Received: 10 July 2018 | Accepted: 2 January 2019

DOI: 10.1111/2041-210X.13150

RESEARCH ARTICLE

Methods in Ecology and Evolution BRITISH ECOLOGICAL SOCIETY

Regularized S-map for inference and forecasting with noisy ecological time series

Simone Cenci[1] | **George Sugihara**[2] | **Serguei Saavedra**[1]

[1]Department of Civil and Environmental Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts

[2]Scripps Institution of Oceanography, University of California San Diego, La Jolla, California

Correspondence
Serguei Saavedra
Email: sersaa@mit.edu

Funding information
Department of Defense SERDP 15 RC-2509, National Science Foundation, Grant/Award Number: NSF-DEB-1020372 and NSF-ABI-Innovation DBI- 1667584, McQuown Chair, UCSD (GS)

Handling Editor: Robert B. O'Hara

Abstract

1. It is well known that fluctuations of species abundances observed in ecological time series emerge from an interplay between deterministic nonlinear dynamics and stochastic forces. Importantly, nonlinearity and stochasticity introduce significant challenges to the analysis of ecological time series, such as the inference of the effect of species interactions on community dynamics and forecasting of species abundances.
2. Local linear fits with state-space-dependent kernel functions, known as S-maps, provide an efficient method to infer Jacobian coefficients (a proxy for the local effect of species interactions) and to make reliable forecasts from nonlinear time series. Yet, while it has been shown that the S-map outperforms existing methods for nonparametric inference and forecasting, the methodology is sensitive to process noise. To overcome this limitation, we integrate the S-map with different regularization schemes.
3. To validate our approach, we test our methodology against different levels of noise and nonlinearity using three standard population dynamics models. We show that an appropriate choice of the regularization scheme, alongside an accurate choice of the kernel functions, can significantly improve the in-sample inference of Jacobian coefficients and the out-of-sample forecast of species abundances in the presence of process noise. We further validate our methodology using two empirical time series of marine microbial communities.
4. Our results illustrate that the regularized S-map is an efficient method for nonparametric inference and forecasting from noisy, nonlinear, ecological time series. Yet, attention must be paid on the regularization scheme and the structure of the kernel for whether inference or forecasting is the ultimate goal of a research study.

KEYWORDS
nonlinear time series, out-of-sample forecast, parameter inference, process noise, regularization, S-map

1 | INTRODUCTION

It is well known that purely deterministic ecological dynamics are mathematical idealizations (Bjørnstad & Grenfell, 2001; Kantz & Schreiber, 2004). Stochastic processes affect populations at all scales, from random realizations of birth and death processes to random environmental variations (Cenci & Saavedra, 2018a; Frentz, Kuehn, & Leibler, 2015; Ranta, Lundberg, Kaitala, & Laakso, 2000;

Spanio, Hidalgo, & Muñoz, 2017). It has been broadly recognized that fluctuations of species abundances observed in ecological time series emerge from the interplay between nonlinear (density-dependent) deterministic dynamics and stochastic forces (Bjørnstad, Fromentin, Stenseth, & Gjøsater, 1999; Coulson, Rohani, & Pascual, 2004; Dixon, Milicich, & Sugihara, 1999; Dobrinevski & Frey, 2012; Higgins, Hastings, Sarvela, & Botsford, 1997; Mutshinda, O'Hara, & Woiwod, 2009), plus unavoidable noise due to measurement errors (Coulson et al., 2004; Mason, Holdaway, & Richardson, 2018; Sugihara & May, 1990).

Intrinsic stochasticity such as demographic stochasticity or, more generally, any type of process noise can induce new dynamics that cannot be predicted from a deterministic analysis: from the reversal of deterministic predictions (Constable, Rogers, McKane, & Tarnita, 2016) to the emergence of alternative stable states (Schooler, Salau, Julien, & Ives, 2011). Furthermore, noise has important implications on the statistical analysis of time-series data. For example, for many inference techniques, process noise can set qualitative lower bounds to the number of data points needed to fully recover the parameters. That is, the number of neighbouring points required in order to fit the real dynamics (rather than just the noise) grows with the level of noise (Bradley & Kantz, 2015). Similarly, process noise can reduce the out-of-sample forecasting skill of learning algorithms by, for example, breaking down the self-similarity of strange attractors in chaotic systems (Kantz & Schreiber, 2004), or, more generally, by increasing the risk of overfitting the training data (Abu-Mostafa, Magdon-Ismail, & Lin, 2012; Clark et al., 2001). Thus, to improve inductive analyses of ecological dynamics, it is necessary to develop robust methodologies that can allow us to analyse ecological time-series data under the presence of process noise.

Two important analyses performed on ecological time series are the inference of the effect of species interactions on community dynamics and the forecast of species abundances. Inferring accurate effects of species interactions from time-series data is important as a first step to find answers to fundamental questions in ecology such as: what is the role of species interactions during community assembly (Enke, Leventhal, Metzger, Saavedra, & Cordero, 2018)? How does the effect of species interactions on community dynamics evolve in changing environments (Cenci, Montero-Castaño, & Saavedra, 2018; Lawrence et al., 2012)? Does interaction variability in changing environments have any impact on species co-evolution (Barraclough, 2015)? Equivalently, a good forecast of changes in species abundances is key to improve the planning and decision-making process of ecosystem management (Clark et al., 2001), and to design successful socio-economic developmental strategies (Burkov, Novikov, & Shchepkin, 2015). However, inference and forecasting from noisy, nonlinear, time series is notoriously challenging (Bjørnstad & Grenfell, 2001; Casdagli, Eubank, Farmer, & Gibson, 1991; Perretti, Munch, & Sugihara, 2013; Wood, 2010).

In standard ecological-inference studies, the local effect of species interactions on community dynamics are approximated by the Jacobian coefficients of the data generating process (Deyle et al., 2013; Ives, Dennis, Cottingham, & Carpenter, 2003; Ushio et al., 2018). Formally, the Jacobian coefficients are the partial derivatives of the vector field of a dynamic model with respect to the state variables. In a population dynamics setting, this corresponds to the change in growth of a species as a result of a change in abundance of another species. Jacobian coefficients can be estimated from nonlinear time-series data using either parametric or nonparametric approaches. Using parametric approaches, the Jacobian coefficients are computed analytically from the assumed model after inference of its parameters (performed using, e.g. state-space Bayesian approaches or Kalman filters [Bucci et al., 2016; Meyer & Christensen, 2000; So, Ott, & Dayawansa, 1994; Stein et al., 2013]). Using nonparametric approaches instead, the Jacobian coefficients are inferred directly from the data with minimal assumptions (such as stationarity and distribution of the noise) on the data generating process (Deyle, May, Munch, & Sugihara, 2016; Ghosh, Mukhopadhyay, Roy, & Bhattacharya, 2014; Ting, D'Souza, Vijayakumar, & Schaal, 2008). Here, we focus on nonparametric approaches for inference and forecasting.

Earlier work has shown that learning algorithms based on nonlinear-state-space reconstruction can be used to infer nonparametrically Jacobian coefficients and to forecast species abundances from noisy ecological time series (Sugihara, 1994). In this context, a commonly used predictive nonlinear function approximation algorithm, known as S-map (a locally weighted multivariate linear regression model with state-space dependent kernel function (Deyle et al., 2016)), has been derived to deal with signals generated by a mixture of linear and nonlinear components contaminated by noise. The S-map involves computing Jacobian matrices sequentially as the system travels along its attractor. The coefficients of these Jacobian matrices vary with each location on the attractor, giving rise to state-dependent interactions (Deyle et al., 2016). The robustness of the S-map to observational noise has been demonstrated; however, its vulnerability to process noise remains a significant open problem that needs to be addressed in order to increase the applicability of this method.

In this manuscript, we show that by integrating appropriate regularization techniques to the S-map, it is possible to increase significantly the performance of both the in-sample inference of the Jacobian coefficients and the out-of-sample forecast of species abundances in the presence of process noise. To validate our results, we use three standard stochastic population dynamics models, which allow us to represent ecological time-series data while knowing the ground truth for inference and forecasting. We further validate our methodology on empirical data.

2 | MATERIALS AND METHODS

2.1 | Noisy nonlinear time series

The goal of this study is to derive an algorithm for an efficient inference of Jacobian matrices and out-of-sample forecast of species

abundance from nonlinear time series in the presence of process noise. Hence, the time series used in this study were generated by numerical integration of Ito stochastic differential equations (SDE). To test the proposed algorithm against different levels of noise and nonlinearity, we considered three parameterizations of a stochastic Lotka–Volterra (SLV) dynamics: a two-species predator–prey model, a three-species rock–paper–scissor (RPS) model and a four-species model that exhibits chaotic trajectories. These models were chosen to (a) generate fluctuations that resemble the behaviour of observed ecological time series (Bjørnstad & Grenfell, 2001; Nahum, Harding, & Kerr, 2011), (b) produce dynamics that are notoriously challenging to forecast in the presence of process noise (Cheng et al., 2015), and (c) to explicitly write down the analytical form of the noise and systematically study its behaviour (see Supplementary Information).

Specifically, time series were generated by numerical integration of:

$$\dot{\vec{x}} = \vec{x}(\vec{r}_\gamma - \mathcal{A}_\gamma \vec{x}) + \frac{1}{\sqrt{V}}\xi_\gamma(\vec{x},t), \tag{1}$$

where $\vec{x}$ is the state vector (i.e. species abundance) and $\vec{r}_\gamma$ and $\mathcal{A}_\gamma$ are the vector of intrinsic growth rates and the interaction matrix of model γ respectively. The noise component $\xi_\gamma(\vec{x},t)$ in Equation (1) is a Gaussian random variable with mean $\langle \xi_\gamma(\vec{x},t)\rangle = 0$ and variance $\langle \xi_{\gamma,i}(\vec{x},t), \xi_{\gamma,j}(\vec{x},t')\rangle = \mathcal{T}_{ij}(\vec{x})\delta(t-t')$ where $\mathcal{T}$ is the covariance matrix of the noise. Note that the noise is multiplicative and depends on the parameters of the model. The covariance matrix of the noise in Equation (1) was derived from the master equation of the microscopic stochastic process using the Van Kampen System size expansion (Van Kampen, 1992; Gardiner, 2004; see Supplementary Information section S2). The noise level was set by adjusting the system size V (note that for $V \to \infty$ we recover a deterministic equation).

To integrate Equation (1) numerically, we used the python library *sdeint*. To compare inferred and analytical Jacobian coefficients, we simulated the noise with a vector of independent Wiener processes $(d\vec{W}(t))$ with zero mean and standard deviation $\sigma = \sqrt{dt}$ where $dt = 0.01$ is the integration step (i.e. the noise component is $\frac{1}{\sqrt{V}}\mathcal{T}(t)d\vec{W}(t)$. Note that additional noise was introduced by the sampling procedure. That is, after numerical integration of Equation (1), we sampled equidistant points ($\Delta T = 200$) and used this sampled time series for the numerical analysis, rather than the whole solution of Equation (1). This procedure accounts for the fact that observational data are sampled far more sparsely than the numerical solutions of an equation. Figure 1 shows a graphical example of the time series generated by integration of these models and Supplementary Information section S2 s the details of the parameters used in this study.

Our proposed methodology for inference and forecasting is an extension of a locally linear regression scheme with state-space-dependent kernel function known as S-map (Sugihara, 1994). For completeness, in the next section, we briefly review the details of the S-map. The readers familiar with the S-map methodology may skip the next section.

2.2 | The S-map

The S-map is a nonlinear extension of standard linear vector autoregressive models (VAR). It differs from VAR and close linear relatives, such as dynamical linear models, in that the coefficients (or weights) of the regression in the S-map depend upon the position of the predictee on an attractor in the state space (Hsieh, Glaser, Lucas, & Sugihara, 2005), and not on its proximity in time. Therefore, the S-map is particularly useful when the observed time series has been generated by nonlinear processes in which state-space dependence is an important factor.

In particular, the S-map is the SVD solution for C to the linear model:

$$B = A \cdot C, \tag{2}$$

with $B_k = w_k x_i(t_k + 1) = w_k Y_i$, $A_{kj} = w_k X_j(t_k)$, and

$$w_k = e^{-\theta \frac{||x(t_k) - x(t^*)||}{\bar{d}}} \tag{3}$$

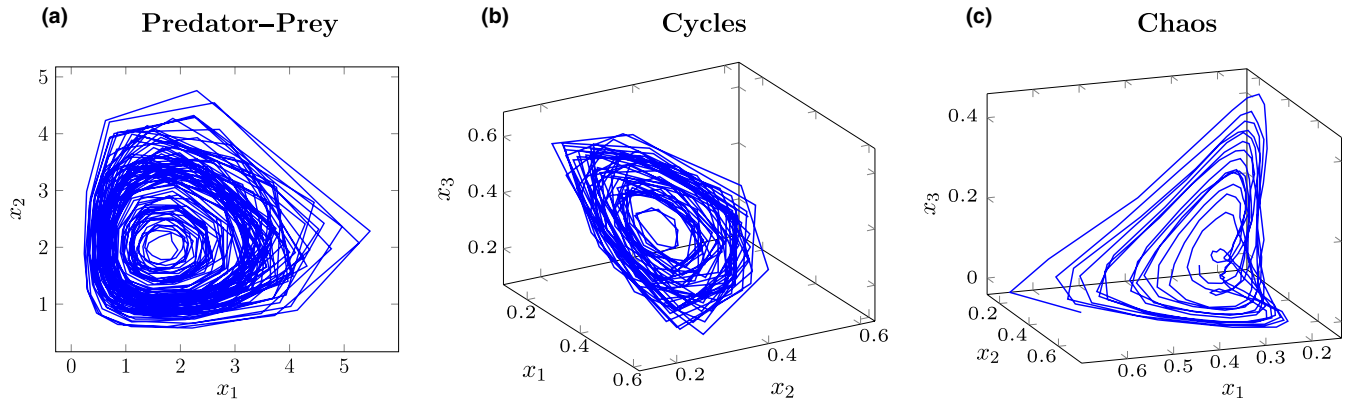

FIGURE 1 Illustrative example of the studied nonlinear, noisy, ecological time series. The figure shows the attractor of a two-species predator–prey model (Panel a), a three-species rock–paper–scissor game which exhibits cyclic dynamics (Panel b) and a four-species chaotic dynamics (Panel c). Note that the chaotic attractor in Panel c is a projection of the 4-dimensional attractor in a 3-dimensional state space. Each axis corresponds to a species abundance

are state-dependent weights. Note that in Equation (3), $x(t^*)$ is the predictee variable and $\bar{d} = \frac{1}{n}\sum_i ||x(t_i) - x(t^*)||$ is the average distance from the target point (Deyle et al., 2016; Dixon, Milicich, & Sugihara, 2001; Sugihara, 1994). The weights then depend upon the distance of the target point in the state space and on a tuning parameter θ that controls the local weighting (Deyle et al., 2016; Sugihara, 1994). That is, θ controls the degree of nonlinearity of the S-map. For this reason, the forecasting skill of the S-map has been extensively used as a test for nonlinearity of the time series (Perretti et al., 2013; Sugihara, 1994) that is time-series generated by nonlinear processes can be better forecast with $\theta \neq 0$ (Dakos, Glaser, Hsieh, & Sugihara, 2017). Note that solving the linear problem above is equivalent to solving a weighted least square minimization problem, which reads in matrix form as:

$$\hat{c} = \underset{c\in\mathbb{R}^d}{\operatorname{argmin}} \frac{1}{n}\sum_j^n w_j(y_j - x_j c)^2 = \min_{c\in\mathbb{R}^d} \frac{1}{n}(Y - Xc)^T W(Y - Xc), \tag{4}$$

where X is the data matrix (for the rest of this manuscript, we will use all the variables of the time series to reconstruct the attractor), W is a diagonal matrix of the weights in Equation (3), and c is one row of the Jacobian matrix. The solution of this minimization problem (i.e. ordinary least square, OLS) is $\hat{c} = (X^T WX)^{-1} X^T WY$. Equivalently, the solution of the SVD for C in the linear model in Equation (2) is

$$\hat{c} = (\tilde{V}\Sigma^{-1}\tilde{U}^T)WY, \tag{5}$$

where the solution is written in terms of the weighted data matrix $\tilde{X} = WX$. It is then straightforward to prove that the two solutions are equivalent (see Section S1). To avoid numerical instabilities in the inversion of the matrix Σ, the S-map methodology removes small singular values before inversion by setting to zero each singular value below a predefined threshold, that is, $\sigma_{ii} < \epsilon$ for some fixed ϵ (typically $\epsilon = 10^{-5}$ (Ye et al., 2017)).

Importantly, as mentioned before, process noise can introduce upper bounds to the forecasting skill of learning algorithms, such as the S-map. Specifically, small perturbations in the input data (induced by noise) can cause large changes in the coefficients of the fit and the forecast values. Indeed, small singular values can drastically increase the prediction variance of the ordinary least square optimization (Kugiumtzis, Lingjærde, & Christophersen, 1998), an effect that is amplified by noise. However, the cut-off value of the singular value is not universal, as in the standard S-map, but depends upon the structure of the data to fit. Generally, depending on the data, one needs to select the cut-off by appropriately balancing the prediction bias and its variance. This is typically done by introducing some form of regularization of the fit. Below, we discuss how to introduce regularization techniques to tackle small singular values, and to improve the performance of the fit along with the out-of-sample generalization in the presence of process noise.

2.3 | Regularized S-map

Regularized loss minimization is a well-established learning rule, in which both a regularization function ($\mathcal{R}$) and the empirical risk ($\mathcal{L}$) are minimized together (Shalev-Shwartz & Ben-David, 2014):

$$\hat{c} = \underset{c\in\mathbb{R}^d}{\operatorname{argmin}}\{\mathcal{L}(\mathcal{D},\vec{c}) + \mathcal{R}(\vec{r},\vec{c})\}, \tag{6}$$

where $\mathcal{D}$ are the data, $\vec{c}$ is the vector of parameters to estimate and $\vec{r}$ is a vector of regularization parameters. Regularized loss minimization is a typical approach used to constrain the coefficients of a regression model in order to make the algorithm more stable in the presence of noise in the data (recall that a learning algorithm is stable if small changes in the input data do not generate large changes in the output (Abu-Mostafa et al., 2012; Shalev-Shwartz & Ben-David, 2014)).

There are a variety of possible regularization functions to choose from; here, we discuss the most common ones. Tikhonov regularization requires to substitute Equation (4) with the following minimization problem called regularized least square or ridge regression (Shalev-Shwartz & Ben-David, 2014):

$$\hat{c} = \underset{c\in\mathbb{R}^d}{\operatorname{argmin}} \frac{1}{n}(Y - Xc)^T W(Y - Xc) + \lambda||c||_2^2. \tag{7}$$

The regularization function $R(\lambda, c) = \lambda||c||_2^2$ constrains the hypothesis space $\mathcal{H}$ (i.e. controls its complexity) or, equivalently, adjusts the least-square solution to avoid numerical instability during matrix inversion. In fact, the solution of Equation (7) is:

$$\hat{c} = (X^T WX + \lambda nI)^{-1} X^T WY, \tag{8}$$

where the regularization parameter λ, selected by running cross-validation on a training set data, avoids possible instabilities during matrix inversion in Equation (8). We chose Tikhonov regularization because the solution of Equation (7) can also be written in terms of singular values, and it can be shown that the regularization parameter λ damps any term in the solution with $\sigma_i < \lambda_i$. Therefore, Tikhonov regularization is strictly related to a truncated singular value decomposition (Gorodnitsky & Rao, 1994). Indeed, Tikhonov regularization introduces the balance between prediction bias and variance that we discussed in the previous section (Abu-Mostafa et al., 2012). This means that Tikhonov regularization is, in fact, the natural type of regularization of the S-map. Note that regularizations of local linear fits (Kugiumtzis et al., 1998) have already been considered in settings different than the S-map and for purposes different than controlling the quality of the inference of the effect of species interactions.

Generally, the choice of the regularization function is heuristic (Abu-Mostafa et al., 2012). The choice depends upon the structure of the available data. For example, ridge regression filters the small singular values that can create instabilities when they are contaminated by noise. However, it may produce suboptimal results if most of the information is contained in these values. Other types of regularizations, such as LASSO penalization (Tibshirani, 1996), can be more suitable when dealing with sparse datasets. For comparison purposes, here, we also use elastic-net regularization (Zou & Hastie, 2005), which is a more flexible function than Tikhonov. Elastic net is a combination of Tikhonov ($L2$ norm) and LASSO ($L1$ norm) regularization with an additional parameter α that controls the relative strength of the two norms:

$$\hat{c} = \underset{c \in \mathbb{R}^d}{\operatorname{argmin}} \frac{1}{n}(Y - Xc)^T W(Y - Xc) + \lambda((1-\alpha)||c||_1 + \alpha||c||_2^2). \quad (9)$$

Note that when $\alpha \to 1$, one recovers Tikhonov regularization. Whereas when $\alpha \to 0$, one recovers LASSO. Importantly, elastic-net regularization has the advantage that it can be adapted to the particular structure of the time series. For example, if the time series is sparse, then α should be small. However, if small constraints are enough to avoid both over-fitting and instability, then α should be close to one. In order to properly compare two different algorithms, one can always keep $\alpha < 1$.

Finally, we discuss the choice of weights in the locally weighted regression (Equation (4)). The standard S-map uses exponentially decay kernels, such as Equation (3). The choice, however, is not unique and generally depends upon the characteristics of the analysed dataset. For example, exponentially decay kernels are motivated by Lyapunov instability, where information decays exponentially. Other kernels, such as the Epanechnikov or the tricubic kernel (Hastie, Tibshirani, & Friedman, 2001), may be more suitable than Equation (3) for different types of nonlinearity in the time series. In general, a preliminary analysis of the data is needed in order to choose the most appropriate regression model. Here we test the following kernels:

$$\begin{aligned} K(x,x') &= (1 - (\tfrac{||x-x'||}{\theta})^3)^3 \Theta(1 - \tfrac{||x-x'||}{\theta}), \\ K(x,x') &= \tfrac{3}{4}(1 - (\tfrac{||x-x'||}{\theta})^2) \Theta(1 - \tfrac{||x-x'||}{\theta}), \\ K(x,x') &= e^{-\theta \frac{||x-x'||}{d}}. \end{aligned} \quad (10)$$

These kernels are the tricubic, Epanechnikov and exponential respectively. Note that $\Theta(\cdot) = 1$ is the Heaviside theta (i.e. $\Theta(n) = 1$ if $n \geq 0$ and zero otherwise). Below, we specify how to measure and compare the inference and forecasting skills of the standard (i.e. non-regularized) and regularized S-maps.

2.4 | Inference of Jacobian coefficients

To infer Jacobian coefficients from the stochastic time series using the standard and regularized S-maps, we ran the following analysis. Firstly, we generated time series of species abundances by integrating numerically Equation (1) and by sampling data points at constant rate as discussed above (see Figure 1 for a graphical example of these time series). Then, we randomly sampled a block of length $L = 100$ of the time series. The block is standardized to have zero mean and standard deviation one. Finally, we fitted the data using the standard and the regularized S-map. Regularization and kernel parameters were chosen by running leave-one-out cross-validation (Hastie et al., 2001) that is by choosing those parameters that minimize the validation error on the training set. Note as well that we fit each species independently so regularization and kernel parameters are d dimensional vectors (where d is the number of species in the system). Because we want to show the effect of different regularization functions on the performance of the S-map, we fixed the parameter α in the elastic-net regularization to be strictly less than one. Specifically, we fixed $\alpha = 0.9$ for the RPS model, and $\alpha = 0.95$ for the chaotic and predator–prey dynamics. These values were chosen after a preliminary analysis of the time series. We did not run cross-validation on α because, contrary to the parameters λ and θ, its optimum value did not vary significantly across the realizations.

The Jacobian coefficients are the coefficients of the fit (Deyle et al., 2016; Ushio et al., 2018). To measure the quality of the inference, we computed the Pearson correlation coefficient (Sugihara, 1994) and the root mean square error (RMSE) between analytical and inferred Jacobian matrices. To obtain a reliable statistics, we repeated this process for 80 random blocks (i.e. for 80 times, we sample a random block, standardized it, run leave-one-out cross-validation to select regularization and kernel parameters, and estimate the Jacobian coefficients). Finally, notice that, given that we are dealing with synthetic data for which we know the true dimensionality, we used the dimension of the model as the embedding dimension for the inference of the Jacobian coefficients. However, recall that for the analysis of empirical data, for which the true embedding dimension is generally unknown, this parameter must be chosen together with the regularization and kernel parameters.

2.5 | Forecasting of species abundances

Using the kernel and regularization parameters estimated with leave-one-out cross-validation in the training set, we built a predictive model and forecast 30 data points out-of- sample. To make out-of-sample forecast, we ran the following analysis: after training over L-data points (i.e. over the whole training set), the S-map provided $(L-1)$ coefficients given that the last data point of the training set was not trained (i.e. we did not have information about the $L + 1$ abundance which was the first data in the test set). Therefore, we used the $(L - 1)$th coefficient and the Lth data point to make predictions about the first data point in the test set (i.e. the $(L + 1)$th point in the time series). Predictions were made with a locally linear model (Ye et al., 2017):

$$x_i(t+1) = c_0 + \sum_{j=1}^{d} J_{ij}(t-1)x_j(t), \quad (11)$$

where c_0 is the intercept fitted adding a column of ones to the data matrix (Hastie et al., 2001) and d is the number of species in the system. The matrix J_{ij} is the Jacobian matrix, whose rows are the vectors ci solutions of the minimization problem in Equations (7) and (9) (and the standard S-map). After each forecast, we used the predicted point to fit the Lth interaction coefficient, and we repeated the operation for as many times as numbers of data points we had in the test set (i.e. 30). The forecast is purely out-of-sample because this procedure never uses the original test set to select the model parameters or to make predictions. In addition, the test set is standardized using the mean and the standard deviation of the training set. Hence, we truly tested for the generalization skills of the algorithms on unseen data.

To measure the quality of the forecasting skill, we used the Pearson correlation coefficient (Sugihara & May, 1990) and the RMSE as we did for the coefficients of the Jacobian matrix. Moreover, as benchmark for the value of the RMSE in the test set, we compared the quality of the forecast with the RMSE of the naive predictor (i.e. the predictor that uses the last point in the training set as prediction for any data in the test set).

To test the robustness of the algorithm, we repeated the inference and forecasting analyses with different training-set lengths ($L=\{20,40,80,160\}$) and different levels of (Gaussian) observational error with standard deviation $\sigma \in [0.05,0.2]$ and $\mu=0$ added to the training data. In order to make the noise comparable to the perturbed value, we added multiplicative noise to the data, that is $\vec{x}_{\text{perturbed}}=\vec{x}_{\text{unperturbed}}+\vec{x}_{\text{unperturbed}}\times\epsilon$, where $\epsilon\sim\mathcal{N}(0,\sigma)$ (notice that $\vec{x}$ come for the integration of an SDE so it already has a intrinsic noise component).

2.6 | Analysis on empirical data

Finally, we tested the performance of the regularized S-map on two empirical time series of marine microbial communities: the Bermuda Atlantic Time Series (BATS) and the Hawaii Ocean Time Series (HOTS). Because in empirical data, we lack the ground truth of Jacobian coefficients, we only tested the forecasting skills of the algorithm. The time series can be freely downloaded from http://batsftp.bios.edu/BATS/bottle/bval_bottle.txt (BATS) and http://hahana.soest.hawaii.edu/hot/hot-dogs/index.html (HOTS).

The BATS data include the abundance of four species: *Prochlorococcus*, *Synechococcus*, *Picoeukaryotes* and *Nanoeukaryotes*. Here, we focused on the time period that goes from September to October 2011 for which data were collected approximately twice per day. This time series has 30 data points. We split our data in 28 points for training (i.e. for cross-validation) and 2 points for testing. The HOTS data include species abundances of: *Prochlorococcus*, *Synechococcus*, *Picoeukaryotes* and heterotrophic bacteria. These data were collected monthly from 2006 to 2016 (93 data points). We used 91 points for training and 2 points for testing.

To analyse the empirical time series, we ran a standard variable selection (Hastie et al., 2001) to subset only those variables and their time-lags that provide the best validation error. Then, we ran the same analysis described above to choose the optimal regularization parameters and kernel function of the regularized S-map. We only report the RMSE of the analysis (note the Pearson correlation) given that the test set only includes 2 data points.

3 | RESULTS

3.1 | Results on synthetic data

Overall, we found that the regularized S-map consistently outperforms the standard S-map in both tasks: the inference of Jacobian coefficients and the forecast of species abundances. Tables 1 and 2 provide the summary results of the best inference and forecasting skills of the two algorithms. Performance is expressed in terms of median RMSE and correlation coefficient (ρ) over an ensemble of 80 random chunks of the time series. Errors around the median are the 95% confidence intervals computed with nonparametric bootstrapping (full results are shown in Figs 2 and S1). The tables show the results of the fit that provides the best trade-off between maximum correlation coefficients and minimum RMSE using the same kernel and regularization function. The RMSE of the standard S-map on the test set of the chaotic model is out of scale and is reported as NA.

Importantly, Figure 2 and Table 3 illustrate that the quality of the inference and the forecasting strongly depends upon the choice of the kernel function used in the weighted regression scheme. For example, Figure 2b shows that for the periodic predator–prey time series, the best forecast is achieved using the Epanechnikov kernel with an elastic-net regularization function. Instead, Figure 2f shows that for the chaotic time series, the exponential kernel exhibits the best inference and forecasting skills.

Moreover, Figure 2 shows that one can obtain different reconstructions of the Jacobian coefficients (left panels) simply by using different regularization schemes or kernel functions for the weights of the linear regression (see e.g. panel e). The figure also shows that the quality of the reconstruction is not necessary informative of the generalization skills of the algorithm (right panels). This effect is clearly illustrated in Figure 2. For example, one can compare the performance of the exponential kernel and an elastic-net regularization function for the inference of Jacobian coefficients in panel (a) with its generalization skills in panel b. Overall, these results pose serious challenges to the problem of model selection. That is, using empirical data one can only select parameters, kernels and regularization

TABLE 1 Inference of the Jacobian coefficients using synthetic data

Inference	Predator–prey		Cyclic		Chaotic	
Regression scheme	ρ	RMSE	ρ	RMSE	ρ	RMSE
Standard S-map	0.61 ± 0.01	0.62 ± 0.01	0.51 ± 0.02	0.37 ± 0.01	0.43 ± 0.05	0.61 ± 0.02
Regularized S-map	**0.64** ± 0.01	**0.58** ± 0.01	**0.61** ± 0.01	**0.31** ± 0.007	**0.5** ± 0.03	**0.36** ± 0.01

For each of the three models analysed (columns), the table shows the performance of the standard and regularized S-map (rows) in inferring the analytic Jacobian coefficients. The performance is computed as the Pearson correlation (ρ) and root mean square error (RMSE), that is, smaller (larger) RMSEs (correlations) correspond to better performances. Note that the regularized S-map consistently outperforms the inference skills of the standard S-map. Bold values refer to optimal performance.

TABLE 2 Out-of-sample forecasting skills using synthetic data

Forecasting Regression scheme	Predator–prey		Cyclic		Chaotic	
	ρ	RMSE	ρ	RMSE	ρ	RMSE
Standard S-map	0.75 ± 0.18	0.88 ± 0.26	0.41 ± 0.23	1.08 ± 0.22	0.03 ± 0.04	NA
Regularized S-map	**0.88** ± 0.08	**0.49** ± 0.06	**0.60** ± 0.06	**0.72** ± 0.08	**0.8** ± 0.07	**0.65** ± 0.11
Naive predictor	NA	1.32 ± 0.06	NA	1.4 ± 0.08	NA	1.35 ± 0.1

For each of the three models analysed (columns), the table shows the performance of the standard S-map, regularized S-map and naive predictor (rows) in predicting species abundances. The performance is computed as the Pearson correlation (ρ) and root mean square error (RMSE), that is, smaller (larger) RMSEs (correlations) correspond to better performances. Note that the regularized S-map consistently outperforms the standard S-map (and the naive predictor). Bold values refer to optimal performance.

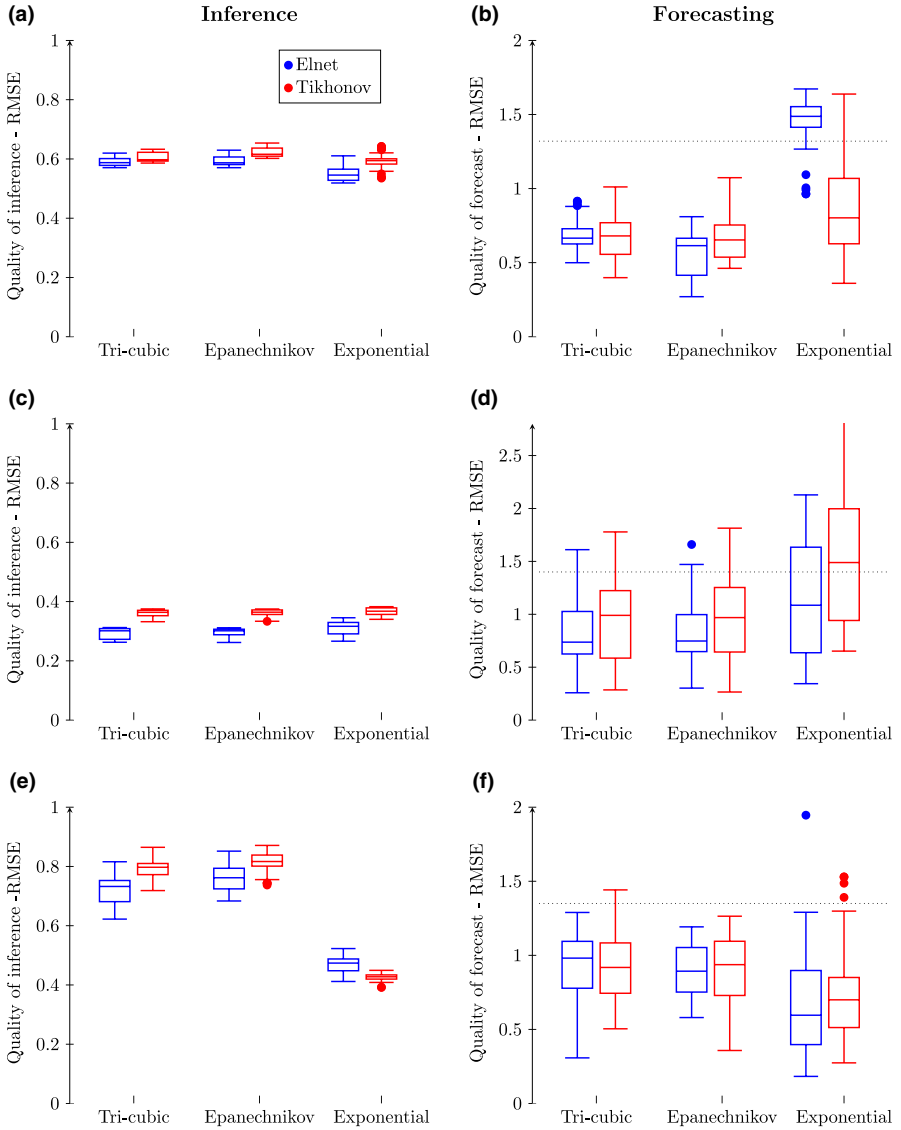

FIGURE 2 Importance of the kernel function. The figure illustrates the full distribution of the inference and forecasting skills of the regularized S-map using the tricubic kernel, the Epanechnikov kernel and the exponential kernel (used in the standard S-map), see Equation (10). Models go from top to bottom: Predator–prey, rock–paper–scissor and chaotic Lotka–Volterra. The left panels show the results of the inference, whereas the right panels show the results of the forecast. The dotted lines in the right panels depict the value of the naive error. All performances are reported as the root mean square error (RMSE), that is, smaller RMSEs correspond to better performances. Overall, the figure illustrates that the choice of the kernel function is critical for the performance of the algorithm

functions based on validation or prediction errors; yet, this may not be the optimal solution for inference purposes.

Finally, Figure 3 shows the quality of the inference and forecast as a function of the level of observational noise added to the training data. Performance is again expressed in terms of median error over 60 random chunks of the time series and errors are the 95% confidence interval around the median error. The figure shows that the median RMSE, while increasing, remains significantly lower than the naive prediction and the standard S-map (compare the Figure with TableS 1 and 2). Figure S2 in the Supplementary Information shows the robustness of the algorithm to different lengths of the training set. Interestingly, in our example, smaller

TABLE 3 Combinations of kernels and regularization functions that provide the best inference and forecasting skills

	Tikhonov	Elastic net
Exponential	Chaos, Chaos, Predator–Prey	Cycles
Epanechnikov		Predator–Prey
Tricubic		Cycles

For each of the three tested models, the table shows which is the best combination of kernels and regularization functions to infer the Jacobian coefficients (red) and to forecast species abundance (blue). Note that the best kernel is always chosen to be the one that provides the best trade-off between maximum correlation coefficient and minimum RMSE.

TABLE 4 Performance of the algorithms on the test set of the empirical time series

Method	BATS	HOTS
Standard S-map	0.21 (0.19, 0.36, 0.07)	0.20 (0.24, 0.16)
Regularized S-map	**0.18** (0.17, 0.3, 0.07)	**0.15** (0.23, 0.07)
Naive forecast	0.32	0.93

Performance is reported as the root mean square error (RMSE) on the test set, that is, smaller RMSEs correspond to better performances. As a benchmark, we also report the error of the naive predictor (see Methods section). The numbers between brackets are the errors on each of the single species analysed. The table illustrates that the regularized S-map outperforms the standard S-map also on the empirical data. Bold values refer to optimal performance.

training sets do not necessarily correspond to larger errors. This is particularly true for the RPS dynamics which is the most noisy model we used. This counter-intuitive result could be explained by the fact that, because the time series, we are analysing are stochastic, training on more data increases the probability of training on noisy regions of the attractor. Overall, Figure 3 and Figure S2 show that the regularized S-map is robust to different levels of noise and sizes of training data.

3.2 | Results on Empirical data

Focusing on the BATS data, we found that the exponential kernel and elastic-net regularization function with $\alpha = 0.85$ provided the best out-of-sample skills (in terms of RMSE) for predicting the abundances of the three species of bacteria. In the analysis, we included the abundance of *Prochlorococcus* (embedding dimension, E = 1), *Synechococcus* (E = 2) and *Picoeukaryotes* (E = 1). Moving to the HOTS data, we found that the Epanechnikov kernel with $\alpha = 0.9$ provided the best performance. In the analysis, we included the abundance of *Prochlorococcus* (E = 2) and *Picoeukaryotes* (E = 1).

Table 4 shows the performance comparisons between the standard and regularized S-maps. We used the naive forecast as a benchmark for the quality of performances (see Methods section). As expected from the results derived from synthetic data, we found that the regularized S-map consistently outperforms the standard S-map on the empirical time series.

4 | DISCUSSION

In this manuscript, we have studied the application of regularized loss minimization methods to a locally weighted linear fit known as S-map developed to perform inference and forecasting from nonlinear ecological time series. The goal of this study has been to investigate if regularization applied to S-maps can improve the inference of the Jacobian coefficients (that are typically used as proxy for the local effect of biotic interactions on community dynamics) and the out-of-sample forecasting skill of species abundances when the time series are contaminated with process noise. In particular, to study the effect of regularization on different types of dynamics and with different levels of process noise, we have investigated three standard nonlinear models: a predator–prey, a cyclic and a chaotic time series. Because the choice of the regularization function is not unique (Abu-Mostafa et al., 2012), we have investigated two different regularization algorithms (Tikhonov and elastic net) and we have compared their performance on simulated and empirical data.

We have found that by imposing different types of regularization functions and appropriately choosing the kernel functions, we can improve the inference and the out-of-sample forecasting skills of the S-map. Generally, we recommend testing different regularization schemes and kernels in a preliminary analysis of a new dataset to determine which one performs best (the code accompanying this manuscript allows to easily perform this task). We speculate that the type of kernel that yields the best forecast could be informative

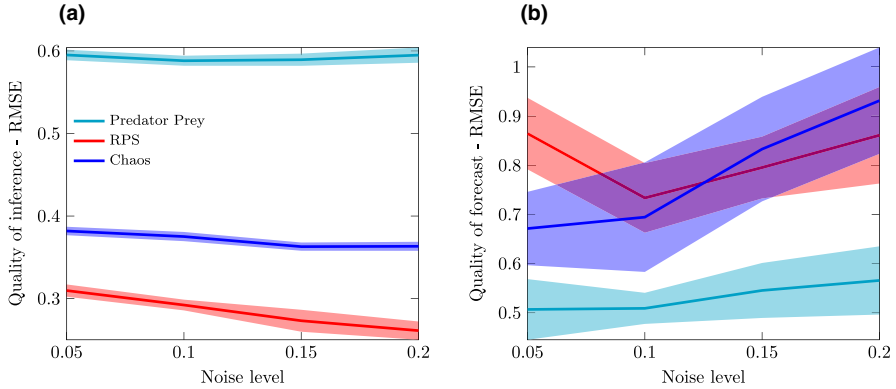

FIGURE 3 Robustness of the regularized S-map to observational noise. The figure shows that both the inference and the forecasting skills of the regularized S-map are robust to observational noise. Note that while the RMSE increases as we increase the level of noise, it still remains significantly lower than the naive error (see Table 2 and Figure 2)

of the type of dynamics generating the time series (this comes directly from the fact that kernels are a way to express prior knowledge about the data). For example, as expected, we have found that exponentially decay kernels perform better when forecasting with chaotic time series—where information decays exponentially (although more research is needed to validate or reject this hypothesis). Moreover, we suggest to choose the kernel based on a trade-off between different performance metrics such as correlation coefficients and the RMSE.

Most of the results of this study have been based on the analysis of synthetic data. We have used synthetic data to show the quality of the inference by comparison with analytic results (this is not possible with empirical data where the true Jacobian coefficients are unknown). Furthermore, the use of synthetic data has allowed us to systematically introduce process noise in the time series. However, to further test the validity of our approach, we have also provided an application on two empirical datasets. We have found that, as expected, the regularized S-map outperforms the standard S-map also on these empirical time series.

It is also worth mentioning that while the S-map has been and can be used on empirical data, there are a number of limitations for its applicability. These limitations apply equivalently to both the standard and the regularized S-maps. Firstly, it is important to stress that for an accurate reconstruction of Jacobian coefficients the S-map needs densely sampled neighbourhoods on the attractor. Therefore, the data requirement for an accurate inference increases exponentially with the system size (Hastie et al., 2001). Generally, if the number of variables (i.e. dimensionality of the system) is large and the time series is short, we discourage the use of the S-map for inference studies. The exact definition of large, however, strongly depends upon the length of the time series. Here, for example, with four species and 50 data points, we have already obtained accurate reconstruction of Jacobian coefficients even in the presence of process and observational noise. Moreover, the dimensionality of the system may be reduced by excluding variables that are not causally related with the rest of the system (Sugihara et al., 2012). Secondly, as for many learning algorithms used in nonlinear time-series analysis, the stationarity of the data is an important requirement. Thus, to guarantee the statistical significance of the inferred Jacobian coefficients, a preliminary check for stationarity is needed (Kantz & Schreiber, 2004). Finally, because of the sparsity and nonlinearity of empirical time series, coefficients inferred with statistical algorithms, as the regularized S-map, may not be practically identifiable (meaning that more than one set of parameters can explain the data approximately equally well) (Saccomani, 2013; Saccomani & Thomaseth, 2016). Hence, inference studies should also include uncertainty quantification analysis (Cenci & Saavedra, 2018b). The applicability of the S-map for predictive studies is, on the other hand, less of a concern because the validity of the results can always be checked by means of cross-validation.

Overall, the increasing amount of available time-series data is paving the way for moving theoretical population biology from being an assumption-driven towards a data-driven science. Because ecological time series are the product of an interplay between nonlinear deterministic and stochastic dynamics, the integration of methods of nonlinear time-series analysis (based on dynamical system theory) with machine learning algorithms (developed to deal with observational data contaminated by noise) is an important step forward towards this goal.

4.1 | Data availability

The code for inference and forecasting with the regularized S-map as well as for the integration of the stochastic model and the computation of the Jacobian matrices is publicly available at https://zenodo.org/badge/latestdoi/164699095. The repository also includes the analysis of the empirical data presented in the main text. The code for the standard, non-regularized, S-map is freely available in the *rEDM* library in R.

ACKNOWLEDGEMENTS

We thank Chuliang Song and Gabriel Leventhal for insightful discussions. The funding was provided by the MIT Research Committee Funds and the Mitsui Chair (SS) and by Department of Defense SERDP 15 RC-2509, National Science Foundation NSF-DEB-1020372 and NSF-ABI-Innovation DBI- 1667584, McQuown Chair, UCSD (GS).

CONFLICT OF INTERESTS

The authors declare no competing financial interests.

AUTHORS' CONTRIBUTIONS

S.C., G.S. and S.S. designed the study; S.C. performed the study; S.S. supervised the study. S.C., G.S. and S.S. wrote the manuscript.

REFERENCES

Abu-Mostafa, Y. S., Magdon-Ismail, M., & Lin, H.-T. (2012) *Learning from data*. AMLBook.

Barraclough, T. G. (2015). How do species interactions affect evolutionary dynamics across whole communities? *Annual Review of Ecology, Evolution, and Systematics*, *46*, 25–48. https://doi.org/10.1146/annurev-ecolsys-112414-054030

Bjørnstad, O. N., Fromentin, J.-M., Stenseth, N. C., & Gjøsater, J. (1999). Cycles and trends in cod populations. *Proceedings of the National Academy of Sciences*, *96*, 5066–5071. https://doi.org/10.1073/pnas.96.9.5066

Bjørnstad, O. N., & Grenfell, B. T. (2001). Noisy clockwork: Time series analysis of population fluctuations in animals. *Science*, *293*, 638–643. https://doi.org/10.1126/science.1062226

Bradley, E., & Kantz, H. (2015) Nonlinear time-series analysis revisited. *Chaos: An Interdisciplinary Journal of Nonlinear Science*, *25*, 097610. https://doi.org/10.1063/1.4917289

Bucci, V., Tzen, B., Li, N., Simmons, M., Tanoue, T., Bogart, E., … Gerber, G. K. (2016). Mdsine: Microbial dynamical systems inference engine for

microbiome time-series analyses. *Genome Biology*, *17*, 121. https://doi.org/10.1186/s13059-016-0980-6

Burkov, V., Novikov, D., & Shchepkin, A. (2015) *Control mechanisms for ecological-economic systems. studies in systems, decision and control*. Springer International Publishing, Switzerland.

Casdagli, M., Eubank, S., Farmer, J., & Gibson, J. (1991). State space reconstruction in the presence of noise. *Physica D: Nonlinear Phenomena*, *51*, 52–98. https://doi.org/10.1016/0167-2789(91)90222-U

Cenci, S., Montero-Castaño, A., & Saavedra, S. (2018). Estimating the effect of the reorganization of interactions on the adaptability of species to changing environments. *Journal of Theoretical Biology*, *437*, 115–125. https://doi.org/10.1016/j.jtbi.2017.10.016

Cenci, S., & Saavedra, S. (2018a). Structural stability of nonlinear population dynamics. *Physical Review E*, *97*, 012401. https://doi.org/10.1103/PhysRevE.97.012401

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

Cheng, C., Sa-Ngasoongsong, A., Beyca, O., Le, T., Yang, H., Kong, Z. J., & Bukkapatnam, S. T. (2015). Time series forecasting for nonlinear and non-stationary processes: A review and comparative study. *IIE Transactions*, *47*, 1053–1071. https://doi.org/10.1080/07408 17X.2014.999180

Clark, J. S., Carpenter, S. R., Barber, M., Collins, S., Dobson, A., Foley, J. A., … Wear, D. (2001). Ecological forecasts: An emerging imperative. *Science*, *293*, 657–660. https://doi.org/10.1126/science.293.5530.657

Constable, G. W. A., Rogers, T., McKane, A. J., & Tarnita, C. E. (2016). Demographic noise can reverse the direction of deterministic selection. *Proceedings of the National Academy of Sciences*, *113*, E4745–E4754. https://doi.org/10.1073/pnas.1603693113

Coulson, T., Rohani, P., & Pascual, M. (2004). Skeletons, noise and population growth: The end of an old debate? *Trends in Ecology & Evolution*, *19*, 359–364. https://doi.org/10.1016/j.tree.2004.05.008

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

Deyle, E. R., Fogarty, M., Hsieh, C.-H., Kaufman, L., MacCall, A. D., Munch, S. B., … Sugihara, G. (2013). Predicting climate effects on pacific sardine. *Proceedings of the National Academy of Sciences*, *110*, 6430–6435. https://doi.org/10.1073/pnas.1215506110

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

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

Dixon, P., Milicich, M. J., & Sugihara, G. (2001) *Noise and nonlinearity in an ecological system*, pp. 339–364. Boston, MA: Birkhäuser. https://doi.org/10.1007/978-1-4612-0177-9

Dobrinevski, A., & Frey, E. (2012). Extinction in neutrally stable stochastic Lotka-Volterra models. *Physical Review E*, *85*, 051903. https://doi.org/10.1103/PhysRevE.85.051903

Enke, T. N., Leventhal, G. E., Metzger, M., Saavedra, J., & Cordero, O. X. (2018). Microscale ecology regulates particulate organic matter turnover in model marine microbial communities. *Nature Communications*, *9*, 2743. https://doi.org/10.1038/s41467-018-05159-8

Frentz, Z., Kuehn, S., & Leibler, S. (2015). Strongly deterministic population dynamics in closed microbial communities. *Phys. Rev. X*, *5*, 041014.

Gardiner, C. W. (2004) *Handbook of stochastic methods for physics, chemistry and the natural sciences*, volume 13 of Springer Series in Synergetics. Springer-Verlag, Berlin, 3rd edition. https://doi.org/10.1007/978-3-662-05389-8

Ghosh, A., Mukhopadhyay, S., Roy, S., & Bhattacharya, S. (2014). Bayesian inference in nonparametric dynamic state-space models. *Statistical Methodology*, *21*, 35–48. https://doi.org/10.1016/j.stamet.2014.02.004

Gorodnitsky, I., & Rao, B. (1994) Analysis of error produced by truncated svd and tikhonov regularization methods. Proceedings of 1994 28th Asilomar Conference on Signals, Systems and Computers, pp. 25 – 29 vol.1.

Hastie, T., Tibshirani, R., & Friedman, J. (2001) *The elements of statistical learning*. Springer Series in Statistics. Springer, New York. https://doi.org/10.1007/978-0-387-21606-5

Higgins, K., Hastings, A., Sarvela, J. N., & Botsford, L. W. (1997). Stochastic dynamics and deterministic skeletons: Population behavior of Dungeness crab. *Science*, *276*, 1431–1435. https://doi.org/10.1126/science.276.5317.1431

Hsieh, C.-H., Glaser, S. M., Lucas, A. J., & Sugihara, G. (2005). Distinguishing random environmental fluctuations from ecological catastrophes for the north pacific ocean. *Nature*, *435*, 336–340. https://doi.org/10.1038/nature03553

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

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

Kugiumtzis, D., Lingjærde, O., & Christophersen, N. (1998). Regularized local linear prediction of chaotic time series. *Physica D: Nonlinear Phenomena*, *112*, 344–360. https://doi.org/10.1016/S0167-2789(97)00171-1

Lawrence, D., Fiegna, F., Behrends, V., Bundy, J. G., Phillimore, A. B., Bell, T., & Barraclough, T. G. (2012). Species interactions alter evolutionary responses to a novel environment. *PLOS Biology*, *10*, 1–11.

Mason, N. W. H., Holdaway, R. J., & Richardson, S. J. (2018). Incorporating measurement error in testing for changes in biodiversity. *Methods in Ecology and Evolution*, *9*, 1296–1307. https://doi.org/10.1111/2041-210X.12976

Meyer, R., & Christensen, N. (2000). Bayesian reconstruction of chaotic dynamical systems. *Physical Review E*, *62*, 3535–3542. https://doi.org/10.1103/PhysRevE.62.3535

Mutshinda, C. M., O'Hara, R. B., & Woiwod, I. P. (2009). What drives community dynamics? *Proceedings of the Royal Society of London B: Biological Sciences*, *276*, 2923–2929. https://doi.org/10.1098/rspb.2009.0523

Nahum, J. R., Harding, B. N., & Kerr, B. (2011). Evolution of restraint in a structured rock–paper–scissors community. *Proceedings of the National Academy of Sciences*, *108*, 10831–10838. https://doi.org/10.1073/pnas.1100296108

Perretti, C. T., Munch, S. B., & Sugihara, G. (2013). Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data. *Proceedings of the National Academy of Sciences*, *110*, 5253–5257. https://doi.org/10.1073/pnas.1216076110

Ranta, E., Lundberg, P., Kaitala, V., & Laakso, J. (2000). Visibility of the environmental noise modulating population dynamics. *Proceedings of the Royal Society of London B: Biological Sciences*, *267*, 1851–1856. https://doi.org/10.1098/rspb.2000.1220

Saccomani, M. P. (2013) Structural vs practical identifiability in system biology. *WBBIO 2013 Proceedings*, p. pp. 305–313.

Saccomani, M. P., & Thomaseth, K. (2016) *Structural vs practical identifiability of nonlinear differential equation models in systems biology*, pp. 31–41. Springer International Publishing, Cham. https://doi.org/10.1007/978-3-319-45723-9

Schooler, S. S., Salau, B., Julien, M. H., & Ives, A. R. (2011). Alternative stable states explain unpredictable biological control of *Salvinia*

molesta in Kakadu. *Nature*, *470*, 86–89. https://doi.org/10.1038/nature09735

Shalev-Shwartz, S., & Ben-David, S. (2014). *Understanding machine learning: From theory to algorithms*. New York, NY: Cambridge University Press. https://doi.org/10.1017/CBO9781107298019

So, P., Ott, E., & Dayawansa, W. P. (1994). Observing chaos: Deducing and tracking the state of a chaotic system from limited observation. *Physical Review E*, *49*, 2650–2660. https://doi.org/10.1103/PhysRevE.49.2650

Spanio, T., Hidalgo, J., & Muñoz, M. A. (2017). Impact of environmental colored noise in single species population dynamics. *Physical Review E*, *96*, 042301. https://doi.org/10.1103/PhysRevE.96.042301

Stein, R. R., Bucci, V., Toussaint, N. C., Buffie, C. G., Rätsch, G., Pamer, E. G., Sander, C., & Xavier, J. B. (2013). Ecological modeling from time-series inference: Insight into dynamics and stability of intestinal microbiota. *PLOS Computational Biology*, *9*, 1–11.

Sugihara, G. (1994) Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *348*, 477–495.

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

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

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society (Series B)*, *58*, 267–288.

Ting, J., D'Souza, A., Vijayakumar, S., & Schaal, S. (2008) A Bayesian approach to empirical local linearization for robotics. Proceeding of the IEEE International Conference on Robotics and Automation (ICRA'08).

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

Van Kampen, N. (1992) *Stochastic processes in physics and chemistry*. North-Holland Personal Library. Elsevier Science, Amsterdam.

Wood, S. N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, *466*, 1102–1104. https://doi.org/10.1038/nature09319

Ye, H., Clark, A., Deyle, E., Munch, S., Keyes, O., Cai, J., White, E., Cowles, J., Stagge, J., Daon, Y., & Sugihara, G. (2017) rEDM: Applications of Empirical Dynamic Modeling from Time Series. R package version 0.6.5.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, *67*, 301–320. https://doi.org/10.1111/j.1467-9868.2005.00503.x

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of the article.

How to cite this article: Cenci S, Sugihara G, Saavedra S. Regularized S-map for inference and forecasting with noisy ecological time series. *Methods Ecol Evol*. 2019;10:650–660. https://doi.org/10.1111/2041-210X.13150