Check for updates

OPEN Circularity in fisheries data weakens real world prediction

Alfredo Giron-Nava[1,2,3], Stephan B. Munch[4,5], Andrew F. Johnson[6], Ethan Deyle[1], Chase C. James[1], Erik Saberski[1], Gerald M. Pao[1,7], Octavio Aburto-Oropeza[1] & George Sugihara[1 ✉]

The systematic substitution of direct observational data with synthesized data derived from models during the stock assessment process has emerged as a low-cost alternative to direct data collection efforts. What is not widely appreciated, however, is how the use of such synthesized data can overestimate predictive skill when forecasting recruitment is part of the assessment process. Using a global database of stock assessments, we show that Standard Fisheries Models (SFMs) can successfully predict synthesized data based on presumed stock-recruitment relationships, however, they are generally less skillful at predicting observational data that are either raw or minimally filtered (denoised without using explicit stock-recruitment models). Additionally, we find that an equation-free approach that does not presume a specific stock-recruitment relationship is better than SFMs at predicting synthesized data, and moreover it can also predict observational recruitment data very well. Thus, while synthesized datasets are cheaper in the short term, they carry costs that can limit their utility in predicting real world recruitment.

Faced with budget reductions for fisheries science and management worldwide, fisheries programs have experienced pressure to systematically replace or augment observational data programs with less expensive indirect data estimation programs that produce so called synthesized data[1,2]. These programs construct continuous time series from sparse observations of standing stock biomass (SSB) or catch data, using model-based estimates to filter noise from the raw observations and to fill in for times that were not directly sampled[3]. Thus, such estimates are necessarily the product of the model assumptions for how fish populations grow[4]. These synthetic time series are commonly used as a record of stock status and as input data for predicting recruitment (fisheries productivity) which in some cases can directly influence management strategies[5,6].

A problem arises when data that were synthesized with explicit assumptions about the stock-recruitment relationship (SRR) are then used to make predictions about that relationship as it occurs in the real world. Although models can often be made to fit and predict model-generated data, the accuracy of such models at predicting real world observations is often very low. There is potential for circularity in the overall approach[3,7], all of which casts doubt about whether a relationship between the SSB and recruitment in nature actually exists[8–10]. Recent studies that find evidence for such a relationship in principle, are beginning to question whether it can be used to improve our practical ability to predict recruitment[11,12]. For example, Pierre et al. find positive evidence of a causal relationship between stock size and recruitment, however, they also conclude that recruitment is largely unpredictable using classical models that are based on stock size alone[11]. This finding is supported by Deyle et al. who uses a non-parametric nonlinear EDM approach (Empirical DynamicModelling, see Box 1) to find that for Atlantic and Gulf Menhaden recruitment is indeed predictable from year to year, but only when allowing for realistic interdependence of adult stock size with other ecological factors[13]. More recently, Munch et al. analyzed a global database with EDM, to find that on average 40% of the variability in recruitment can be explained by previously observed recruitment fluctuations[12], which according to theory should contain information about the relevant environmental drivers[12,13].

Here, we examine the Ransom Myers database[14], a global repository of stock sizes and recruitment estimates for over 600 marine and freshwater fish populations (>100 species) to ask the following questions: (1) How well do standard fisheries models (SFMs) predict the number of recruits when the data are synthesized by an

[1]Scripps Institution of Oceanography, University of California San Diego, 9500 Gilman Dr, CA, La Jolla, 92093, USA. [2]National Center for Ecological Analysis and Synthesis, University of California Santa Barbara, 735 State St #300, Santa Barbara, CA, 93101, USA. [3]Future Earth, School of Global Environmental Sustainability, Colorado State University, 108 Johnson Dr, Fort Collins, 80523, CO, USA. [4]Fisheries Ecology Division. Southwest Fisheries Science Center. National Oceanographic and Atmospheric Administration, 110 Shaffer Rd, Santa Cruz, CA, 95060, USA. [5]Department of Ecology and Evolutionary Biology, University of California Santa Cruz, 1156 High Street, Santa Cruz, CA, 95064, USA. [6]MarFishEco, Portland, Oregon, USA, London, UK. [7]Salk Institute of Biological Sciences, 10010 North Torrey Pines Rd, La Jolla, CA, 92037, USA. ✉e-mail: gsugihara@ucsd.edu

assessment method that incorporates an explicit stock-recruitment model (synthetic data, SD) versus data that come directly from surveys or statistically denoised estimates such as those that come from a sequential population analysis that do not have density dependence built-in (direct data, DD)? And (2) can an equation-free EDM approach provide better predictions than SFMs? We selected all populations from this database with at least 25 years of both stock size and recruitment, representing 134 populations from 36 species, spanning 8 orders. The datasets were classified according to their origin as SD or DD, see *Materials and Methods* for further details.

BOX 1 Empirical dynamic modeling. Empirical Dynamic Modeling (EDM) is a data-driven approach to understand mechanisms, identify causal variables and make predictions on complex systems. It relies on extracting system behavior directly from observed time series[15–17]. The rationale for this approach is that dynamical relationships that are too complex or subtle to capture in a simple set of equations can instead be recovered empirically from time series observations[18]. EDM is specifically capable of describing nonlinear, state-dependent interactions (e.g. where the effect of one variable on another can depend on a third or fourth dynamically changing factor etc.). Such interactions are difficult to describe with models having constant coefficients and thus may not yield to standrad model-building approaches (e.g. those that use single-factor experiments to estimate fixed rate parameters)[10]. The essential ideas and concepts of EDM are briefly summarized in three 1-minute videos at http://tinyurl.com/EDM-intro.

The basic premise of EDM is to view a dynamic system from a geometrical perspective – the attractor. The attractor can be generated from some underlying set of equations or from observational time series data, and describes how variables change with respect to each other. Thus, as the system changes over time, its trajectory moves across different points in the state-space, and over time the paths form a geometric attractor (Fig. B1). A time series of a state variable can then be viewed as a projection of the motion on an attractor onto the coordinate axes recording how that variable changes through time. A natural way to view the system is in its native state space where each coordinate axis is an essential active causal variable – this primary form of EDM attractor reconstruction requires multiple time series, one for each active causal variable[16,19,20]. However, Takens theorem[21] also allows one to construct equivalent attractors using lags of just one time series. Here, time lags of a single variable can serve as proxies for the other variables (Fig. B1).
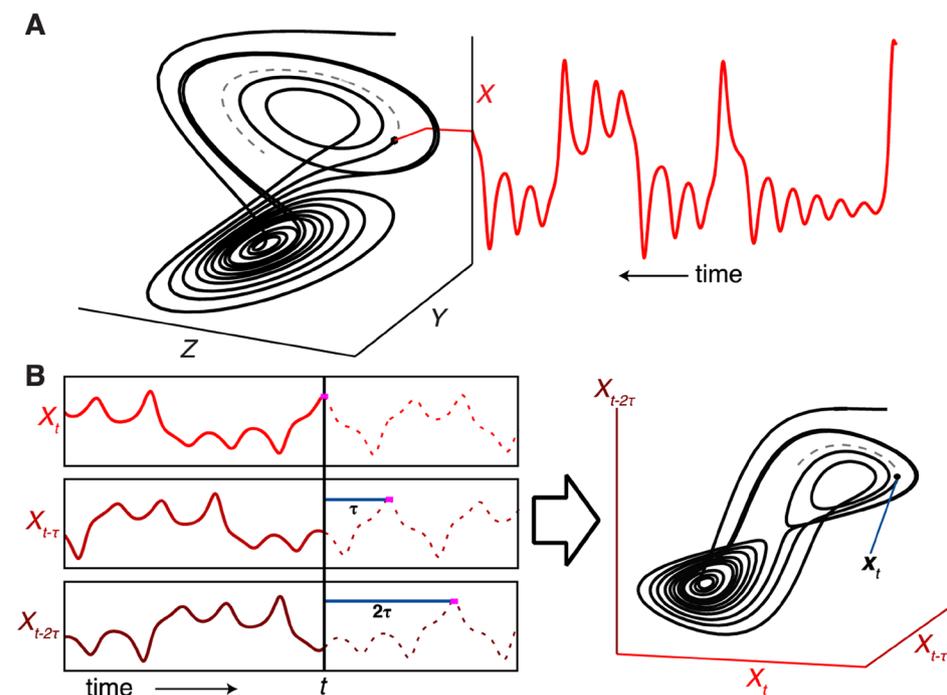

Figure B1. Reconstruction of System Dynamics from a Time Series (using the canonical Lorenz attractor example) (**A**) Projecting the motion of the Lorenz attractor onto the x-axis yields a time series for variable x. (**B**) Successive lags (with time step τ) of the time series x_t are plotted as separate coordinates to form a reconstructed (and visually similar) "shadow" attractor that preserves essential mathematical properties of the original system.

Again, the primary motivation for using EDM is to address nonlinear, non-equilibrium behavior. In contrast, many statistical methods (e.g. generalized linear models, structural equation modeling, dynamic linear models) rely on describing dynamics in terms of relationships that are constant or fixed in time, rather than in terms of relationships that are interdependent (state dependent) and that vary in time. For example, linear dynamics can occur with a static equilibrium when the dynamics are essentially viewed as stochastic perturbations around a fixed point, as in a typical linear stability analysis. Because nonlinear behavior appears to be ubiquitous in nature, EDM has been applied broadly across domains, from astrophysics[22], geophysics and climate[23], to epidemiology[24], cardiology[25] and medicine[26,27].
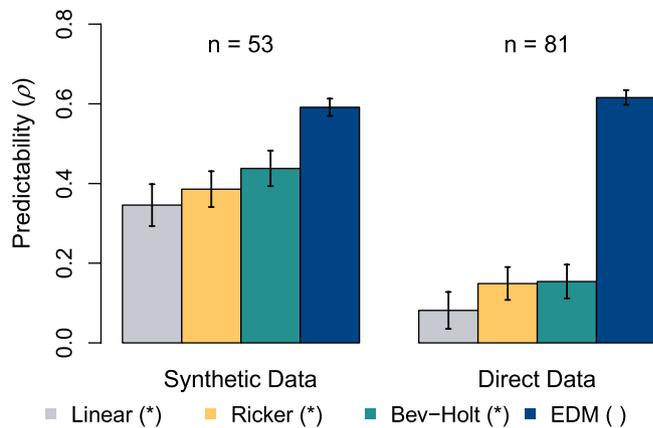
Figure 1. Potential circularity arises when models predict model output (SD). Comparison of predictability of spawner-recruit relationships for three standard fisheries models (Linear, Ricker, and Beverton-Holt) to equation-free EDM (S-maps). The y-axis represents the average predictability as measured by the correlation coefficient between 25 observed data points and their corresponding modeled predictions. The error bars represent the standard error. Asterisk labels next to the assessment method labels indicate a significant difference (P < 0.05) between thepredictability for SD as compared to DD.

Results

SFMs reported higher predictability (higher Pearson correlation coefficient between observed and predicted recruitment, ρ) when using SD as compared to DD. In the case of SD, the average Pearson correlation ρ values were equal to 0.35, 0.38 and 0.43 for the density independent, Ricker and Beverton-Holt models respectively (Fig. 1), with an overall average value of 0.39. In the case of the DD, average ρ values were equal to 0.08, 0.15 and 0.15respectively for the same models (Fig. 1), with an overall average value of 0.13. This represents an overall average decrease of 65% in the predictive capacity of SFMs when using DD versus SD. When confronted with real-world direct data, the reduction in predictability was significant (P < 0.05) for each of the three SFMs.

Regardless of the type of data used, EDM outperformed the three SFMs in predicting the number of recruits. On average, EDM reported Pearson correlation ρ values equal to 0.59 and 0.62 for SD and DD respectively (Fig. 1). We found no significant differences in the predictability of EDM between both types of data (P > 0.05).

Discussion

The use of synthetic data (SD) can be misleading when SFMs are used to study the stock-recruitment problem. Circularity arises when SD estimated through models is used as input into SFMs to predict SD.

Nonetheless, SD filtered with an explicit Stock-Recruitment model has been a critically useful assessment procedure when applied consistently. The majority of our stock assessment records have been derived this way, which allows us, importantly, to track stock status and productivity, albeit ultimately with a proxy for actual recruitment. Thus the consistent use of this procedure allows for historical comparisons to be made, which has ongoing merit in a management context. However, issues arise in terms of the original objective if we literally want to be able to make forecasts of real-world recruitment to generate management advice. The inability of SFM's to accomplish this original aim comes in part because the assumptions of the models are uncertain (e.g. the hypothesized functional form of the model, the classic assumption of near-equilibrium linear dynamics, etc. Box 1)[3,9,28,29]. Nonetheless the SFM's usefulness to provide an internally consistent record on stock status and productivity, though the ability to predict data it has produced should not be taken as validation of the models, and the SD should not be used in place of more direct observations in studies of fisheries recruitment.

As we move forward, if one of our main scientific management objectives is to improve our ability to predict recruitment as a real-world quantity, the potential problems with data circularity need to be recognized and constructively addressed, and alternatives such as EDM need to be considered. EDM derives the dynamic mechanisms and causes directly from the data – a capability that allowed it to accommodate the dynamics introduced to the SD by the SFM's (Fig. 1). More importantly, EDM was able to perform well with the DD and thus presumably is able to predict recruitment in the real fishery (Fig. 1). EDM methods have previously been used to improve recruitment predictability for a variety of stocks, for example: tuna in the North Pacific[30], sockeye salmon in the Fraser River system in British Columbia[10], red snapper in the Gulf of Mexico[31], and menhaden from the Gulf of Mexico and the Atlantic menhaden[13]. Although these results seem promising, the challenge remains to integrate these benchmark predictions more broadly into specific enacted management schemes, and particularly ones that are sustainably adaptive to non-stationary harvest targets[12].

Methods

Ransom myers database. We compared the predictions of the numbers of recruits through time from stock assessments using Standard Fisheries Models (SFMs)[4] and an Empirical Dynamic Modelling (EDM) technique known as S-maps[16,19,32]. To do this, we used the Ransom Myers database[14], a global repository of stock sizes and recruitment estimates for over 600 marine and freshwater fish populations (>100 species). All populations

from this database with at least 25 years of both stock size and recruitment data were included in our analysis, representing 134 populations from 36 species, spanning 8 orders. We classified each time series into one of two categories based on the nature of the data: data reanalyzed with an explicit Stock-Recruitment model (synthetic data, SD) (n = 53) and data from direct or statistically denoised observations (direct data, DD) (n = 81). The SD included datasets derived from Biomass Dynamic Models (BDM), while the DD included datasets derived from Sequential Population Analyses (SPA) or Direct Observations (DO). SPA datasets were classified as DD given that they only make an assumption about constant natural mortality to back calculate recruitment from landings data, which does not introduce an explicit assumption for the SR relationship. There were 3 datasets derived from Statistical Catch-at-Age (SCA) analyses that also met the requirement of having at least 25 years of data, however, because no information about whether explicit assumptions about the SR relationship was given, they were not included in our analyses. Supplementary Table I presents a summary of the original method used in each stock assessment, the classification as either SD or DD, and time series' length.

Predictability–standard fisheries models (SFMs). We evaluated the performance of three SFMs to predict the spawner-recruit relationship in the 134 populations from the Ransom Myers database: density independence, Ricker, and Beverton-Holt[4]. These models assume that the number of recruits is a function of the current stock size. All models can be written in the general form $R_t = \alpha S_t g(S_t)$, where R is recruitment, S is stock size, α is the maximum rate of reproduction, and $g(S_t)$ is a function that accounts for density-dependent processes[14]. In the case of the density-independent model, the function $g(S_t) = 1$ and the model is a straight line that intercepts the origin (0,0) with slope α. The Ricker and Beverton-Holt models introduce the term β, which is proportional to the product of fecundity and density-dependent mortality (see e.g. Quinn & Deriso[33]). The three models are presented below.

$$R_t = \alpha S_t - \text{Density independent} \tag{1}$$

$$R_t = \alpha S_t e^{-\beta S_t} - \text{Ricker} \tag{2}$$

$$R_t = \alpha S_t \left(\frac{1}{1 + \beta S_t} \right) - \text{Beverton} - \text{Holt} \tag{3}$$

The Ricker and Beverton-Holt models were fitted on a log scale, re-written so that $y_t = \ln[R_t/S_t]$[33]. All models were fit using the function 'fminsearch' in Matlab R2015b.

To calculate the predictability achieved by each model, we performed leave-one-out cross validation. The minimum length of any time series in our dataset was 25 years; thus, to ensure sample sizes were consistent across models, for time series with more than 25 points, we randomly selected 25 points as targets to be predicted. For each prediction on a target, the point one timestep before the target and 23 other, randomly selected points were used to fit model parameters, and used to make a prediction on the target. For time series with exactly 25 points, all points were used as targets to be predicted, and the other 24 points were used to fit model parameters on each iteration. Even though the Ricker and Beverton-Holt models were fitted on a log scale, all predictions were made in the original recruitment scale. We then calculated the predictability (ρ) as the Pearson correlation coefficient between the 25 observations and their respective predicted values.

Predictability–empirical dynamic modelling (EDM). EDM is based on the idea that time series are one-dimensional projections (a time record of some coordinate or variable) of a dynamic system (see introductory video https://youtu.be/fevurdpiRYg). If there are "n" relevant variables the trajectory produced as the system evolves in this n-dimensional space would produce a geometric shape or an "attractor". Following the trajectories at locations on an attractor nearby to a current state allows one to predict future states[15,16]. Because in practice we may not know what all the relevant variables are or even how many relevant variables there are (the "n" of the n-dimensional coordinate space), we can use Takens' Theorem[21] to construct a shadow version of the original attractor from a single time series or single variable that we want to predict. Thus, assuming that the single time series is x_t, one can reconstruct a "shadow" version of the original attractor by using lagged time series (eg. x_{t-1}, x_{t-2}) as proxies for other unknown time series of the same system and predict future values of x_t. Again, the number of time-lagged proxies required (the embedding dimension) corresponds to the number of active causal variables – or number of coordinates required to embed the attractor. The principles and mechanics of Takens' theorem and EDM are illustrated in Box 1 and further explained in Deyle and Sugihara[34] and Sugihara et al.[17] and in a series of short animations (http://tinyurl.com/EDM-intro).

Although it is possible to construct an attractor from a single time series (univariate reconstruction), predictability can often be improved by using multiple time series of different active causal variables measured from the same system (multivariate reconstruction)[16,35]. For example, for modelling fish stocks a useful multivariate reconstruction may involve a time series for fish stock biomass (S_t), another time series for the number of recruits (R_t), as well as time-lagged time series of both S_t and R_t. We tested all the possible combinations of these 6 time series ($R_t, R_{t-1}, R_{t-2}, S_t, S_{t-1}, S_{t-2}$) to make multivariate reconstructions, going from using 1 to 6 time series at a time.

This generalized embedding scheme is used to make forecasts using S-maps[32], which is a standard weighted kernel regression scheme that controls local weights with a tuning parameter θ. When $\theta = 0$ all the points on the attractor are equally weighted to generate a single global linear map. When $\theta > 0$ more weight is given to points

nearby each predictee on the attractor, so that the map produced for each forecast differs with location on the attractor (map varies with the system state). Finding that prediction improves for any $\theta > 0$ indicates curvature (nonlinearity) in the attractor. All our results report the predictability (ρ) achieved when θ is optimized.

Thus, as with the SFMs, to avoid overfitting we perform a leave-one-out cross validation by excluding the single time point that we are trying to predict from the data used to build the forecast model. We calculate ρ as the maximum Pearson correlation coefficient between the observations and their respective predicted values for each of the 134 analyzed fish stocks. All analyses were performed using the *rEDM* package in CRAN (*v. 0.7.2*).

Differences in predictability between SD and DD.   An unpaired t-test was used to test whether a particular model's predictions were significantly different when using SD versus DD. In Fig. 1 an asterisk indicates when the differences in predictability (ρ) between SD versus DD for each model type are significant ($P < 0.05$).

Data availability
Should the manuscript be accepted, the data supporting the results will be archived in an appropriate public repository such as Dryad or Figshare and the data DOI will be included at the end of the article.

References
1. Levine, C. R. *et al*. Evaluating the efficiency of environmental monitoring programs. *Ecol. Indic.* **39**, 94–101 (2014).
2. Lovett, G. M. *et al*. Who needs environmental monitoring? *Front. Ecol. Environ.* **5**, 253–260 (2007).
3. Brooks, E. N. & Deroba, J. J. When "data" are not data: the pitfalls of post hoc analyses that use stock assessment model output. *Can. J. Fish. Aquat. Sci.* **72**, 634–641 (2015).
4. Quinn, T. J. Ruminations on the development and future of population dynamics models in fisheries. *Nat. Resour. Model.* **16**, 341–392 (2003).
5. Mäntyniemi, S. Bayesian fisheries stock assessment: integrating and updating knowledge. (University of Helsinki (2006).
6. Punt, A. E. & Hilborn, R. Fisheries stock assessment and decision analysis: the Bayesian approach. *Rev. Fish Biol. Fish.* **7**, 35–63 (1997).
7. Anderson, S. C. *et al*. Improving estimates of population status and trend with superensemble models. *Fish Fish.* **18**, 732–741 (2017).
8. Myers, R. A. When do environment-recruitment correlations work? *Rev. Fish Biol. Fish.* **8**, 285–305 (1998).
9. Schindler, D. E. & Hilborn, R. Prediction, precaution, and policy under global climate change. *Science (80-.).* **347**, 953–954 (2015).
10. Ye, H. *et al*. Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proc. Natl. Acad. Sci.* **112**, E1569–E1576 (2015).
11. Pierre, M., Rouyer, T., Bonhommeau, S. & Fromentin, J.-M. Assessing causal links in fish stock-recruitment relationships. *ICES J. Mar. Sci.*1–9, https://doi.org/10.1093/icesjms/fsxw02 (2017).
12. Munch, S. B., Giron-Nava, A. & Sugihara, G. dynamics and noise in fisheries recruitment: A global meta-analysis. *Fish Fish.* 1–10, https://doi.org/10.1111/faf.12304 (2018).
13. Deyle, E. R., Schueller, A. M., Ye, H., Pao, G. M. & Sugihara, G. Ecosystem-based forecasts of recruitment in two menhaden species. *Fish Fish.* 1–13, https://doi.org/10.1111/faf.12287 (2018).
14. Myers, R. A., Barrowman, N., Hutchings, J. A. & Rosenberg, A. A. Population Dynamics of Exploited Fish Stocks at Low Population Levels. *Science (80-.).* **269**, 1106–1108 (1995).
15. Sugihara, G. & May, R. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **344**, 734–741 (1990).
16. Dixon, P., Milicich, M. & Sugihara, G. Episodic Fluctuations in Larval Supply. *Science (80-.).* **283**, 1528–1530 (1999).
17. Sugihara, G. *et al*. Detecting causality in complex ecosystems. *Science* **338**, 496–500 (2012).
18. DeAngelis, D. L. & Yurek, S. Equation-free modeling unravels the behavior of complex ecological systems. *Proc. Natl. Acad. Sci.* **112**, 3856–3857 (2015).
19. Deyle, E. R., May, R. M., Munch, S. B. & Sugihara, G. Tracking and forecasting ecosystem interactions in real time. *Proc. R. Soc. B Biol. Sci.* **283**, 20152258 (2016).
20. Chang, C.-W., Ushio, M. & Hsieh, C. Empirical dynamic modeling for beginners. *Ecol. Res.*, https://doi.org/10.1007/s11284-017-1469-9 (2017).
21. Takens, F. Detecting strange attractors in turbulence. In *Rand DA*, Young *LS, eds. Symposium on Dynamical Systems and Turbulence*. 366–381 (1981).
22. Kilcik, A. Nonlinear prediction of solar cycle 24. *Astrophys. J.* **693**, 1173 (2009).
23. Tsonis, A. A. *et al*. Dynamical evidence for causality between galactic cosmic rays and interannual variation in global temperature. *Proc. Natl. Acad. Sci.* **112**, 3253–3256 (2015).
24. Deyle, E. R., Maher, M. C., Hernandez, R. D., Basu, S. & Sugihara, G. Global environmental drivers of influenza. *Proc. Natl. Acad. Sci.* **113**, 13081–13086 (2016).
25. Sugihara, G., Allan, W., Sobel, D. & Allan, K. D. Nonlinear control of heart rate variability in human infants. *Proc. Natl. Acad. Sci.* **93**, 2608–2613 (1996).
26. Olde Rikkert, M. G. M. *et al*. Slowing Down of Recovery as Generic Risk Marker for Acute Severity Transitions in Chronic Diseases. *Crit. Care Med.* **44**, 601–606 (2016).
27. McBride, J. C. Sugihara causality analysis of scalp EEG for detection of early Alzheimer's disease. *Neuroimage Clinical.*, 258–265 (2015).
28. Lowerre-Barbieri, S. *et al*. Reproductive resilience: A paradigm shift in understanding spawner-recruit systems in exploited marine fish. *Fish Fish.* 285–312, https://doi.org/10.1111/faf.12180 (2016).
29. Glaser, S. M. *et al*. Complex dynamics may limit prediction in marine fisheries. *Fish Fish.* **15**, 616–633 (2014).
30. Glaser, S. 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.* **68**, 400–412 (2011).
31. Liu, H., Karnauskas, M., Zhang, X., Linton, B. & Porch, C. Forecasting dynamics of red snapper (Lutjanus campechanus) in the U.S. Gulf of Mexico. *Fish. Res.* **187**, 31–40 (2017).
32. Sugihara, G. Nonlinear Forecasting for the Classification of Natural Time Series. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **348**, 477–495 (1994).
33. Quinn, T. J. & Deriso, R. B. *Quantitative fish dynamics*. (Oxford University Pres (1999).
34. Deyle, E. R. & Sugihara, G. Generalized theorems for nonlinear state space reconstruction. *Plos one* **6** (2011).
35. Ye, H. & Sugihara, G. Information leverage in interconnected ecosystems: Overcoming the curse of dimensionality. *Science (80-.).* **353**, 922–925 (2016).

Author contributions

A.G.N. designed the study, performed the analyses and drafted the manuscript, G.S. suggested the study, helped guide the analysis, helped draft the manuscript and the discussion; A.F.J., E.D., S.M., E.S., G.P., O.A.O. all helped with the manuscript; E.D. and S.M. helped design the study, S.M. provided the data and the standard fisheries methods, C.C.J. prepared markdowns and supplementary materials, E.S. performed analyses, and O.A.O. and G.S. supported the study. All authors contributed to revisions.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at https://doi.org/10.1038/s41598-020-63773-3.

Correspondence and requests for materials should be addressed to G.S.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.