounted to 312.98 million pounds or 141.97 million kilograms of tails.

Discussion

Our simulations demonstrate that GP-EDM can yield robust estimates of MSY that closely approximate the theoretical MSY values. This holds true not only for single-species fishery dynamics, where potential species interactions are treated as environmental (or process) noise, but also for multi-species systems, where species interactions are implicitly modeled with lags in GP-EDM (Fig. 2). Furthermore, unlike traditional parametric production models, GP-EDM displays a notable resilience in scenarios with chaotic dynamics under unidirectional (one-way trip) dynamics (Fig. S2). This observation aligns with prior findings (Brias and Munch 2021, Giron-Nava et al. 2021), suggesting that non-parametric approaches may offer greater robustness in the face of dynamic history of harvesting, compared to traditional parametric methods where statistical identifiability issues often hinder the accurate estimation of parameter aggregates like MSY.

Recently, Boettiger (2022) identified a "forecast trap" where using prediction accuracy to choose among a small handful of parametric models led to significantly worse management benchmarks. In contrast, our results indicate that using out-of-sample (LOO) prediction skill for model selection is entirely satisfactory and better than other selection criteria (Figs 2, S2). Careful consideration of the issue of forecast trap suggests that it arises from choosing among several candidate models, none of which can reconstruct the true mechanistic dynamics (Paniw et al. 2023). On the other hand, the non-parametric approach of GP-EDM, provided with sufficient data, can reconstruct the underlying dynamics in a wide range of systems (Rogers and Munch 2020, Brias and Munch 2021, Tsai et al. 2023). We hypothesize that this flexibility circumvents the forecast trap, though more detailed analysis is warranted.

While our findings hold promise, it is crucial to acknowledge several important caveats. First, the efficacy of EDM and any resulting management recommendations depends on the availability of data. Meta-analyses suggest that time series should be several multiples of the generation time to provide reasonable prediction accuracy (Munch et al. 2018). Second, if there is relatively little variation in historical dynamics between fishing and stock abundance (i.e. variations in catch/landings subtracted from CPUE), estimates of MSY may be accompanied by increased uncertainty. In situations where the generation time is long or the historical variation in fishing and stock abundance dynamics is limited, resorting to traditional parametric approaches, such as stock synthesis and data-limited methods, may help (Methot and Wetzel 2013, Chong et al. 2020, Pons et al. 2020, Legault et al. 2023). However, we caution that the accuracy of extrapolation from traditional parametric approaches depends heavily on structural assumptions which may be difficult to justify (Thorson et al. 2014). Although beyond the scope of this work, additional simulations with a variety of models and time series lengths may help clarify conditions under which GP-EDM or traditional assessments are preferable. Finally, we note that the present analyses have focused solely on constant harvest rate policies. These are likely to be sub-optimal in cyclical fisheries—overharvesting when biomass is low and underharvesting when it is high. In these situations, a state-dependent policy constructed using EDM (Brias and Munch 2021, 2024) may be a significant improvement.

Several extensions to the current framework may enhance utility of GP-EDM. Here, in keeping with the shrimp case study (Tsai et al. 2023), we used landings and abundance indices to estimate MSY. However, with appropriate algebraic adjustments the GP-EDM framework can readily use landings and some proxy for fishing effort to estimate MSY. Indeed, in keeping with the traditional assumption that catch is a function of abundance and effort, GP-EDM only requires any two measures among catch, effort, and CPUE to generate predictions. Given the greater availability of fishery-dependent CPUE compared to fishery-independent survey data, we propose that our framework can be readily expanded to encompass numerous other fisheries practices. However, we caution that our methodology has certain imprecisions, leading to wide uncertainty ranges in the MSY estimates (Fig. 2). This approach may also fail in cases of strong hyperdepletion, where assumptions about the proportionality of the abundance index can introduce bias in the MSY estimates (Fig. S2). Future research should take these limitations into account. One potential solution might be to relax the assumed relationship between catch, effort, and biomass in a non-parametric way or to expand the pool of candidate models. In addition, age-structured EDM (Dolan et al. 2023) and spatial EDM (Johnson et al. 2021) leverage multiple short series to improve prediction accuracy. We expect that age-structured EDM could improve MSY estimates for longer-lived species. Additionally, since spatial EDM can account for spatial heterogeneity, it may be useful for setting spatial management targets. Given Brown shrimp are more susceptible to environmental factors at early ages, while resident in semi-enclosed, estuarine waters (Schlenker et al. 2023), future work may consider the applicability of spatial EDM in further improving Gulf shrimp fishery management.

Turning to the Gulf of Mexico Brown shrimp case study, we find that the predictive capability of GP-EDM utilizing time lags significantly outperforms that of the single time lag model (Fig. 3, Table 1). Particularly, we find that four lags of escapement [i.e. catch/landings subtracted from CPUE as

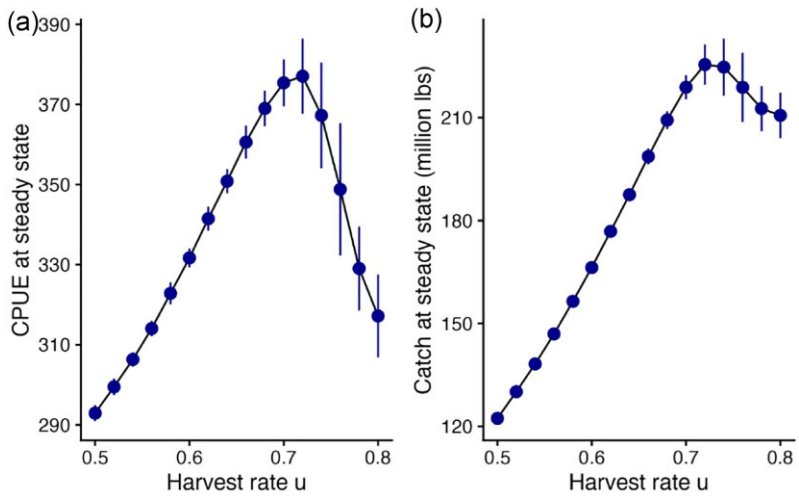


Figure 4. GP-EDM MSY for Gulf of Mexico's Brown Shrimp. The optimal GP-EDM prediction (with an LOO $R^2 = 0.71$ as detailed in Table 1) provides the foundation for determining reference points. (a) The relationship between CPUE and the harvest rate at steady state. The GP-EDM's projected long-term average CPUE at MSY is approximately 376.94 +/− 5.64 tails per tow. (b) The relationship between catch (or landings) and harvest rate at steady state. The GP-EDM MSY is approximately 225.35 +/− 3.52 million pounds or 102.22 ± 1.60 million kilograms of tails. The solid circle indicates the average estimation, while the bar shows a range of two standard errors for these estimates.

per (2)] provides the best prediction accuracy for year-ahead Brown shrimp CPUE (Table 1). This underscores the importance of incorporating historical data on both fishery operations and fishery-independent abundance proxies to account for hidden state-variables beyond the current stock size (e.g. considering unobserved species interactions and environmental feedbacks; Fujiwara et al. 2016, Schlenker et al. 2023), which is also a key characteristic of our new fishery-centric GP-EDM framework (Fig. 1). Additionally, it enables a more complicated production function or yield curve compared with the typically considered parametric family (Fig. S4). Specifically, the expected Brown shrimp CPUE is not a monotonic function of the first lag of escapement, and seems to depend on the longer time lags (e.g. the second time lags) that might account for unobserved species and environmental feedbacks (Fig. S4).

The GP-EDM MSY for brown shrimp (circa 225 million pounds or 102 million kilograms of tails; Fig. 4), appears to have a higher value, but within the same order of magnitude, compared to previous estimates obtained from production models (circa 88 million pounds or 40 million kilograms of tails) and stock synthesis models (circa 147 million pounds or 67 million kilograms of tails). Note that, GP-EDM MSY and other previous MSY estimates are all above the current catch level for Brown shrimp, which is reasonable given that Penaeid shrimp catch in the Gulf of Mexico has only been limited by total fishing effort due to the bycatch mortality on other species with longer life span. Based on our simulation results, we expect that the derived GP-EDM MSY value for Brown shrimp is at least as robust as that derived from traditional assessments (Peterson and Walter 2023). However, we caution that EDM MSY derived from real data may be sensitive to data transformation and model complexity (e.g. effective embedding dimension), where numerical instability may arise from model candidates lacking sufficient prediction skills (Table 1). We suggest that a representative model chosen based on both prediction skill and parsimony, as well as expert knowledge, are important considerations for a consensus EDM MSY.

Interestingly, we found that augmenting our analysis with time series for recruitment and rainfall did not significantly improve prediction accuracy (Tables S1 and S2). This aligns with our earlier study (Tsai et al. 2023) where we observed that environmental factors, including salinity, dissolved oxygen, and temperature, have limited impact on GP-EDM prediction accuracy. As these drivers have clear ecological relevance (Schlenker et al. 2023), it is important to consider why they do not improve forecast accuracy (Peterson and Walter 2023). One potential explanation is that these variables are not measured in the optimal location or time of year or with enough precision to be most informative. Alternatively, we note that Granger-style causality arguments, in which including causal drivers improves time series prediction, do not work with time delay embedding (Sugihara et al. 2012). Thus, a second, non-exclusive, possibility is that these drivers are indeed important but that their influence is already captured in the lagged patterns of CPUE and catch/landings (Munch et al. 2023, Tsai et al. 2023).

In summary, we advocate for adopting MSY estimates derived from GP-EDM as a robust framework for establishing fishery reference points that can effectively account for unobserved species and environmental feedbacks and underlying historical exploitation dynamics. This positions GP-EDM as a reliable scientific foundation to re-evaluate current methodology for estimating stock status (i.e. SDC), particularly crucial for addressing the inadequacies observed in managing short-lived Penaeid shrimp species in the Gulf of Mexico (Peterson and Walter 2023). Furthermore, while GP-EDM does require reasonably extensive time series, the data demands are notably more manageable when compared with those of conventional data-rich assessment methods. This suggests that GP-EDM could serve as a practical alternative for other data-moderate assessments, especially in cases where species exhibit short lifespans and their true population dynamics are influenced by substantial abundance fluctuations stemming from unobserved species interactions and complex ecosystem feedbacks (Glaser et al. 2014, Fogarty et al. 2016, Munch et al. 2018, 2020).

Acknowledgements

We thank Lewis Coggins and Bethany Johnson for helping with preliminary analyses. We would like to thank the anonymous reviewers and NOAA's internal reviewers for their valuable comments and suggestions, which greatly improved the quality of this manuscript. This study and the development of fishery-centric GP-EDM framework were supported by the Lenfest Oceans Program.

Author contributions

C.-H.T., S.B.M., M.D.M., and M.H.S. devised the research program; C.-H.T. and S.B.M. performed analysis with help from M.D.M. and M.H.S; C.-H.T. and S.B.M. wrote the first draft of manuscript; and all authors were involved in interpreting the results and contributed to the final draft of manuscript.

Supplementary data

Supplementary data is available at *ICES Journal of Marine Science* online.

Conflict of interest: The authors declare no competing interests.

Data availability

R code and data required for reproducing analyses are available at https://github.com/TsaiCH/Fishery-GPEDM. Data generated or analyzed in this study can be obtained from the authors, MM and MS, upon a reasonable request and with SEAMAP's permission.

References

Arkhipkin AI, Hendrickson LC, Payá I *et al.* Stock assessment and management of cephalopods: advances and challenges for short-lived fishery resources. *ICES J Mar Sci 2021*;**78**:714–30. https://doi.org/10.1093/icesjms/fsaa038 (27 September 2023, date last accessed).

Boettiger C. The forecast trap. *Ecol Lett 2022*;**25**:1655–64. https://doi.org/10.1111/ele.14024

Brias A, Munch SB. Ecosystem based multi-species management using Empirical Dynamic Programming. *Ecol Model 2021*;**441**:109423. https://doi.org/10.1016/j.ecolmodel.2020.109423

Chang C, Ushio M, Hsieh C. Empirical dynamic modeling for beginners. *Ecol Res* 2017;**32**:785–96. https://doi.org/10.1007/s11284-017-1469-9

Chong L, Mildenberger TK, Rudd MB *et al.* Performance evaluation of data-limited, length-based stock assessment methods. *ICES J Mar Sci 2020*;**77**:97–108. https://doi.org/10.1093/icesjms/fsz212

Cleveland WS, Grosse E. Computational methods for local regression. *Stat Comput 1991*;**1**:47–62. https://doi.org/10.1007/BF01890836

Deyle ER, Fogarty M, Hsieh C *et al.* Predicting climate effects on Pacific sardine. *Proc Natl Acad Sci 2013*;**110**:6430–5. https://doi.org/10.1073/pnas.1215506110

Deyle ER, Schueller AM, Ye H *et al.* Ecosystem-based forecasts of recruitment in two menhaden species. *Fish Fish 2018*;**19**:769–81. https://doi.org/10.1111/faf.12287

Dolan TE, Palkovacs EP, Rogers TL *et al.* Age structure augments the predictive power of time series for fisheries and conservation. *Can J Fish Aquat Sci 2023*;**80**:795–807. https://doi.org/10.1139/cjfas-2022-0219

Engelhard GH, Peck MA, Rindorf A *et al.* Forage fish, their fisheries, and their predators: who drives whom? *ICES J Mar Sci 2014*;**71**:90–104. https://doi.org/10.1093/icesjms/fst087

Essington TE, Moriarty PE, Froehlich HE *et al.* Fishing amplifies forage fish population collapses. *Proc Natl Acad Sci 2015*;**112**:6648–52. https://doi.org/10.1073/pnas.1422020112

Fogarty MJ, Gamble R, Perretti CT. Dynamic complexity in exploited marine ecosystems. *Front Ecol Evol 2016*;**4**. http://journal.frontiersin.org/Article/10.3389/fevo.2016.00068/abstract (27 September 2023, date last accessed).

Fujiwara M, Zhou C, Acres C *et al.* Interaction between penaeid shrimp and fish populations in the Gulf of Mexico: importance of shrimp as forage species. *PLoS One 2016*;**11**:e0166479. https://doi.org/10.1371/journal.pone.0166479

Garcia SP, DeLancey LB, Almeida JS *et al.* Ecoforecasting in real time for commercial fisheries: the Atlantic white shrimp as a case study. *Mar Biol 2007*;**152**:15–24. https://doi.org/10.1007/s00227-007-0622-3

Giron-Nava A, Ezcurra E, Brias A *et al.* Environmental variability and fishing effects on the Pacific sardine fisheries in the Gulf of California. *Can J Fish Aquat Sci 2021*;**78**:623–30. https://doi.org/10.1139/cjfas-2020-0010

Glaser SM, Fogarty MJ, Liu H *et al.* Complex dynamics may limit prediction in marine fisheries. *Fish Fish 2014*;**15**:616–33. https://doi.org/10.1111/faf.12037

Glaser SM, Ye H, Maunder M *et al.* Detecting and forecasting complex nonlinear dynamics in spatially structured catch-per-unit-effort time series for North Pacific albacore (Thunnus alalunga). *Can J Fish Aquat Sci 2011*;**68**:400–12. https://doi.org/10.1139/F10-160

Harley SJ, Myers RA, Dunn A. Is catch-per-unit-effort proportional to abundance? *Can J Fish Aquat Sci 2001*;**58**:1760–72. https://doi.org/10.1139/f01-112

Hsieh C, Reiss CS, Hunter JR *et al.* Fishing elevates variability in the abundance of exploited species. *Nature 2006*;**443**:859–62. https://doi.org/10.1038/nature05232

Johnson B, Gomez M, Munch SB. Leveraging spatial information to forecast nonlinear ecological dynamics. *Methods Ecol Evol 2021*;**12**:266–79. https://doi.org/10.1111/2041-210X.13511

Johnson B, Munch SB. An empirical dynamic modeling framework for missing or irregular samples. *Ecol Model 2022*;**468**:109948. https://doi.org/10.1016/j.ecolmodel.2022.109948

Krämer N, Sugiyama M. The degrees of freedom of partial least squares regression. *J Am Statist Assoc 2011*;**106**:697–705. https://doi.org/10.1198/jasa.2011.tm10107

Legault CM, Wiedenmann J, Deroba JJ *et al.* Data-rich but model-resistant: an evaluation of data-limited methods to manage fisheries with failed age-based stock assessments. *Can J Fish Aquat Sci 2023*;**80**:27–42. https://doi.org/10.1139/cjfas-2022-0045

Li C, Liu H. Comparative ecosystem modelling of dynamics and stability of subtropical estuaries under external perturbations in the Gulf of Mexico. *ICES J Mar Sci 2023*;**80**:1303–18. https://doi.org/10.1093/icesjms/fsad056

Lindegren M, Checkley DM, Rouyer T *et al.* Climate, fishing, and fluctuations of sardine and anchovy in the California Current. *Proc Natl Acad Sci 2013*;**110**:13672–7. https://doi.org/10.1073/pnas.1305733110

Liu H, Fogarty M, Glaser S *et al.* Nonlinear dynamic features and co-predictability of the Georges Bank fish community. *Mar Ecol Prog Ser 2012*;**464**:195–207. https://doi.org/10.3354/meps09868

Masi MD, Ainsworth CH, Kaplan IC *et al.* Interspecific interactions may influence reef fish management strategies in the Gulf of Mexico. *Mar Coast Fish 2018*;**10**:24–39. https://doi.org/10.1002/mcf2.10001

Methot RD, Wetzel CR. Stock synthesis: a biological and statistical framework for fish stock assessment and fishery management. *Fish Res 2013*;**142**:86–99. https://doi.org/10.1016/j.fishres.2012.10.012

Munch SB, Brias A, Sugihara G *et al.* Frequently asked questions about nonlinear dynamics and empirical dynamic modelling. *ICES J Mar Sci 2020*;**77**:1463–79. https://doi.org/10.1093/icesjms/fsz209

Munch SB, Brias A. Empirical dynamic programming for model-free ecosystem-based management. *Methods Ecol Evol 2024*;**15**:769–78. https://doi.org/10.1111/2041-210X.14302

Munch SB, Giron-Nava A, Sugihara G. Nonlinear dynamics and noise in fisheries recruitment: a global meta-analysis. *Fish Fish 2018*;**19**:964–73. https://doi.org/10.1111/faf.12304

Munch SB, Poynor V, Arriaza JL. Circumventing structural uncertainty: a Bayesian perspective on nonlinear forecasting for ecology. *Ecol Complex 2017*;**32**:134–43. https://doi.org/10.1016/j.ecocom.2016.08.006

Munch SB, Rogers TL, Johnson BJ *et al.* Rethinking the prevalence and relevance of chaos in ecology. *Annu Rev Ecol Evol Syst 2022*;**53**:227–49. https://doi.org/10.1146/annurev-ecolsys-111320-052920

Munch SB, Rogers TL, Sugihara G. Recent developments in empirical dynamic modelling. *Methods Ecol Evol 2023*;**14**:732–45. https://doi.org/10.1111/2041-210X.13983

Nance JM. Estimation of effort for the Gulf of Mexico shrimp fishery. NOAA Technical Memorandum, NMFS-SEFSC-300, 12 pp, 1992.

Neal RM. Priors for infinite networks. In: *Bayesian Learning for Neural Networks*. New York, NY: Springer, 1996.

Paniw M, García-Callejas D, Lloret F *et al.* Pathways to global-change effects on biodiversity: new opportunities for dynamically forecasting demography and species interactions. *Proc R Soc B Biol Sci 2023*;**290**:20221494. https://doi.org/10.1098/rspb.2022.1494

Peterson CD, Walter JF. *Southeast fisheries science center management strategy evaluation strategic plan*. NOAA Tech. Memo. NMFS-SEFSC-TM-766, 27pp, 2023.

Pinsky ML, Byler D. Fishing, fast growth and climate variability increase the risk of collapse. *Proc R Soc B Biol Sci 2015*;**282**:20151053. https://doi.org/10.1098/rspb.2015.1053

Pons M, Cope JM, Kell LT. Comparing performance of catch-based and length-based stock assessment methods in data-limited fisheries. *Can J Fish Aquat Sci 2020*;**77**:1026–37. https://doi.org/10.1139/cjfas-2019-0276

Rogers TL, Munch SB. Hidden similarities in the dynamics of a weakly synchronous marine metapopulation. *Proc Natl Acad Sci 2020*;**117**:479–85. https://doi.org/10.1073/pnas.1910964117

Schlenker LS, Stewart C, Rock J *et al.* Environmental and climate variability drive population size of annual penaeid shrimp in a large lagoonal estuary. *PLoS One 2023*;**18**:e0285498. https://doi.org/10.1371/journal.pone.0285498

Siple MC, Koehn LE, Johnson KF *et al.* Considerations for management strategy evaluation for small pelagic fishes. *Fish Fish 2021*;**22**:1167–86. https://doi.org/10.1111/faf.12579

Sugihara G, May R, Ye H *et al.* Detecting causality in complex ecosystems. *Science 2012*;**338**:496–500. https://doi.org/10.1126/science.1227079

Takens F. Detecting strange attractors in turbulence. In: *Dynamical Systems and Turbulence, Warwick 1980*. Berlin, Heidelberg: Springer, 1981, 366–81.

Thorson JT, Cope JM, Branch TA *et al.* Spawning biomass reference points for exploited marine fishes, incorporating taxonomic and body size information. *Can J Fish Aquat Sci 2012*;**69**:1556–68. https://doi.org/10.1139/f2012-077

Thorson JT, Ono K, Munch SB. A Bayesian approach to identifying and compensating for model misspecification in population models. *Ecology 2014*;**95**:329–41. https://doi.org/10.1890/13-0187.1

Tsai C-H, Munch SB, Masi MD *et al.* Predicting nonlinear dynamics of short-lived penaeid shrimp species in the Gulf of Mexico. *Can J Fish Aquat Sci 2023*;**80**:57–68. https://doi.org/10.1139/cjfas-2022-0029

Ushio M, Hsieh C, Masuda R *et al.* Fluctuating interaction network and time-varying stability of a natural fish community. *Nature 2018*;**554**:360–3. https://doi.org/10.1038/nature25504

Vaida F, Blanchard S. Conditional Akaike information for mixed effects models. *Biometrika 2005*;**92**:351–70.

Ye H, Beamish RJ, Glaser SM *et al.* Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proc Natl Acad Sci 2015*;**112**:E1569–76. https://pnas.org/doi/full/10.1073/pnas.1417063112 (27 September 2023, date last accessed).

Handling Editor: Ruben Roa-Ureta