

Equation-free modeling unravels the behavior of complex ecological systems

Donald L. DeAngelis^{a,b,1} and Simeon Yurek^b

^a*Southeast Ecological Science Center, US Geological Survey, Gainesville, FL 32653; and* ^b*Biology Department, University of Miami, Miami, FL 33124*

Ye et al. (1) address a critical problem confronting the management of natural ecosystems: How can we make forecasts of possible future changes in populations to help guide management actions? This problem is especially acute for marine and anadromous fisheries, where the large interannual fluctuations of populations, arising from complex nonlinear interactions among species and with varying environmental factors, have defied prediction over even short time scales. The empirical dynamic modeling (EDM) described in Ye et al.'s report, the latest in a series of papers by Sugihara and his colleagues, offers a promising quantitative approach to building models using time series to successfully project dynamics into the future.

With the term “equation-free” in the article title, Ye et al. (1) are suggesting broader implications of their approach, considering the centrality of equations in modern science. From the 1700s on, nature has been increasingly described by mathematical equations, with differential or difference equations forming the basic framework for describing dynamics. The use of mathematical equations for ecological systems came much later, pioneered by Lotka and Volterra, who showed that population cycles might be described in terms of simple coupled nonlinear differential equations. It took decades for Lotka–Volterra-type models to become established, but the development of appropriate differential equations is now routine in modeling ecological dynamics. There is no question that the injection of mathematical equations, by forcing “clarity and precision into conjecture” (2), has led to increased understanding of population and community dynamics. As in science in general, in ecology equations are a key method of communication and of framing hypotheses. These equations serve as compact representations of an enormous amount of empirical data and can be analyzed by the powerful methods of mathematics.

However, mathematics has not had the “unreasonable effectiveness” in ecology that it has had in physics. Critics point out that

models in ecology have not passed the criterion of predictive ability (3, 4). There are many reasons for this, one being the highly nonlinear nature of ecological interactions. This has led to arguments over whether the “right” models are being used, but also to broad opinion that, unlike in physics, there are no right models to describe the dynamics of ecological systems, and that the best that can be done is to find models that are at least good approximations for the phenomena they are trying to describe. It is common in introductions of ecological modeling to find descriptions of the “modeling cycle,” in which a question is formulated, hypotheses are made, a model structure is chosen in the form of variables and equations, the equations are parameterized according to best information, and the model is analyzed and compared with patterns in nature. The cycle can be repeated again and again to obtain the best fit or validation by data. Although this methodology may work for some cases, it has become clear that there are limits on the accuracy of models applied to systems, with many variables interacting nonlinearly to create complex dynamics (e.g., marine ecosystems). Complete parameterization of models of such systems is nearly impossible. Furthermore, the nonlinearities of ecological systems cause models to be so sensitive to structure and parameters that even the most thoroughly and carefully developed model can hardly be expected to be predictive (5–7).

The message of Ye et al.'s report (1) is that there are ways to make predictive forecasts that do not rely on specifying equations at all, thus avoiding the problems outlined above. One can say generally that “the study of natural systems begins and ends with the specification of observables describing such a system, and a characterization of the manner in which these observations are linked” (8). In traditional models, equations are formulated with certain functional forms characterizing linkages of variables. For example, the Ricker difference equation model links the number of recruits to a fish population, R_t ,

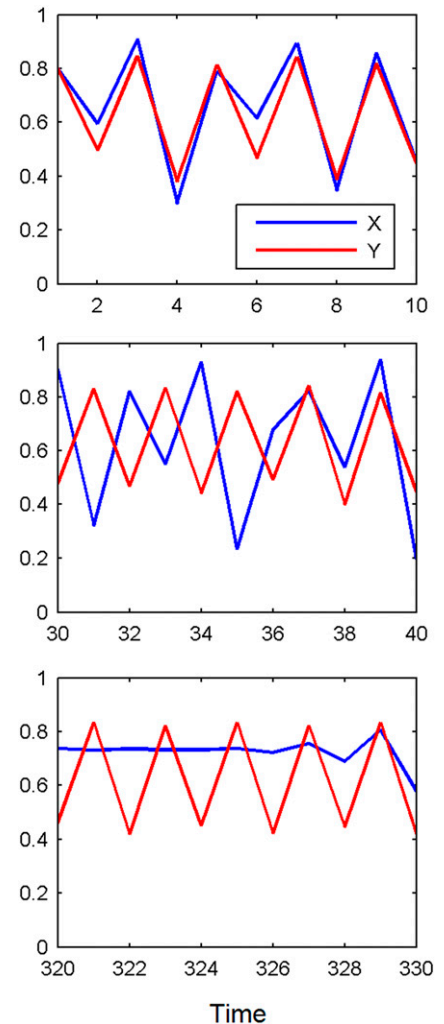


Fig. 1. Two-species competition model of Sugihara et al. (10) showing two populations going in and out of mirage correlations at different periods of the simulation.

to population size S_t through the relationship $R_t = S_t \exp(r - \alpha S_t)$, so that population size can be projected ahead to S_{t+1} at time $t + 1$. Given a set of time series data from a fishery, the modeler attempts to find values of r and α from the best fit for the equation. Most models are more complex, linking a population's

Author contributions: D.L.D. and S.Y. wrote the paper.

The authors declare no conflict of interest.

See companion article on page E1569.

¹To whom correspondence should be addressed. Email: don_deangelis@usgs.gov.

change through time to its own current size, the sizes of other populations, and relevant abiotic conditions. The functional forms describing these linkages can in principle be estimated based on measured correlations between variables over time. Although such correlations can in principle be extracted from data, meta-analyses on marine fisheries show that apparent correlations between populations and environmental variables have not stood up when additional data were analyzed (9). This is an important part of the argument of Ye et al. (1) in favor of equation-free modeling. To illustrate this point, Sugihara et al. (10) generated time series from a nonlinear model of two marine populations fluctuating through time, showing that what appear to be tightly positive correlations, over a period, suddenly switch to no correlation or even a negative correlation for another period (Fig. 1). Parameterizing equations based on these “mirage correlations” from one part of the time series would produce models that make false predictions.

The mirage correlations reflect deeper causes involving nonlinear interactions within the system. Rather than declaring analysis impossible, Ye et al. (1) show that there is an alternative that starts with the time series, without a presupposed set of equations. In this equation-free or nonparametric approach, no equation is assumed that needs to be parameterized. Instead, the dataset is used to build a model independent of equations. The basic idea is related to a question asked by Schaffer and Kot (11): “Do strange attractors govern ecological systems?” The authors showed that even time series that appear to be completely random could be following an underlying deterministic dynamic. The time map of any single population is an emergent property of interactions with other populations and abiotic conditions. Understanding the dynamics would require plotting all N -interacting variables in an N -dimensional phase space, something that is practically impossible for any real system because of the lack of data on most species in a community. However, if there is a sufficiently long time series for one population, a theorem of Takens (12) solves this problem. Takens showed that an N -dimensional phase portrait having the same dynamic properties as the portrait from the N -independent variables could be constructed by plotting the population size $x(t)$ versus $x(t+T)$, vs. $x(t+2T)$. . . vs. $x(t+(M-1)T)$, where T is a time delay. Using an appropriate value of T , a subset of these time-lag variables could be plotted

(or embedded) in phase space. An ecological time series plotted in this way will usually collapse to a geometric shape of dimensions less than the number of relevant variables would suggest. Often, three dimensions are sufficient to view the trajectories forming an “attractor,” a criterion for which is that the trajectories do not cross. What is important is that this preserves the essential features of the dynamics of the N variables. Sugihara

The broader message of the Ye et al. report is that science may be moving into a period where equations do not play the central role in describing dynamic systems that they have played in the last 300 years.

and his colleagues (10) made important additions to this embedding theorem of Takens; in particular, showing that it is possible to create a composite time series from ecologically similar species in the same geographic region to form a longer time series, and thus create a better model. Ye et al. (1) apply this approach to sockeye salmon stocks in the Fraser River. Dividing a time series of nine salmon stocks into “library data” used to calibrate, and “predict data” used to test the model, the authors show that their results are substantially better for most stocks than the parametric models, Ricker and extended Ricker. This approach allows projection into the future. One might quibble that it is well known that the Ricker is poor at describing population dynamics, so it is an easy target, but the comparison clearly demonstrates that, in this case, assuming

no equations is better than assuming some traditional equations.

The broader message of the Ye et al. (1) report is that science may be moving into a period where equations do not play the central role in describing dynamic systems that they have played in the last 300 years. This is largely the result of the rapidly increasing power of computers. Another sign of this trend is that, from a very different direction, two related classes of equation-free models have emerged as computational power increased. The methodology of individual- or agent-based modeling does not specify equations at the level of populations and communities, and instead simulates the individual organisms and their interactions with each other and the environment (5). A related approach, cellular automata models, is based on local laws determining transitions of spatial cells between different states. Although it would not seem intuitively that cellular automata models can describe biological systems, Wolfram (13), through his Principle of Computational Equivalence, has proposed that it can form a universal foundation for all phenomena of nature. Both individual- or agent-based models and cellular automata models are bottom-up approaches for constructing models of phenomena, and although deeply different conceptually from EDM, are similar in not specifying equations for population level variables.

The EDM approach demonstrates progress toward solving important but seemingly intractable problems of ecology; predicting behavior of complex nonlinear dynamic systems with limited information. This approach also suggests that, with high-performance computation, the study of dynamic systems is moving away from formulation and parameterization of equations and toward letting data directly determine the model. Because of the central role equations hold in science, it also raises questions: How will these changes affect the way scientists communicate, the way they understand, and the way they think?

- 1 Ye H, et al. (2015) Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proc Natl Acad Sci USA* 112:E1569–E1576.
- 2 May RM (2004) Uses and abuses of mathematics in biology. *Science* 303(5659):790–793.
- 3 Peters RH (1991) *A Critique for Ecology* (Cambridge Univ Press, Cambridge, UK).
- 4 Pilkey OH, Pilkey-Jarvis L (2007) *Useless Arithmetic: Why Environmental Scientists Can't Predict the Future* (Columbia Univ Press, New York).
- 5 Grimm V, Railsback SF (2005) *Individual-Based Modeling and Ecology* (Princeton Univ Press, Princeton, NJ).
- 6 Yodzis P (1988) The indeterminacy of ecological interactions as perceived through perturbation experiments. *Ecology* 69(2):508–515.

- 7 Wood SN, Thomas MB (1999) Super-sensitivity to structure in biological models. *Proc Biol Sci (B)* 266(1419):565–570.
- 8 Casti JL (1992) *Reality Rules: I. Picturing the World in Mathematics. The Fundamentals* (Wiley-InterScience, New York).
- 9 Myers RA (1998) When do environment-recruitment correlations work? *Rev Fish Biol Fish* 8(3):285–305.
- 10 Sugihara G, et al. (2012) Detecting causality in complex ecosystems. *Science* 338(6106):496–500.
- 11 Schaffer WM, Kot M (1985) Do strange attractors govern ecological systems? *Bioscience* 35(6):342–350.
- 12 Takens F (1981) Detecting strange attractors in turbulence. *Dynamic Systems and Turbulence*, eds Rand DA, Young LS (Springer, New York), pp 366–381.
- 13 Wolfram S (2002) *A New Kind of Science* (Wolfram Media, Inc., Champaign, IL).