Chaos is not rare in natural ecosystems

Tanya Rogers
National Marine Fisheries Service
Bethany Johnson
University of California, Santa Cruz
Stephan Munch (smunch@ucsc.edu)
University of California, Santa Cruz

Article

Keywords: chaotic dynamics, chaos, natural ecosystems, natural populations

DOI: https://doi.org/10.21203/rs.3.rs-888047/v1

Title: Chaos is not rare in natural ecosystems

Abstract: Chaotic dynamics are thought to be rare in natural populations, but this may be due to methodological and data limitations, rather than the inherent stability of ecosystems. Following extensive simulation testing, we applied multiple chaos detection methods to a global database of 175 population time series and found evidence for chaos in >30%. In contrast, fitting traditional one-dimensional models identified <10% as chaotic. Chaos was most prevalent among plankton and insects and least among birds and mammals. Lyapunov exponents declined with generation time and scaled as the -1/6 power of mass among chaotic populations. These results demonstrate that chaos is not rare in natural populations, indicating that there may be intrinsic limits to ecological forecasting and cautioning against the use of steady-state approaches to conservation and management.

Main

Chaos was introduced to ecology nearly 50 years ago[1,2] to provide an explanation for widespread fluctuations in abundance of natural populations. If common, chaos would offer the promise of short-term predictability while setting hard limits on long-term forecasting[3]. It would also mean that the "stable ecosystem" paradigm – the theoretical justification for linear statistical models of ecological dynamics[4] and steady-state management policies[5] – would need rethinking. However, despite considerable effort, the evidence for chaos in natural populations remains limited to a handful of examples (e.g.[6–9]); the most recent global meta-analysis of chaos concluded that only 1 out of 634 ecological time series was chaotic[10].

But, the apparent rarity of chaos in natural populations is a mystery. Nonlinear dynamics are common in ecological time series[11], and abiotic drivers of population dynamics are themselves chaotic[12]. Moreover, ecosystems involve tens to thousands of species, and large complex systems are prone to chaos[13–15]. In light of this, we hypothesize that the dearth of evidence for ecological chaos reflects methodological and data limitations, rather than genuine rarity. Although prototypical models of chaos are one-dimensional[1], using 1-d models to classify natural populations[10,16,17] treats ecological complexity (e.g. species interactions) as noise, thereby hindering chaos detection[18]. Importantly, flexible methods for detecting chaos developed outside ecology do not require a 1-d population model[19–22], though how most will perform on ecological time series is unknown. Moreover, compared with earlier studies[18,23] many more time series of sufficient length are now available – a critical factor for detecting chaos[24].

Here, we revisit whether chaos is, in fact, rare in ecological systems. We tested 4 modern algorithms for chaos detection (recurrence quantification analysis[19], permutation entropy[20], horizontal visibility graphs[21], the chaos decision tree[22]), and 2 classical methods of estimating Lyapunov exponents (direct[25] and Jacobian[24]) on data simulated with a variety of chaotic, periodic, and stochastic models in order to benchmark misclassification rates under ecologically-relevant time series lengths and levels of noise (Supplementary Text B-D, Fig. S1-6). We tested the generality of this classification accuracy using two additional suites of simulation models. Three methods had error rates greater than 0.5 and so were not pursued further (Table 1). We applied the remaining 3 methods (Jacobian, recurrence quantification, and permutation entropy) to time series from the Global Population Dynamics Database (GPDD)[26]. The GPDD aggregates 4471 time series from 1891 taxa. Previous analyses of the GPDD have concluded that the majority of these time series are too noisy to permit accurate modeling[27,28]. Therefore, we restricted our attention to the subset of the GPDD where chaos could be detected if present, i.e.

relatively long time series of good data quality without any major gaps (see Methods for more details). Applying these criteria produced a dataset of 175 time series representing 138 different taxa with between 30 and 197 observations. To confirm the prevalence of chaos among plankton in the GPDD, we analyzed an independent dataset of 34 additional zooplankton time series from 3 lakes with between 138 and 639 observations.

We then explored how Lyapunov exponents (LEs) varied among taxa and depended on intrinsic timescale (generation time), body size (mass), time series length (generations sampled), and embedding dimension (*E*). The LE measures the average rate of divergence between nearby points in phase space and is the most widely used index of chaos[29] (see Supplementary Text B for more details). We estimated LE values using a locally linear model fit to lags of the time series[30], where *E* is the number of lags needed to reconstruct the dynamics (see Supplementary Text A for more details).

Table 1. False negative and false positive rates for 6 chaos detection methods, ranked by reliability, across all simulated datasets (pooled test and 2 validation datasets, see Supplementary Text C for details and Table S11 for disaggregated results on each dataset), and rates of positive chaos detection in the empirical GPDD dataset using the 3 most reliable methods. Values in italics indicate misclassification rates >0.5.

Chaos detection method	False negative rate	False positive rate	GPDD fraction chaotic (number of series)
1. Jacobian LE	0.29	0.04	0.33 (58)
2. Recurrence quantification analysis	0.37	0.13	0.42 (74)
3. Permutation entropy	0.26	0.18	0.51 (89)
4. Direct LE	0.08	*0.66*	
5. Horizontal visibility algorithm	*0.62*	0.10	
6. Chaos decision tree	*0.73*	0.02	

Results

Across 3 independent classification methods, at least ⅓ of the GPDD time series were classified as chaotic (Table 1). The most conservative estimate (33%) was obtained with the Jacobian LE method, which was the most robust to process noise and *underestimated* the frequency of chaos in the presence of substantial observation error in our simulations (Supplementary Text D). We focus most of our remaining analyses on the Jacobian LE estimates.

Noise and nonstationarity can affect the classification of time series. Based on tests with low-d parametric models, these were found to be present in the GPDD[27,31]. However, the time series we selected for chaos detection only partially overlap these prior studies. Hence we needed to address the role that noise and nonstationarity play in our specific results. If noisy time series were being incorrectly classified as chaotic, we would expect a higher frequency of chaos among series with lower prediction accuracy. However, the fraction classified as chaotic by the Jacobian method did not vary with prediction R^2 (logistic regression, $p>0.1$, $n=175$, $df = 173$; Fig. S7A), and series with high prediction error did not have higher LEs (Fig. S7B). The frequency of chaos (33%) also did not change if only series with prediction $R^2 > 0.25$ were considered. So although

chaotic series were more variable than non-chaotic series (Fig. 1A), they were actually somewhat more predictable (Fig. 1B); hence observation error is not inflating the frequency of chaos.

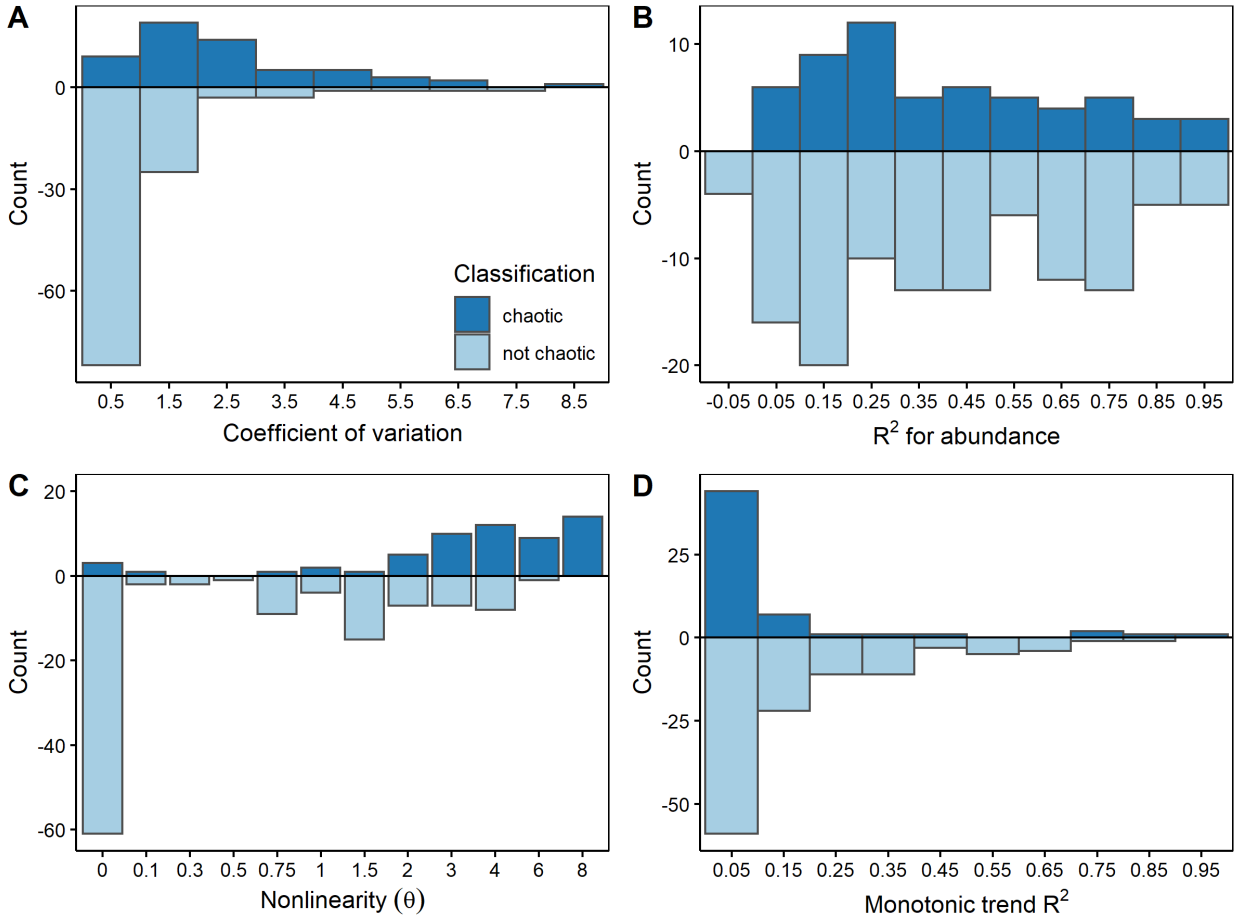


Fig. 1. Histograms showing the number of chaotic and non-chaotic time series in relation to (A) variability, as measured by the coefficient of variation; (B) predictability, as measured by the leave-one-out prediction R^2 for abundance; (C) nonlinearity, as measured by the local weighting parameter (θ)[30]; and (D) monotonic trend, as measured by the squared Spearman rank correlation coefficient. Horizontal axis labels give the midpoint of each bin with the exception of (C) which displays the discrete values that were used. Key in (A) applies to all panels.

If nonstationarity was driving the results, we should expect chaotic series to exhibit strong monotonic trends, exponential growth, or nearly linear dynamics. Only 6 time series that had either strong monotonic trends and/or near-linear dynamics were misclassified as chaotic (Fig. 1C,D). Reclassifying these series (4 birds, 1 mammal, 1 insect) as not chaotic reduced the frequency of chaos to 30%. The majority of chaotic series, however, were strongly nonlinear (Fig. 1C), did not display a strong monotonic trend (Fig. 1D, S8), and had a median growth rate near 0. Hence, nonstationarity and exponential growth are not responsible the observed frequency of chaos.

These observations in the GPDD time series are consistent with our simulations and previously published results[32]: The Jacobian method was *less* likely to find chaos as observation noise increased (Fig. S1-6), was minimally affected by process noise, rarely classified long term trends as chaotic, and effectively discriminated between chaos and stochastic linear dynamics

with seasonality. Taken together, these analyses indicate that the frequency of ecological chaos is not an artifact.

So, why is chaos more prevalent in our study than in previous meta-analyses (e.g.[10,16,17])? Most earlier analyses fit one-dimensional population models to classify series, whereas the methods used here make minimal assumptions about the dynamics. To evaluate this, we first constrained the Jacobian method to 1-d (E=1, see Methods). This reduced the apparent frequency of chaos from 33% to 9.1%, with reductions seen across all taxonomic groups (Fig. 2). Changes in classification were most common among populations in which the optimal E was high (Fig. S9), consistent with the hypothesis that reducing dimensionality inhibits chaos detection. Using the set of 1-d parametric models used in previous meta-analyses further reduced the apparent frequency of chaos to 6% or less (Supplementary Text E).

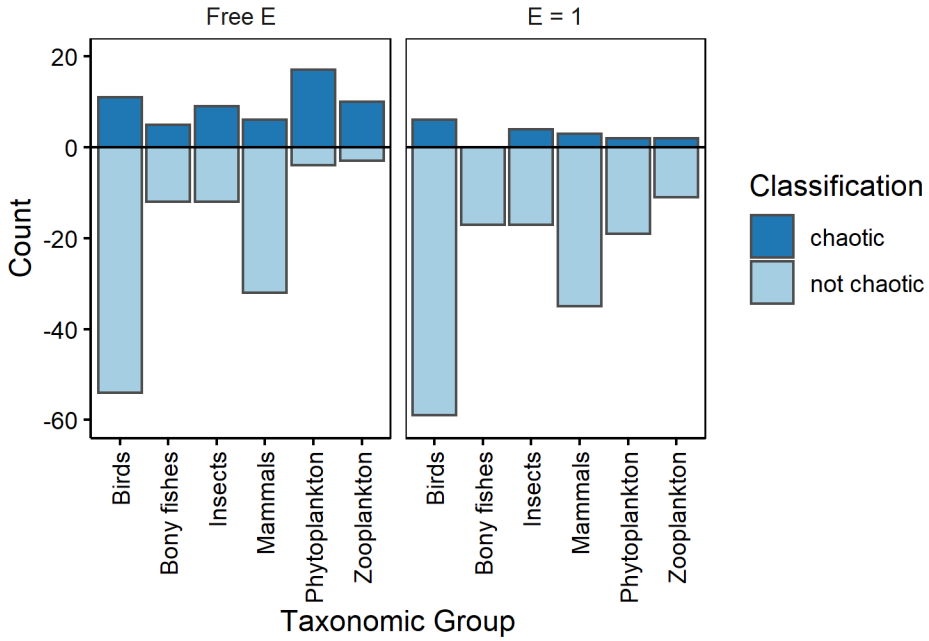


Fig. 2. Number of chaotic and non-chaotic time series by taxonomic group with unconstrained embedding dimension (Free E) and with embedding dimension fixed to 1 (E=1) using the Jacobian method.

Having allayed most reasonable qualms about statistical artifacts, we further explored the biological contexts in which chaos occurs. The frequency of chaos differed among taxonomic groups; phytoplankton had the greatest proportion of chaotic series (81%), followed by zooplankton (77%), insects (43%), bony fishes (29%), birds (17%), and mammals (16%) (Fig. 2). The prevalence of chaos decreased in species with longer generation times (Fig. 3A) which tended to have lower LEs as well (Fig. 3B). E also tended to decrease with increasing generation time (Pearson r = -0.31) and was lowest among birds (Fig. S10). Long-lived species definitionally have lower average mortality rates. Hence, on a per unit time basis (but not per generation), we might expect long-lived species to have relatively weaker interactions with other species, leading to lower LE and E compared with short-lived taxa. Long-lived species may also be better insulated from chaotic environmental drivers[17,33].

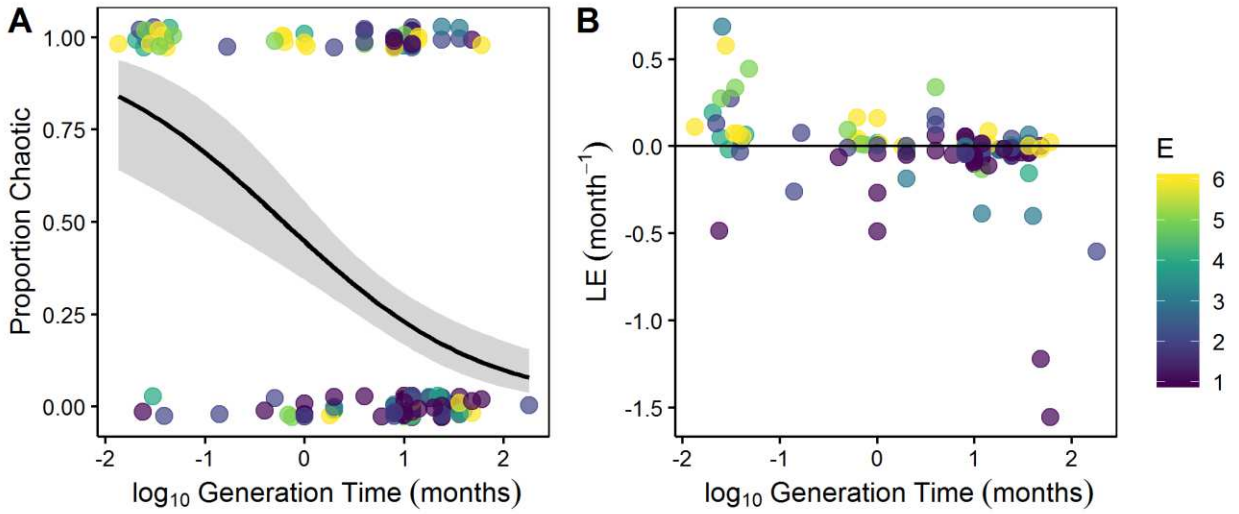


Fig. 3. Chaotic dynamics in relation to generation time. (A) Proportion of time series classified as chaotic using the Jacobian method, and (B) values of the Lyapunov exponent (LE) with color indicating embedding dimension (*E*). In (A), line is logistic regression and 95% confidence interval, and points are vertically offset by a random distance to reduce overplotting.

Data limitations might also account for the paucity of chaos in long-lived species, as chaos detection depends on the time series length relative to the intrinsic time scale for the system. To evaluate this effect, we artificially shortened the chaotic time series and reapplied the Jacobian method. When truncated to 30 data points, 42% of the chaotic time series were no longer classified as chaotic. Although not statistically significant, species with longer generation times were somewhat more likely to be reclassified as non-chaotic (logistic regression, p=0.38, df=46, n=48; Fig. S11). Data limitation increased false negatives for simulated data as well (Supplementary Text D, Fig. S1-5). Thus with more data, we expect to see more populations classified as chaotic.

Recent evidence indicates that LEs scale with body mass in experimental systems[34]. To determine how generally this applies to natural populations, we evaluated whether variation in LEs among chaotic species exhibited analogous mass (*M*) scaling. Combining LEs from our study and [34], we fit the model $\log_{10}(\text{LE})=a+b\log_{10}(M)$ for LE>0 and found evidence for consistent scaling with $b=-0.16$ (±0.02, p<0.001, df=74; Fig. 4). For LEs from the GPDD, including taxon or a taxon × mass interaction did not significantly improve the model (taxon+$\log_{10}(M)$ v. $\log_{10}(M)$, likelihood ratio=4.13, df=5, p=0.14; taxon*$\log_{10}(M)$ v. $\log_{10}(M)$, likelihood ratio=6.52, df=10, p=0.22) suggesting that differences in mass account for most of the average differences in LE among broad taxonomic groups. The consistency of LE scaling between lab and natural populations cannot readily be explained as a statistical artifact. Moreover, since the laboratory-derived LE's are not artifacts of observation noise or nonstationarity, consistent scaling with the field data provides additional evidence for chaos in natural populations.

The last meta-analysis to employ higher dimensional models and Jacobians[23] found evidence for chaos in 23% of 31 field time series, noting that this was likely an underestimate. This is close to our estimate of 33% and the difference may simply reflect an absence of plankton in their database. As the plankton in the GPDD are from relatively open marine systems, it is plausible that what appears here as chaos reflects advection of patchily distributed

populations. To address this, we evaluated the frequency of chaos and mass scaling of LEs in additional time series of lake zooplankton. The prevalence of chaos for the lake zooplankton was 47%. Among the chaotic taxa, none of these new LE estimates were significantly different from the mass scaling derived from the GPDD (deviation of observed values from regression predictions, $p>0.05$ for all; Fig. 4); It is exceedingly unlikely that advection would result in LEs that scale with mass consistently across 3 datasets, though it may contribute to the frequency of chaos.

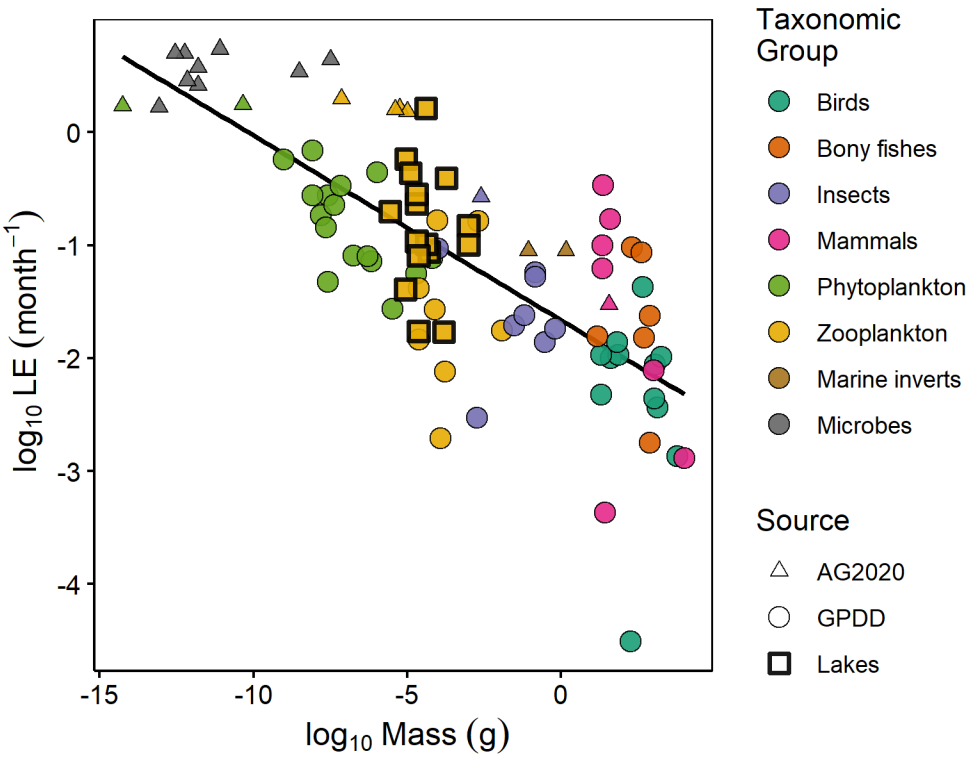


Fig. 4. Positive Lyapunov exponents (LEs) in relation to body mass, color distinguishing broad taxonomic groups. Includes data from this study (GPDD and supplemental results from 3 lake systems) and positive LEs compiled by [34] (AG2020). Note that the lake data (squares) were not used to fit the regression line.

Discussion

Single species models are routinely used to evaluate population status in applied fields such as fisheries[5] and conservation biology[35]. However, our results clearly show that scalar population models typically mischaracterize dynamics, treating complexity as noise, and leading to the conclusion that chaos is rare[10,16,17,36]. As May noted, such models "do great violence to reality"[37]. More flexible methods (e.g.[23,30,32]) are better able to characterize complex dynamics and integrating these into population status assessments is an important area for future research.

Reflecting on the frequency of chaos in natural populations, we note that birds and mammals, the least chaotic taxa, make up 59% of the time series we analyzed but represent less than 1% of the species on earth[38]; chaos may be considerably more common than the ⅓ presented here. Diseases, genetic variants, species, and statistical events are labeled "rare" using thresholds ranging from 0.001% to 5%. By these standards, chaos in natural ecosystems is far

from rare. This presents both challenges and opportunities for ecology as a predictive science; although short-term forecasting is feasible[39], precise long term prediction is likely to be impossible and management should avoid defining objectives in terms of equilibrium conditions. However, with increasing amounts of data and modern learning algorithms, new frontiers are open for characterizing the complex, non-equilibrium, and high-dimensional dynamics of ecology which will advance both our understanding of natural variability and improve our ability to manage ecosystems.

Methods

Data

We obtained abundance time series data from the Global Population Dynamics Database (GPDD)[26] accessed through the R package 'rgpdd'[40]. Our analyses required reasonably long and continuous time series and for organisms to be detected with sufficient frequency to reconstruct their dynamics. Consequently, we selected series with a reliability score of at least 2, at least 30 non-missing time points, at least 5 unique abundance values, less than 60% zeros, and less than 22% missing time points (in our dataset, this resulted in time series having no more than 11 missing values). We used only field-collected survey data (we excluded laboratory and harvest data), excluded human diseases, and excluded the shorter and lower-quality of 3 duplicate time series that passed our filtering.

Our final dataset contained 175 time series representing 138 different taxa. Of these series, there were 112 sampled annually, 53 monthly, 8 semiannually, and 2 bimonthly. There were 65 series from birds (Aves), 38 from mammals (Mammalia), 21 from insects (Insecta), 21 from phytoplankton (Bacillariophyceae, Dinophyceae), 17 from bony fishes (Osteichthyes), and 13 from zooplankton (Bivalvia, Crustacea, Echinoidea, Gastropoda, Polychaeta, Scyphozoa, Chaetognatha). Time series lengths ranged from 30 to 197 timesteps.

Prior to analysis, all untransformed abundance time series were rescaled to unit variance by dividing by the standard deviation. To allow for log transformations and calculations of population growth rate, $\ln(x_t/x_{t-\tau})$, all time series containing zeros were rescaled after adding a constant (1 if all values were integers, the minimum non-zero value if the series contained non-integers). Leaving the zeros intact and using only model forms that did not require log transformations produced similar results.

As a measure of organismal intrinsic timescale, generation time was obtained from published sources for all species in our dataset. We used the age at first reproduction as a proxy for generation time, unless direct estimates of generation time or doubling time were available. Wet body mass was obtained from published sources, or if unavailable, was estimated from volume assuming organisms have the same density as water. Generation time and mass data were not included for 7 taxa which were not finely resolved enough taxonomically to obtain this information. Sources for generation time and mass were the following: birds, mammals, fish, insects:[41]; diatoms:[42,43]; insect masses not included in [41]:[44–49]; copepods:[50–52]; dinoflagellates:[53,54].

The plankton data in the GPDD are marine and thus from relatively open systems. Hence, it is possible that their dynamics reflect water movement in addition to population growth. However, these time series display seasonal peaks and troughs that persist for months, rather than the more ephemeral fluctuations expected from water mass movement and it seems reasonable to assume these represent population dynamics over a large spatial area as opposed to

fluid dynamics. Nevertheless, to assess the robustness of these plankton results, we performed a supplemental analysis on 34 monthly zooplankton time series data from 3 lake systems which are arguably more "closed" than the marine environment. These systems were Lake Zurich (Wasserversorgung Zürich), Lake Geneva (© SOERE OLA-IS, AnaEE-France, INRA of Thonon-les-Bains, CIPEL, Dec 19 2019, developed by the Eco-Informatique ORE system of the INRA[55]), and Oneida Lake[56]. We also required these series to have less than 60% zeros. Mass data for these species were obtained from [44,57–60]. If only dry mass was available, we assumed dry mass was 20% of wet mass for arthropods and 4% of wet mass for rotifers[57,61].

Analysis

Our goal was to use modern and classical developments for detecting chaos to characterize ecological time series. However, most of these methods were developed in data-rich fields and tested on finely spaced time series with thousands to millions of observations. Therefore, we began by testing 6 chaos detection methods on simulated data from 37 stochastic, periodic, and chaotic models with ecologically relevant time series lengths and levels of observation and process noise. These simulations included both a test set of models used for tuning and novel set of models used for validation and evaluation of generality. The specific classification methods we tested were the "direct" method of estimating Lyapunov exponents (DLE)[25], the Jacobian method of estimating Lyapunov exponents (JLE)[24,30], recurrence quantification analysis (RQA)[19,62], permutation entropy (PE)[20], the horizontal visibility algorithm (HVA)[21,63], and the chaos decision tree (CDT)[22]. Note that the more traditional DLE and JLE have been tested previously[23,24,32,64,65], and we re-test them here for comparison with the newer methods. Supplementary Text A provides a brief background on time-delay embedding and a comparison of methods for selecting the embedding dimension and time delay which are used in many of the detection methods. Supplementary Text B provides the mathematical definition for Lyapunov exponent (LE) and full details on our implementation of each detection method. Supplementary Text C provides details on the simulation models, and Supplementary Text D summarizes results of the simulation testing.

Under the conditions of our simulations, DLE, HVA, and CDT had either false positive or false negative rates greater than 0.5 in both the test and validation datasets and so were not pursued further. We applied the remaining methods to the empirical dataset to estimate the frequency of chaos in natural populations.

The JLE method derives Lyapunov exponents (LEs) from the Jacobian matrices of a local linear time-delay embedding model, discounting for sampling variability, and classifies series with LE significantly >0.01 as chaotic (Supplementary Text B.2). This proved to be the most accurate index of chaos in the simulated test and validation datasets with the lowest false positive rate. Since the LE also is the most widely used index of chaos and provides a quantitative, scale-invariant measure of divergence rate, we used the numeric values of the LEs to further explore the relationship between chaotic dynamics, intrinsic timescale, body size, time series length, and embedding dimension in the GPDD dataset, after converting the LE from units of timestep^{-1} to units of month^{-1}. To calculate the mass scaling of the LE, we followed prior work[34,66] and used ordinary least squares regression on $\log_{10}$ transformed data for LE >0.

To test the effect of time series length on inferred LE (and subsequent classification), we truncated all time series that had been classified as chaotic to the last 30 observations and recomputed the LE. To test whether nonstationarity or long-term trends affected our results, we examined whether LEs were greater in series with stronger monotonic trends. We assessed the

degree of monotonic trend using the squared Spearman rank correlation between abundance and time. To test whether the restriction of dimensionality affects the inferred LE, we recomputed the LE with the embedding dimension set to 1. To test whether the restriction of model form, in addition to restricting dimensionality, affects the inferred LE, we fit a set of common 1-d population models with the form $x_{t+1} = x_t \exp[f(x_t, \mathbf{q})]$ to each time series and used the fitted model to estimate the LE (see Supplementary Text E for details).

References

1. May, R. M. Biological Populations with Nonoverlapping Generations: Stable Points, Stable Cycles, and Chaos. *Science* **186**, 645–647 (1974).
2. Beddington, J. R., Free, C. A. & Lawton, J. H. Dynamic complexity in predator-prey models framed in difference equations. *Nature* **255**, 58–60 (1975).
3. Hastings, A., Hom, C. L., Ellner, S., Turchin, P. & Godfray, H. C. J. Chaos in Ecology: Is Mother Nature a Strange Attractor? *Annu. Rev. Ecol. Syst.* **24**, 1–33 (1993).
4. Cressie, N. & Wikle, C. K. *Statistics for spatio-temporal data*. (John Wiley & Sons, 2011).
5. FAO. The State of World Fisheries and Aquaculture 2020. Sustainability in action. Rome. (2020).
6. Tilman, D. & Wedin, D. Oscillations and chaos in the dynamics of a perennial grass. *Nature* **353**, 653–655 (1991).
7. Turchin, P. & Ellner, S. P. Living on the Edge of Chaos: Population Dynamics of Fennoscandian Voles. *Ecology* **81**, 3099–3116 (2000).
8. Ferrari, M. J. *et al.* The dynamics of measles in sub-Saharan Africa. *Nature* **451**, 679–684 (2008).
9. Benincà, E., Ballantine, B., Ellner, S. P. & Huisman, J. Species fluctuations sustained by a cyclic succession at the edge of chaos. *Proc. Natl. Acad. Sci.* **112**, 6389–6394 (2015).
10. Sibly, R. M., Barker, D., Hone, J. & Pagel, M. On the stability of populations of mammals, birds, fish and insects. *Ecol. Lett.* **10**, 970–976 (2007).
11. Clark, T. J. & Luis, A. D. Nonlinear population dynamics are ubiquitous in animals. *Nat. Ecol. Evol.* **4**, 75–81 (2020).
12. Sivakumar, B., Berndtsson, R., Olsson, J. & Jinno, K. Evidence of chaos in the rainfall-runoff process. *Hydrol. Sci. J.* **46**, 131–145 (2001).
13. Gross, T., Ebenhöh, W. & Feudel, U. Long food chains are in general chaotic. *Oikos* **109**, 135–144 (2005).
14. Ispolatov, I., Madhok, V., Allende, S. & Doebeli, M. Chaos in high-dimensional dissipative dynamical systems. *Sci. Rep.* **5**, 12506 (2015).
15. Pearce, M. T., Agarwala, A. & Fisher, D. S. Stabilization of extensive fine-scale diversity by ecologically driven spatiotemporal chaos. *Proc. Natl. Acad. Sci.* **117**, 14572–14583 (2020).

16. Hassell, M. P., Lawton, J. H. & May, R. M. Patterns of Dynamical Behaviour in Single-Species Populations. *J. Anim. Ecol.* **45**, 471–486 (1976).

17. Shelton, A. O. & Mangel, M. Fluctuations of fish populations and the magnifying effects of fishing. *Proc. Natl. Acad. Sci. U. S. A.* **108**, 7075–7080 (2011).

18. Turchin, P. & Taylor, A. D. Complex Dynamics in Ecological Time Series. *Ecology* **73**, 289–305 (1992).

19. Webber, C. L. & Zbilut, J. P. Dynamical assessment of physiological systems and states using recurrence plot strategies. *J. Appl. Physiol.* **76**, 965–973 (1994).

20. Bandt, C. & Pompe, B. Permutation Entropy: A Natural Complexity Measure for Time Series. *Phys. Rev. Lett.* **88**, 174102 (2002).

21. Luque, B., Lacasa, L., Ballesteros, F. & Luque, J. Horizontal visibility graphs: Exact results for random time series. *Phys. Rev. E* **80**, 46103 (2009).

22. Toker, D., Sommer, F. T. & D'Esposito, M. A simple method for detecting chaos in nature. *Commun. Biol.* **3**, 11 (2020).

23. Ellner, S. P. & Turchin, P. Chaos in a Noisy World: New Methods and Evidence from Time-Series Analysis. *Am. Nat.* **145**, 343–375 (1995).

24. Nychka, D., Ellner, S., Gallant, A. R. & McCaffrey, D. Finding Chaos in Noisy Systems. *J. R. Stat. Soc. Ser. B* **54**, 399–426 (1992).

25. Rosenstein, M. T., Collins, J. J. & De Luca, C. J. A practical method for calculating largest Lyapunov exponents from small data sets. *Phys. D Nonlinear Phenom.* **65**, 117–134 (1993).

26. Prendergast, J., Bazeley-White, E., Smith, O., Lawton, J. & Inchausti, P. The Global Population Dynamics Database. Knowledge Network for Biocomplexity. doi:10.5063/F1BZ63Z8. (2010).

27. Thibaut, L. M. & Connolly, S. R. Hierarchical modeling strengthens evidence for density dependence in observational time series of population dynamics. *Ecology* **101**, e02893 (2020).

28. Knape, J. & de Valpine, P. Are patterns of density dependence in the Global Population Dynamics Database driven by uncertainty about population abundance? *Ecol. Lett.* **15**, 17–23 (2012).

29. Pikovsky, A. & Politi, A. *Lyapunov exponents: a tool to explore complex dynamics*. (Cambridge University Press, 2016).

30. Sugihara, G. Nonlinear forecasting for the classification of natural time series. *Philos. Trans. R. Soc. A Math. Phys. Eng. Sci.* **348**, 477–495 (1994).

31. Loh, J. *et al.* The Living Planet Index: using species population time series to track trends in biodiversity. *Philos. Trans. R. Soc. B Biol. Sci.* **360**, 289–295 (2005).

32. Kendall, B. E. Cycles, chaos, and noise in predator–prey dynamics. *Chaos, Solitons & Fractals* **12**, 321–332 (2001).

33. Anderson, C. N. K. *et al.* Why fishing magnifies fluctuations in fish abundance. *Nature*

452, 835–839 (2008).

34. Anderson, D. M. & Gillooly, J. F. Allometric scaling of Lyapunov exponents in chaotic populations. *Popul. Ecol.* **62**, 364–369 (2020).

35. IUCN. The IUCN Red List of Threatened Species. Version 2020-2. https://www.iucnredlist.org. (2020).

36. Freckleton, R. P. & Watkinson, A. R. Are weed population dynamics chaotic? *J. Appl. Ecol.* **39**, 699–707 (2002).

37. May, R. M. Simple mathematical models with very complicated dynamics. *Nature* **261**, 459–467 (1976).

38. Mora, C., Tittensor, D. P., Adl, S., Simpson, A. G. B. & Worm, B. How many species are there on Earth and in the ocean? *PLoS Biol.* **9**, e1001127 (2011).

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

40. Boettiger, C., Harte, T., Chamberlain, S. & Ram, K. rgpdd: R Interface to the Global Population Dynamics Database. https://docs.ropensci.org/rgpdd, https://github.com/ropensci/rgpdd. (2019).

41. Brook, B. W., Traill, L. W. & Bradshaw, C. J. A. Minimum viable population sizes and global extinction risk are unrelated. *Ecol. Lett.* **9**, 375–382 (2006).

42. Baars, J. W. M. Autecological investigations of marine diatoms, 2. Generation times of 50 species. *Hydrobiol. Bull.* **15**, 137–151 (1981).

43. Lavigne, A. S., Sunesen, I. & Sar, E. A. Morphological, taxonomic and nomenclatural analysis of species of Odontella, Trieres and Zygoceros (Triceratiaceae, Bacillariophyta) from Anegada Bay (Province of Buenos Aires, Argentina). *Diatom Res.* **30**, 307–331 (2015).

44. Anderson, D. M. & Gillooly, J. F. Physiological constraints on long-term population cycles: a broad-scale view. *Evol. Ecol. Res.* **18**, 693–707 (2017).

45. Janes, M. J. Oviposition studies on the chinch bug, Blissus leucopterus (Say). *Ann. Entomol. Soc. Am.* **28**, 109–120 (1935).

46. Cook, L. M. Food-Plant Specialization in the Moth Panaxia dominula L. *Evolution (N. Y).* 478–485 (1961).

47. Casey, T. M. Flight energetics of sphinx moths: power input during hovering flight. *J. Exp. Biol.* **64**, 529–543 (1976).

48. Kobayashi, A., Tanaka, Y. & Shimada, M. Genetic variation of sex allocation in the parasitoid wasp Heterospilus prosopidis. *Evolution (N. Y).* **57**, 2659–2664 (2003).

49. Hozumi, N. & Miyatake, T. Body-size dependent difference in death-feigning behavior of adult Callosobruchus chinensis. *J. Insect Behav.* **18**, 557–566 (2005).

50. Huntley, M. E. & Lopez, M. D. G. Temperature-Dependent Production of Marine Copepods: A Global Synthesis. *Am. Nat.* **140**, 201–242 (1992).

51. Cohen, R. E. & Lough, R. G. Length-Weight Relationships for Several Copepods

Dominant in the Georges Bank-Gulf of Maine Area. *J. Northwest Atl. Fish. Sci.* **2**, 47–52 (1981).

52. WoRMS Editorial Board. World Register of Marine Species. Available from http://www.marinespecies.org at VLIZ. Accessed 2020-11-01. doi:10.14284/170. (2020).

53. Nakamura, Y. Growth and grazing of a large heterotrophic dinoflagellate, Noctiluca scintillans, in laboratory cultures. *J. Plankton Res.* **20**, 1711–1720 (1998).

54. Boulding, E. G. & Platt, T. Variation in photosynthetic rates among individual cells of a marine dinoflagellate. *Mar. Ecol. Prog. Ser* **29**, 199–203 (1986).

55. Rimet, F. *et al.* The Observatory on LAkes (OLA) database: Sixty years of environmental data accessible to the public: The Observatory on LAkes (OLA) database. *J. Limnol.* **79**, (2020).

56. Rudstam, L. Zooplankton survey of Oneida Lake, New York, 1964 to present. kgordon.17.67 (https://knb.ecoinformatics.org/knb/metacat/kgordon.17.67/default). (2020).

57. Dumont, H. J., Van de Velde, I. & Dumont, S. The dry weight estimate of biomass in a selection of Cladocera, Copepoda and Rotifera from the plankton, periphyton and benthos of continental waters. *Oecologia* **19**, 75–97 (1975).

58. Geller, W. & Müller, H. Seasonal variability in the relationship between body length and individual dry weight as related to food abundance and clutch size in two coexisting Daphnia species. *J. Plankton Res.* **7**, 1–18 (1985).

59. Branstrator, D. K. Contrasting life histories of the predatory cladocerans Leptodora kindtii and Bythotrephes longimanus. *J. Plankton Res.* **27**, 569–585 (2005).

60. Rosen, R. A. Length-Dry Weight Relationships of Some Freshwater Zooplankton. *J. Freshw. Ecol.* **1**, 225–229 (1981).

61. Peters, R. H. & Downing, J. A. Empirical analysis of zooplankton filtering and feeding rates. *Limnol. Oceanogr.* **29**, 763–784 (1984).

62. Eckmann, J. P., Kamphorst, S. O. & Ruelle, D. Recurrence Plots of Dynamical Systems. *Europhys. Lett.* **4**, 973–977 (1987).

63. Luque, B., Lacasa, L., Ballesteros, F. J. & Robledo, A. Analytical properties of horizontal visibility graphs in the Feigenbaum scenario. *Chaos An Interdiscip. J. Nonlinear Sci.* **22**, 13109 (2012).

64. McCaffrey, D. F., Ellner, S., Gallant, A. R. & Nychka, D. W. Estimating the Lyapunov exponent of a chaotic system with nonparametric regression. *J. Am. Stat. Assoc.* **87**, 682–695 (1992).

65. Dämmig, M. & Mitschke, F. Estimation of Lyapunov exponents from time series: the stochastic case. *Phys. Lett. A* **178**, 385–394 (1993).

66. Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. & West, G. B. Toward a metabolic theory of ecology. *Ecology* **85**, 1771–1789 (2004).

67. Takens, F. Detecting strange attractors in turbulence. in *Dynamical Systems and Turbulence* (eds. Rand, D. A. & Young, L. S.) 366–381 (Springer, 1981).

68. Munch, S. B., Poynor, V. & Arriaza, J. L. Circumventing structural uncertainty: A Bayesian perspective on nonlinear forecasting for ecology. *Ecol. Complex.* **32**, 134–143 (2017).

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

70. Fraser, A. M. & Swinney, H. L. Independent coordinates for strange attractors from mutual information. *Phys. Rev. A* **33**, 1134–1140 (1986).

71. Kennel, M. B., Brown, R. & Abarbanel, H. D. I. Determining embedding dimension for phase-space reconstruction using a geometrical construction. *Phys. Rev. A* **45**, 3403–3411 (1992).

72. Marwan, N. How to avoid potential pitfalls in recurrence plot based data analysis. *Int. J. Bifurc. Chaos* **21**, 1003–1017 (2011).

73. Albers, D. J. & Hripcsak, G. Using time-delayed mutual information to discover and interpret temporal correlation structure in complex populations. *Chaos An Interdiscip. J. Nonlinear Sci.* **22**, 13111 (2012).

74. Van Kampen, N. G. *Stochastic processes in physics and chemistry*. vol. 1 (Elsevier, 1992).

75. Stark, J. Delay Embeddings for Forced Systems. I. Deterministic Forcing. *J. Nonlinear Sci.* **9**, 255–332 (1999).

76. Stark, J., Broomhead, D. S., Davies, M. E. & Huke, J. Delay Embeddings for Forced Systems.II. Stochastic Forcing. *J. Nonlinear Sci.* **13**, 519–577 (2003).

77. Ragwitz, M. & Kantz, H. Markov models from data by simple nonlinear time series predictors in delay embedding spaces. *Phys. Rev. E* **65**, 56201 (2002).

78. Kantz, H. & Ragwitz, M. Phase space reconstruction and nonlinear predictions for stationary and nonstationary Markovian processes. *Int. J. Bifurc. Chaos* **14**, 1935–1945 (2004).

79. Munch, S. B., Brias, A., Sugihara, G. & Rogers, T. L. Frequently asked questions about nonlinear dynamics and empirical dynamic modelling. *ICES J. Mar. Sci.* **77**, 1463–1479 (2020).

80. Abarbanel, H. D. I. *Analysis of Observed Chaotic Data*. *Institute for Nonlinear Science* (Springer New York, 1996).

81. Kantz, H. & Schreiber, T. *Nonlinear Time Series Analysis*. (Cambridge University Press, 2003).

82. Garland, J., James, R. G. & Bradley, E. Leveraging information storage to select forecast-optimal parameters for delay-coordinate reconstructions. *Phys. Rev. E* **93**, 22221 (2016).

83. Cheng, B. & Tong, H. On consistent nonparametric order determination and chaos. *J. R. Stat. Soc. Ser. B* **54**, 427–449 (1992).

84. Cenci, S., Sugihara, G. & Saavedra, S. Regularized S-map for inference and forecasting with noisy ecological time series. *Methods Ecol. Evol.* **10**, 650–660 (2019).

85. Kantz, H. A robust method to estimate the maximal Lyapunov exponent of a time series.

Phys. Lett. A **185**, 77–87 (1994).

86. Graham, D. W. *et al.* Experimental demonstration of chaotic instability in biological nitrification. *ISME J.* **1**, 385–393 (2007).

87. Benincá, E. *et al.* Chaos in a long-term experiment with a plankton community. *Nature* **451**, 822–825 (2008).

88. Becks, L. & Arndt, H. Different types of synchrony in chaotic and cyclic communities. *Nat. Commun.* **4**, 1359 (2013).

89. R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/. (2019).

90. Ushio, M. *et al.* Fluctuating interaction network and time-varying stability of a natural fish community. *Nature* **554**, 360–363 (2018).

91. Abarbanel, H. D. I., Brown, R. & Kennel, M. B. Local Lyapunov exponents computed from observed data. *J. Nonlinear Sci.* **2**, 343–365 (1992).

92. Ye, H., Clark, A., Deyle, E. & Munch, S. rEDM: Applications of Empirical Dynamic Modeling from Time Series. https://ha0ye.github.io/rEDM, https://github.com/ha0ye/rEDM. (2019).

93. Poincaré, H. Introduction. *Acta Math.* **13**, 5–7 (1890).

94. Riley, M. A. & Turvey, M. T. Variability and Determinism in Motor Behavior. *J. Mot. Behav.* **34**, 99–125 (2002).

95. Anderson, N. C., Bischof, W. F., Laidlaw, K. E. W., Risko, E. F. & Kingstone, A. Recurrence quantification analysis of eye movements. *Behav. Res. Methods* **45**, 842–856 (2013).

96. Madeo, D., Castellani, E., Santarcangelo, E. L. & Mocenni, C. Hypnotic assessment based on the Recurrence Quantification Analysis of EEG recorded in the ordinary state of consciousness. *Brain Cogn.* **83**, 227–233 (2013).

97. Karagianni, S. & Kyrtsou, C. Analysing the Dynamics between U.S. Inflation and Dow Jones Index Using Non-Linear Methods. *Stud. Nonlinear Dyn. Econom.* **15**, (2011).

98. Zbilut, J. P. Use of Recurrence Quantification Analysis in Economic Time Series. in *Economics: Complex Windows* (eds. Salzano, M. & Kirman, A.) 91–104 (Springer-Verlag, 2005).

99. Dippner, J. W., Heerkloss, R. & Zbilut, J. P. Recurrence quantification analysis as a tool for characterization of non-linear mesocosm dynamics. *Mar. Ecol. Prog. Ser.* **242**, 29–37 (2002).

100. Proulx, R., Côté, P. & Parrott, L. Multivariate recurrence plots for visualizing and quantifying the dynamics of spatially extended ecosystems. *Ecol. Complex.* **6**, 37–47 (2009).

101. Medvinsky, A. B. *et al.* Chaos far away from the edge of chaos: A recurrence quantification analysis of plankton time series. *Ecol. Complex.* **23**, 61–67 (2015).

102. Marwan, N., Carmenromano, M., Thiel, M. & Kurths, J. Recurrence plots for the analysis

of complex systems. *Phys. Rep.* **438**, 237–329 (2007).

103. Trulla, L. L., Giuliani, A., Zbilut, J. P. & Webber, C. L. Recurrence quantification analysis of the logistic equation with transients. *Phys. Lett. A* **223**, 255–260 (1996).

104. Marwan, N. CRP Toolbox 5.22. http://tocsy.pik-potsdam.de/CRPtoolbox. (2020).

105. Amigó, J. M., Zambrano, S. & Sanjuán, M. A. F. Combinatorial detection of determinism in noisy time series. *EPL (Europhysics Lett.* **83**, 60005 (2008).

106. Riedl, M., Müller, A. & Wessel, N. Practical considerations of permutation entropy: A tutorial review. *Eur. Phys. J. Spec. Top.* **222**, 249–262 (2013).

107. TOCSY - Toolbox for Complex Systems (Recurrence Plots, Cross Recurrence Plots, System Identification, ACE, Nonlinear Wavelet Analysis, Nonlinear Regression Analysis, Adaptive Filtering, Coupling Direction). https://tocsy.pik-potsdam.de/ (2020).

108. Lacasa, L., Luque, B., Ballesteros, F., Luque, J. & Nuño, J. C. From time series to complex networks: The visibility graph. *Proc. Natl. Acad. Sci.* **105**, 4972–4975 (2008).

109. Zhu, G., Li, Y., Wen, P. & Wang, S. Analysis of alcoholic EEG signals based on horizontal visibility graph entropy. *Brain Informatics* **1**, 19–25 (2014).

110. Suyal, V., Prasad, A. & Singh, H. P. Visibility-Graph Analysis of the Solar Wind Velocity. *Sol. Phys.* **289**, 379–389 (2014).

111. Yu, Z. G., Anh, V., Eastes, R. & Wang, D.-L. Multifractal analysis of solar flare indices and their horizontal visibility graphs. *Nonlinear Process. Geophys.* **19**, 657–665 (2012).

112. Mali, P., Mukhopadhyay, A., Manna, S. K., Haldar, P. K. & Singh, G. Multifractal analysis of charged particle distributions using horizontal visibility graph and sandbox algorithm. *Mod. Phys. Lett. A* **32**, 1750024 (2017).

113. Rong, L. & Shang, P. Topological entropy and geometric entropy and their application to the horizontal visibility graph for financial time series. *Nonlinear Dyn.* **92**, 41–58 (2018).

114. Lacasa, L. & Toral, R. Description of stochastic and chaotic series using visibility graphs. *Phys. Rev. E* **82**, 36120 (2010).

115. Iacobello, G. Fast Horizontal Visibility Graph (HVG) for MATLAB (https://www.mathworks.com/matlabcentral/fileexchange/72889-fast-horizontal-visibility-graph-hvg-for-matlab), MATLAB Central File Exchange. (2020).

116. Theiler, J., Eubank, S., Longtin, A., Galdrikian, B. & Doyne Farmer, J. Testing for nonlinearity in time series: the method of surrogate data. *Phys. D Nonlinear Phenom.* **58**, 77–94 (1992).

117. Jamšek, J., Paluš, M. & Stefanovska, A. Detecting couplings between interacting oscillators with time-varying basic frequencies: Instantaneous wavelet bispectrum and information theoretic approach. *Phys. Rev. E* **81**, 36207 (2010).

118. Schreiber, T. Extremely simple nonlinear noise-reduction method. *Phys. Rev. E* **47**, 2401–2404 (1993).

119. Gottwald, G. A. & Melbourne, I. A new test for chaos in deterministic systems. *Proc. R. Soc. London. Ser. A Math. Phys. Eng. Sci.* **460**, 603–611 (2004).

120. Gottwald, G. A. & Melbourne, I. Testing for chaos in deterministic systems with noise. *Phys. D Nonlinear Phenom.* **212**, 100–110 (2005).

121. Toker, D. The chaos decision tree algorithm. https://doi.org/10.6084/m9.figshare.7476362.v7. (2019).

122. Ahrestani, F. S., Hebblewhite, M. & Post, E. The importance of observation versus process error in analyses of global ungulate populations. *Sci. Rep.* **3**, 3125 (2013).

123. Francis, C. R. I. C., Hurst, R. J. & Renwick, J. A. Quantifying annual variation in catchability for commercial and research fishing. *Fish. Bull.* **101**, 293–304 (2003).

124. Kamarainen, A. M., Rowland, F. E., Biggs, R. & Carpenter, S. R. Zooplankton and the total phosphorus – chlorophyll a relationship: hierarchical Bayesian analysis of measurement error. *Can. J. Fish. Aquat. Sci.* **65**, 2644–2655 (2008).

125. Ricker, W. E. Stock and recruitment. *J. Fish. Board Canada* **11**, 559–623 (1954).

126. Hénon, M. A two-dimensional mapping with a strange attractor. *Commun. Math. Phys.* **50**, 69–77 (1976).

127. Zhang, H. *et al.* Complex Dynamics on the Routes to Chaos in a Discrete Predator-Prey System with Crowley-Martin Type Functional Response. *Discret. Dyn. Nat. Soc.* **2018**, 1–18 (2018).

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

129. Venkatesan, A. & Lakshmanan, M. Interruption of torus doubling bifurcation and genesis of strange nonchaotic attractors in a quasiperiodically forced map: Mechanisms and their characterizations. *Phys. Rev. E* **63**, 26219 (2001).

130. Ikeda, K. Multiple-valued stationary state and its instability of the transmitted light by a ring cavity system. *Opt. Commun.* **30**, 257–261 (1979).

131. Galias, Z. Rigorous investigation of the Ikeda map by means of interval arithmetic. *Nonlinearity* **15**, 1759–1779 (2002).

132. Guevara, M. R. & Glass, L. Phase locking, period doubling bifurcations and chaos in a mathematical model of a periodically driven oscillator: A theory for the entrainment of biological oscillators and the generation of cardiac dysrhythmias. *J. Math. Biol.* **14**, 1–23 (1982).

133. Timmer, J. Power of surrogate data testing with respect to nonstationarity. *Phys. Rev. E* **58**, 5153–5156 (1998).

134. Zhivomirov, H. A method for colored noise generation. *Rom. J. Acoust. Vib.* **15**, 14–19 (2018).

135. Ali, I., Saeed, U. & Din, Q. Bifurcation analysis and chaos control in discrete-time system of three competing species. *Arab. J. Math.* **8**, 1–14 (2019).

136. Hilborn, R. C. *Chaos and nonlinear dynamics: an introduction for scientists and engineers*. (Oxford University Press, 2000).

137. Yuan, S., Jiang, T. & Jing, Z. Bifurcation and chaos in the tinkerbell map. *Int. J. Bifurc.*

Chaos **21**, 3137–3156 (2011).

138. Turchin, P. & Hanski, I. An Empirically Based Model for Latitudinal Gradient in Vole Population Dynamics. *Am. Nat.* **149**, 842–874 (1997).

139. Costantino, R. F., Desharnais, R. A., Cushing, J. M. & Dennis, B. Chaotic Dynamics in an Insect Population. *Science* **275**, 389–391 (1997).

140. Edwards, C. A., Powell, T. A. & Batchelder, H. P. The stability of an NPZ model subject to realistic levels of vertical mixing. *J. Mar. Res.* **58**, 37–60 (2000).

Acknowledgments: We thank S. Salinas, S. Newkirk, A. Hein, N. Lustenhouwer, A.M. Kilpatrick, and M. O'Farrell for comments which improved the manuscript and C. Symons for assisting with access to the lake data. This work was supported by the NOAA Office of Science and Technology, SeaGrant number NA19OAR4170353, and the Lenfest Oceans Program.

Author contributions: All authors contributed to study design, simulations, data analysis, and writing. TLR made the figures.

Competing interests: Authors declare no competing interests.

Data and code availability: All data used are publicly available from the sources cited herein. The cleaned datasets and all analysis code used is available at https://github.com/tanyalrogers/chaos_GPDD, which will be made public upon publication.

Supplementary Text

A: Time-delay embedding and methods for selecting E and τ

Takens' theorem[67] provides a foundation for the analysis of time series from nonlinear systems. It demonstrates that under fairly generic conditions it is possible to reconstruct a system's dynamics using lagged observations from one (or a subset) of the state variables. Specifically, Takens[67] shows that for an M-dimensional dynamical system that converges to an attractor A of dimension $d < M$ that $\{x_t, x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-E\tau}\}$ is an embedding of A provided that $E > 2d$. What this means in practice is that for some variable x we can write $x_t = f(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-E\tau})$, use data on x to estimate f with some non-parametric regression or machine learning method (e.g.[64,68]), and the function f will retain all of the properties of the original system of which x is part (including chaos). This reconstruction of the state space from lags of a single time series is called time-delay embedding. The models for f most commonly used for ecological time series are piecewise constant (simplex)[69], locally linear (s-map)[30], Gaussian processes[68], and neural networks[24]. For computational speed and relative ease of implementation, we focused on s-map, modified as described below.

Of course, we usually do not know d or M, and some means of empirically determining the time delay τ and embedding dimension E is required. The direct LE method (DLE), Jacobian LE method (JLE), and recurrence quantification analysis (RQA) all employ time-delay embedding to reconstruct the state space, and so rely on a choice embedding parameters E and τ. To allow for a fair comparison, we used the same E and τ values in all 3 methods.

There are several standard and widely-used approaches for selecting optimal embedding parameters, including mutual information[70] or autocorrelation for selecting τ, and false nearest neighbors[71] or simplex projection[69] for selecting E. Since poor choices of τ and E can lead to incorrect conclusions about dynamics[72], we began by testing the effectiveness of several different procedures. We generated an *embedding dataset* consisting of 20 replicate time series (length $T = 100$) from each of nine different chaotic models (Table S1) with prescribed values of E and τ (each ranging from 1-3) and tested the ability of these procedures to correctly identify them. A description of the procedures tested and their performance is given below. Although mutual information is widely used in nonlinear time series analysis, it was not considered here because the data requirements, on the order of 10^4 observations, are far too high for ecological data[73]. The modified s-map regression method (section A.3) was the most accurate and so was used for the remainder of the analyses.

We note here that although Takens' theorem was originally derived for deterministic systems, it may provide a reasonable approximation for stochastic systems, although Van Kampen[74] showed that infinite memory is the general case for incompletely observed stochastic systems. Stark et al.[75,76], extended Takens' theorem to forced and stochastic systems, albeit with either severe constraints on admissible forcing (e.g. periodic drivers) or additional data requirements. Ragwitz and Kantz[77,78] showed that delay embedding for stochastic systems can sometimes be made explicitly Markovian, and Munch et al.[79] showed that delay embedding accurately reconstructed the conditional expectation for a class of stochastic population models, though neither of these results are generic. Ragwitz and Kantz[77,78] point out that although not exact, error from the finite-memory assumption in delay embedding for a stochastic nonlinear system is often small relative to the other sources of error.

A.1 Autocorrelation/partial autocorrelation

One method to select τ is to use the first zero crossing of the autocorrelation function [80,81] because this makes the coordinates $x_t, x_{t-\tau}$ linearly uncorrelated. In our analysis, we selected the first time lag at which either the autocorrelation function (ACF) or the partial autocorrelation function (PACF) switched from positive to negative, taking the smaller lag when they differed. When tested on the embedding dataset, this method accurately identified τ when its true value was 1 approximately 95% of the time but tended to underestimate τ when its true value was greater than 1. This is likely because the ACF/PACF method is intended for continuous time series and does not extend easily to discrete time series.

A.2 False nearest neighbor algorithm

The method of false nearest neighbors is commonly used to select the embedding dimension[71]. The basic idea is that if two points are true neighbors in a space of dimension E, then they will still be close in $E + 1$ dimensions. For each point in the time series, we found the closest neighbor in E dimensions and calculated the ratio of distances between the points in E and $E + 1$ dimensions. Neighbors were classified as false if the ratio exceeded $R_{tol}(= 15)$. Starting with $E = 1$, E is then increased until the proportion of false nearest neighbors is sufficiently close to zero or we reach some maximum E. In our simulations, this method consistently overestimated E even given the correct τ. This could be due to the method's sensitivity to the choice of R_{tol}, and it is difficult to choose a value that works well for many different systems.

A.3 Simplex and s-map regression

Another method for selecting embedding parameters is to fit models using multiple combinations of E and τ and to select those values that maximize the out-of-sample model fit[82]. We used leave-one-out prediction R^2 as our measure of model fit. Sugihara[30] recommends using simplex (i.e. a piecewise constant model) to select E and τ because it has no additional parameters. Unfortunately, multiple combinations of E and τ can give nearly identical model fits. For example, a periodic time series with a period of 4 can be perfectly described with E/τ combinations of 2/1, 4/1, 2/2, 1/4, 1/8, etc. Any apparent differences in model fit are due to numerical rounding error or observation noise, and algorithms which selected an apparent maximum gave inconsistent results across replicates of the same model.

To alleviate this problem, we made two modifications of the algorithm which substantially improved performance both for embedding parameter selection and chaos classification. The first was to evaluate model fit using local linear regression (s-map), rather than simplex, over a constrained grid of E, τ, and the local weighting parameter θ. Since the identifiable E scales as the square root of time series length[83] we considered E and τ ranging from 1 to 6, requiring $E^2 \leq T$ and $E\tau/T \leq 0.2$, and for each E and τ combination, considered 10 values of θ: 0, 0.1, 0.3, 0.5, 0.75, 1, 1.5, 2, 3, 4, 6, 8. The second modification was to ignore differences in prediction $R^2 < 0.01$ and to select from among equivalent "best" models the one that has the lowest τ, then θ, then E. This prioritizes simpler, more linear models, generally avoids models with no dynamics (e.g. $E = 1$ and τ equal to period length), and selects an optimal value for θ along with E and τ. These modifications together resulted in 100% correct identification of E and τ in the embedding dataset. It also led to more consistent identification of E and τ values across replicates in the test dataset, as well as more accurate classification of dynamics, particularly for periodic time series.

Recently, Cenci et al.[84] introduced the "regularized s-map" and showed that this method improved forecasting and Jacobian estimation. However a major conclusion of [84] was that the best weighting kernel and regularization depended on both the simulation model and whether one is trying to forecast or infer Jacobians. Unfortunately, conditions that optimize forecasting introduce bias in Jacobian estimation. In light of this, we used s-map without explicit regularization. However, we note that the selection criteria we used to choose among models with statistically negligible differences in forecast performance (i.e. favoring low E, low θ models) serves a very similar purpose and performed well in extensive simulations.

B: Chaos detection methods

B.1 Direct LE estimation (DLE)

The dominant Lyapunov exponent (LE), typically denoted as λ, is the most commonly used indicator of chaos[29]. The LE generalizes the dominant eigenvalue for linear systems and measures the rate at which nearby trajectories diverge or converge, averaged over the attractor. Specifically, an initial infinitesimal perturbation, $\Delta x(0)$, will grow or shrink approximately exponentially as $||\Delta x(t)|| \sim e^{\lambda t}||\Delta x(0)||$ where $\| \|$ is the Euclidean distance between points. Positive LEs are indicative of chaos (sensitive dependence on initial conditions) while negative LEs are indicative of stable (non-chaotic) dynamics.

This formal definition only holds in the limit as $\Delta x(0) \longrightarrow 0$. In multidimensional systems, the initial growth rate will be less than λ unless the initial perturbation happens to be in the direction of greatest growth. Moreover, since the attractor is bounded, two trajectories cannot get infinitely far apart and so the period of exponential growth will be finite. Nevertheless, it is possible, using this definition, to estimate the LE directly from data by measuring the divergence rate of nearest neighbors over a finite time horizon (the "direct" LE method)[25,85]. In ecology, this approach has primarily been used to characterize results for experimental systems (e.g.[86–88]).

For each point in the time series, we found its nearest neighbor in delay embedding space (with embedding dimension E and lag τ) and followed all pairs forward in time, computing the distance between them at each step. We regressed the log of the mean distance t-steps ahead, $\ln[d(t)]$, on the number of timesteps into future, t:

$$\ln[d(t)] = \ln[d(0)] + \lambda t + error$$

and use the slope, λ, as our estimate of the Lyapunov exponent. Since the attractor is bounded, the distance between points will ultimately saturate, so the maximum t must be set relatively low to avoid underestimating λ. In our implementation, we used t up to 4 steps into the future. Although this method does not depend on the details of the underlying dynamics and is fairly robust to choices of E and τ, it is known to be sensitive to noise[65]. Therefore, to ensure a conservative estimate of the frequency of chaos, only time series with $p(\lambda>0.01)>0.975$ were classified as chaotic, i.e. $\lambda - 1.96 \times SE(\lambda) > 0.01$.

These analyses were performed in R version 3.6.3[89].

B.2 Jacobian LE estimation (JLE)

The LE may also be estimated by fitting a model of the form

$$x_t = f(x_{t-\tau}, x_{t-2\tau}, \ldots, x_{t-E\tau})$$

to the available time series (for embedding dimension E and lag τ) and computing the LE from the model's Jacobian matrices[23,24]. A variety of methods may be used to estimate f (e.g. polynomials, splines, GAMs, neural networks, Gaussian processes). Following Ushio et al.[90], we used local linear regression (s-map)[30] to estimate f, obtaining embedding parameters E and τ and local weighting parameter θ as described in Supplementary Text A.3.

In order to find the best description of the dynamics, we fit 3 different forms of the delay embedding model (first difference as a function of abundance, population growth rate as a function of abundance, and population growth rate as a function of log abundance; Table S2) and selected the one with the best leave-one-out prediction R^2 for abundance. Previous meta-analyses using delay embedding (e.g.[23]) did not perform this model selection step, which may, in part explain the difference in our results.

For simulation models that did not generate strictly positive values, only the first model was fit; all 3 were considered for the remaining simulation models and all of the empirical data. We had also considered models of abundance as a function of abundance and log abundance as a function of log abundance, but these produced identical abundance predictions and LE estimates as models using first difference as a function of abundance, and population growth rate as a function of log abundance, respectively. We opted for the forms we used because they allowed us to also compute the leave-one-out prediction R^2 for growth rate or first difference. Although this quantity was not used to select the best model form, it offered another measure of predictability.

Although not strictly necessary, Jacobian matrices for all models were formulated in terms of abundance (as opposed to log abundance or growth rate). The Jacobians are constructed from the local regression coefficients (partial derivatives) from the model and (depending on the model formulation) the predicted growth rate and/or past observations of abundance.

The LE, λ, is computed by multiplying sequential Jacobian matrices and taking the log absolute value of the dominant eigenvalue (Λ_1) of this product. Formally, the LE is defined (and will converge) in the limit as $T \to \infty$, but for finite time series, it is calculated over the available data.

$$\lambda = \frac{1}{T} \ln \left| \Lambda_1 \left(\prod_{t=1}^{T} J(x_t) \right) \right|$$

In cases where $\tau > 1$, Jacobians every τ steps were multiplied, the LEs were divided by τ, and then the τ different LEs were averaged. The multiplication and eigen decomposition were done using the QR procedure for numerical stability. For a 1-d system (E = 1), the LE is simply the arithmetic mean of the log absolute value of the derivatives.

$$\lambda = \frac{1}{T} \sum_{t=1}^{T} \ln \left| \frac{\partial f}{\partial x_t} \right|$$

We found that the most accurate classifications could be obtained by computing the LE over several long sub-segments of the time series and using the variance among these to compute an approximate "confidence interval," rather than computing a single LE from the full time series. Local LEs computed from segments of a time series converge to the global LE as segment length increases. For sufficiently long segments, local LEs are approximately normally distributed around the global LE with variance inversely proportional to segment length[91]. From

the full Jacobian sequence of length $n = T - E\tau$, we computed LEs for all $i + 1$ possible sequences of length $n - i$ for $i = 3,4,5,6$. For each value of i, we then computed one-tailed 95% confidence intervals (as we are interested only in whether LEs are positive). As a further buffer against false positives, we took the minimum lower bound of these intervals as our estimate of the LE. If this value was >0.01, the time series was classified as chaotic. Note that intervals constructed in this way are conditional on E, τ, and θ, and ignore uncertainty in these embedding parameters. However, we are primarily using this as a classification rule, as opposed to a statistical statement about plausible values for the "true" Lyapunov exponent and, as with the other methods, we benchmarked the performance of this classifier using our simulations.

For empirical time series that contained missing values (35% of GPDD series), we fit the model using all data (skipping over delay vectors with missing elements), but computed the LE only over the longest string of consecutive, non-missing values, which encompassed 73% of non-missing data points on average (range: 35-98%), and was 35 timepoints on average (range: 12-108). Since we are using variability in LEs over subsegments to establish approximate confidence intervals, classification for shorter series is likely to be conservative. Consistent with this, all time series with a longest string <30 (8% of GPDD series) were classified as not chaotic.

These analyses were performed in R version 3.6.3[89]. The s-map models were fit using the package 'rEDM' version 0.7.4[92].

B.3 Recurrence quantification analysis (RQA)

Recurrence quantification analysis (RQA) is based on the notion that most deterministic systems tend to revisit regions of state space[93]. RQA has been used to test for chaos in physiological and financial data[19,94–98], but has only rarely been used in ecology[99–101].

RQA begins with the construction of the recurrence matrix, R, from the observed time series $\overrightarrow{x_t}$ for $t = 1, \ldots, T$. An entry of the $T \times T$ matrix, R, is 1 whenever two time points are within a threshold distance in state space and 0 otherwise. That is, $R_{t,s} = 1$ when $\|\overrightarrow{x_t} - \overrightarrow{x_s}\| \leq r$ and 0 otherwise, where $\| \|$ indicates Euclidean distance. We set r equal to 0.2 times the standard deviation of the data and used time-delay embedding was used to reconstruct the state space with the same embedding parameters E and τ as used for direct and Jacobian LE estimation (Supplementary Text A.3).

Several metrics can be derived from R [102]. After trying several, we selected 3 metrics that consistently partitioned dynamical regimes in our simulated data. Specifically, we used the RQA metrics "determinism," "entropy," and "average length," which are based on the distribution $P(l)$ of diagonal segments of length l contained in R. Determinism[19] measures the percentage of recurrence points which form diagonal lines greater than a threshold length, l_{min}, i.e.

$$DET = \frac{\sum_{l=l_{min}}^{T} lP(l)}{\sum_{l=1}^{T} lP(l)}$$

DET helps distinguish between determinism and noise. The RQA metric "entropy" is the Shannon entropy of $P(l)$, i.e.

$$ENTR = -\sum_{l=l_{min}}^{T} p(l) \ln p(l)$$

and measures the complexity of the deterministic structure. Chaotic time series tend to have higher *ENTR* than periodic time series. "Average length" is simply the mean length of diagonal line segments in R [103], i.e.

$$L = \frac{\sum_{l=l_{min}}^{T} lP(l)}{\sum_{l=l_{min}}^{T} P(l)}$$

L helps distinguish between chaotic and periodic dynamics because the reciprocal of L is related to the largest positive Lyapunov exponent[62].

Unfortunately, the ranges of *DET*, *ENTR*, and L that are best for classifying dynamics as periodic, chaotic, or stochastic are case-specific. We used the first 20 replicates of the test dataset to establish useful thresholds. Based on this analysis, a time series was classified as chaotic if $0.45 < DET < 0.99$ and $0.39 < ENTR < 2.3$ and $1.9 < L < 5.3$. These thresholds were used to classify the remaining simulations and the GPDD time series.

All RQA analyses were performed in MATLAB R2019a with the Cross Recurrence Plot (CRP) Toolbox version 5.22 (R32.4)[104]. We did not make any modifications to the toolbox and handled missing values according to the procedures built into the code. The CRP toolbox removes any missing values from the time series prior to analysis and performs the standard procedure on the remaining points.

B.4 Permutation entropy (PE)

Bandt and Pompe[20] introduced permutation entropy (PE) as a measure of time series complexity that can be used to distinguish between periodic, chaotic, and random time series. Pennekamp et al. (2019) recently used PE to quantify predictability of ecological time series.

The PE of a time series x_t for $t = 1, \dots, T$ is computed by creating embedding vectors of order E given by $\overrightarrow{x_t}(E) = \{x_{t-1}, \dots, x_{t-E}\}$. There are $E!$ possible permutations corresponding to the rank order of the values. Each embedding vector in the time series maps to one of the $E!$ possible permutations. The permutation entropy of the data is determined from the observed frequency distribution of the permutations, P_i, $i = 1, \dots, E!$ as

$$PE = -\frac{\sum_{i=1}^{E!} P_i \ln P_i}{E-1}.$$

PE tends to be high for stochastic time series, low for periodic time series, and intermediate for chaotic time series.

To use the PE as a classification tool in our analysis, we had to make choices regarding the order of embedding vectors and thresholds of classification. Although the rule of thumb is that E should be the largest value such that $5E! \leq N$ [105], PE is badly biased in short time series. To reduce this bias, we fixed $E = 3$ for all simulated and empirical time series. As with RQA, PE thresholds are operationally determined, and we used the first 20 simulations in the test set to determine thresholds for classifying time series as chaotic. Based on this, dynamics were classified as chaotic if the permutation entropy was between 1.06 and 1.23. We used this criterion to classify the remaining simulation data and empirical time series.

The permutation entropy computations were performed in MATLAB R2019a with the Toolboxes for Complex Systems (TOCSY) *petropy* function[106,107]. This function retains missing values when creating the embedding vectors. It ranks missing values higher than the numeric

values in the vectors, and if two missing values occur adjacently, the one that occurs first is given the lower rank.

B.5 Horizontal visibility graphs (HVG)

The visibility algorithm[108] constructs a mapping between a time series and a network, the horizontal visibility graph (HVG), which allows one to use network theory to characterize time series. While there have been several applications of the HVG in physiology[109], physics[110–112], and economics[113] there is, to our knowledge, no ecological application of this method.

To transform a time series into a network by means of the HVG, the following procedure is performed. Each point in the time series is a node on the visibility graph. Two nodes are connected based on whether they are "visible" to one another, i.e. if it is possible to draw a horizontal line between the two points in the time series without intersecting any points in the middle. Mathematically, this occurs when $x_t, x_s > x_n$ for all $t < n < s$. This criterion is checked for every pair of points in the time series to construct the HVG.

The visibility algorithm considers the "degree of connectivity" of each node in the HVG (i.e. the number of nodes each node is connected to). The degree distribution for white noise follows $P(k) \sim \exp(-zk)$ with $z = z_{un} = \ln(3/2)$ [21]. Correlated noise and deterministic dynamics deviate from this power law distribution with $z < z_{un}$ for chaos and $z > z_{un}$ for correlated noise[114]. The Shannon entropy of the HVG degree distribution

$$h = -\sum_{k=2}^{\infty} P(k) \ln P(k)$$

behaves like the LE with $h > \ln 4$ indicating chaotic dynamics[63]. Therefore, by using a combination of the deviation of the degree distribution from the power law (mean squared error, MSE) and the entropy of the degree distribution, we can differentiate chaotic dynamics from periodic and stochastic dynamics. As with RQA and PE, we used the first 20 replicates of the test dataset to tune the thresholds of *h* and MSE. Based on this analysis, time series were classified as chaotic if the entropy of the HVG was between 0.32 and 0.46 and the MSE was between 0.21 and 0.48. These thresholds were used to classify the remaining simulation data.

All visibility algorithm analyses were performed in MATLAB R2019a with the Fast Horizontal Visibility Graph (HVG) for MATLAB file exchange[115].

B.6 Chaos Decision Tree (CDT)

The chaos decision tree (CDT), proposed by Toker et al.[22], is a method that combines several tools into one algorithm to detect the presence or absence of chaos in time series. It has not been applied to ecological data. The CDT has the following procedure.

1. The CDT tests for stochasticity in the data by comparing the permutation entropy of the original time series to the permutation entropy of 1000 random Amplitude Adjusted Fourier Transform surrogates[116] and 1000 Cyclic Phase Permutation surrogates[117]. If the permutation entropy of the original time series falls within either of the surrogate distributions, the data are classified as stochastic. Otherwise, the algorithm proceeds.
2. The CDT de-noises the time series with Schreiber's noise-reduction algorithm[118].
3. The algorithm checks if the data are over-sampled. If so, the time series is downsampled by a factor of 2 and the process is iterated until it is not over-sampled.

4. Finally, the CDT performs a 0-1 test for chaos[119] modified to account for observation noise[120].

In our analysis, we used all of the default settings in the publicly available MATLAB code[121] and focused on the classification output of the algorithm. The classification output of the CDT is either "stochastic," "chaotic," "periodic," or "nonstationary". We aggregated the "stochastic," "periodic," and "nonstationary" time series as "not chaotic" in order to easily compare the output to the other chaos detection methods.

Analyses were performed in MATLAB R2019a with the Chaos Decision Tree algorithm[121].

C: Simulation testing

Several of the chaos detection algorithms we used originated in the nonlinear dynamics / physics literature where they were benchmarked using relatively noise-free datasets with thousands to millions of observations. Therefore, before analyzing the GPDD time series (which had from 30 to 197 observations), we tested the accuracy of each of the 6 methods (Supplementary Text B) on simulated datasets with limited time series lengths and relevant levels of noise. We did this with two sets of simulations which we refer to as the test and validation datasets. All simulated data were generated in MATLAB R2019a.

As the *test dataset*, we simulated data from 21 different models with periodic dynamics (Table S3), chaotic dynamics (Table S4), and stochastic dynamics (Table S5). For generality, we included both ecological models and more generic models[22] for which the dynamical regimes had been previously determined. For each model, we generated time series of 5 different lengths (T = 25, 50, 75, 100, and 250) crossed with 4 different levels of white observation noise (1, 10, 20, and 30% of the standard deviation of the data). This range of observation noise was chosen to cover the range that has been empirically estimated in several natural systems, which is about 1-17% in large ungulates[122] and 10-15% in fish and plankton[123,124] on average. We simulated 100 replicate time series for each combination of model, time series length, and observation error. Each replicate was started from random initial conditions, and for the chaotic and periodic models, the first 500 timepoints were discarded to avoid transients. For models with more than one state variable, we only used data from the first variable. The observation noise was lognormal for models that produced strictly positive values, and Gaussian otherwise. For the colored noise stochastic models, we followed Toker et al.[22] and did not add observation noise to them in order to avoid interfering with the power spectra used in the CDT.

We used each of the 6 methods to classify each simulated time series as "chaotic" or "not chaotic." We used the first 20 replicates of each model/length/noise combination in the test dataset to train the detection methods. This involved tuning threshold parameters to maximize classification accuracy for the RQA, PE, and HVG methods, and making methodological modifications to maximize classification accuracy for the JLE method.

Without modifying the methods or tuned threshold parameters, we then applied the trained methods to the remaining 80 replicates of the test dataset. We recorded the overall classification rates for each method in aggregate and for each model, time series length, and noise level.

We also generated two *validation datasets* using independent sets of periodic, chaotic, and stochastic models. Since the purpose of the validation datasets was to evaluate the robustness of our classification rules on a completely novel set of models, the thresholds established on the test data were applied without modification. Validation dataset #1 had the same characteristics (dynamical regimes, time series lengths, levels of observation noise, number of replicates) as the test dataset, but were generated using different models (Table S6-S8).

Note that, although the test dataset and validation dataset #1 include models with nonlinearity, process stochasticity, and observation error, it does not include models that have all three simultaneously. Therefore, to test whether classification accuracy extends to this more challenging case, we conducted as second validation test. Validation dataset #2 was constructed using 3 additional simulation models (Table S9-S10) with both periodic and chaotic dynamics, crossed with 4 levels of lognormal observation noise (1, 10, 20, 30%) and 4 levels of process noise (0, 10, 20, 30%). Process noise was generated with random perturbations to the dynamics in the discrete time model and random perturbations to a parameter in the continuous time models, and we measured the (relative) level of process noise as $\frac{1}{T}\sum_{i=1}^{T}\frac{\mathrm{std}\left(f(x_i,\theta)\right)}{f(x_i,\bar{\theta})}$ where T is the length of the time series, $f(x_i,\bar{\theta})$ is the i^{th} point in the time series with no noise and $\mathrm{std}\left(f(x_i,\theta)\right)$ is the standard deviation at the i^{th} point with noise included. We simulated 100 replicates for each combination of model, dynamical regime (periodic or chaotic), observation noise, and process noise, and used $T = 100$ for all time series.

D: Simulation results

All methods performed better with lower observation noise and longer time series lengths but differed in sensitivity and overall classification accuracy. JLE, RQA, and PE were the most reliable. DLE, HVG, and CDT had misclassification rates of 0.5 or more. Classification results for the simulated test and validation datasets are presented in Figs. S1-6. Performance was reasonably similar across individual models within a given dynamical regime (Figs. S3, S5). The methods successfully classified long term trends (e.g. RandomWalkTrend) and seasonal dynamics (e.g. SinForcedAR1) as non-chaotic.

Performance on validation dataset #1 was broadly similar to the results for the test data set (Table S11), indicating that the classification thresholds established during training generalize reasonably well to new simulation models. Under validation dataset #2, which included observation noise, process noise, and nonlinearity simultaneously, the performance of all methods was somewhat worse on average. Nevertheless, JLE and PE performed acceptably with overall false positive rates of 10 and 15%, respectively. To provide more context, typical values of observation and process noise in natural populations, as quantified by [122] for large ungulates, range from 0.1 to 17% for observation error (median 1.3%) and 0.6 to 25% for process error (median 2.8%). For fish and plankton time series, average estimates of observation noise are 10 to 15% [123,124]. In our simulations with 10% observation error and 30% process error, we obtained false positive rates of 0.040, 0.043, and 0.45 for JLE, PE, and RQA respectively, with corresponding false negative rates of 0.36, 0.22, and 0.28. Thus, we expect the frequency of chaos estimated for the GPDD, using JLE in particular, to be fairly conservative.

Results by method

The DLE method had the highest true positive rate but also had a very high false positive rate. This method frequently classified stochastic series as chaotic and also struggled to properly classify the periodic series. As this method is known to be sensitive to noise and to be limited to estimating positive LEs[65], this result is not unexpected.

The JLE method had the best overall performance of all of the methods and the best classification accuracy at short time series lengths. Increasing observation noise made it more likely for chaotic series to be classified as non-chaotic, with limited effect on the classification of non-chaotic series. Process noise had a similar effect and also did not increase the false positive rate of chaos detection.

RQA had a lower true positive rate than the Jacobian method and was more sensitive to time series length than to observation noise. RQA tended to misclassify periodic series as chaotic in the presence of high observation noise, but misclassification rates for stochastic series were relatively low. RQA showed the most sensitivity to process noise with higher false positive rates as process noise increased.

For PE, the false positive rate was fairly low and was relatively insensitive to observation noise and time series length. As with RQA, this method tended to misclassify periodic series as chaotic in the presence of high observation noise, but it was not highly sensitive to process noise. Stochastic series were misclassified as chaotic at short time series lengths.

HVG was insensitive to observation noise, and could correctly classify nearly all periodic series, but chaos detection was very sensitive to time series length. This method tended to classify stochastic series as chaotic at longer time series lengths, and chaotic series as non-chaotic at shorter time series lengths.

The chaos DT had the lowest false positive rate but rarely detected chaos when it was present. For long time series ($T = 250$), the algorithm performed well (comparably to JLE, with similar sensitivity to observation noise), but its ability to detect chaos declined rapidly as time series length decreased.

E. Classifying time series using parametric 1-d models

To test whether the restriction of model form, in addition to restriction of dimensionality, affects the inferred LE, we fit a set of 1-d models, $x_{t+1} = x_t \exp[f(x_t, \boldsymbol{q})]$, to each of the empirical (GPDD) time series, where x_t is population size and q is a vector of parameters (Table S12). The set of models included those used in other meta-analyses[10,16,17], and all were capable of generating chaos for some values of q. We estimated the parameters by minimizing

$$SS = \sum_{t=1}^{T-1} \left\{ \ln\left(\frac{x_{t+1}}{x_t}\right) - f(x_t, \boldsymbol{q}) \right\}^2$$

using fminsearch in MATLAB R2020b, ignoring zeros. In keeping with the Jacobian-based method for quantifying chaos, the LE for this model was estimated by taking the average of the absolute value of the derivative over the observed states, i.e.

$$\lambda = \frac{1}{T} \sum_{t=1}^{T} f(x_t, \boldsymbol{q}) + \ln\ |1 + x_t f'(x_t, \boldsymbol{q})|$$

where the prime denotes the derivative with respect to x. When process noise is present, this approach is more appropriate for characterizing dynamics than computing the LE for the deterministic skeleton[29].

The frequency of chaos in the empirical dataset was 6% or less using this set of 1-d models (Table S12).

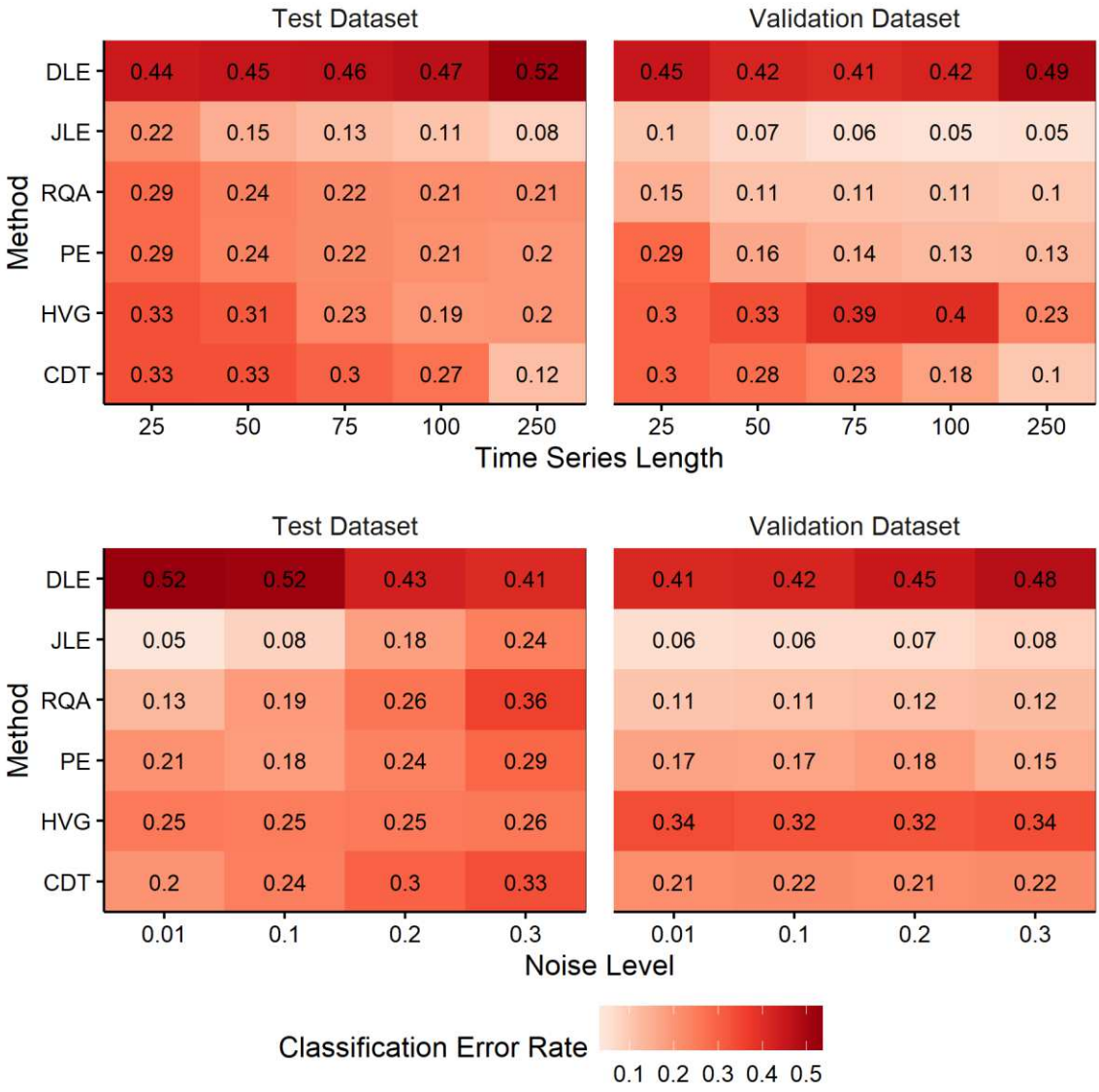


Fig. S1. Classification error rates for each chaos detection method, marginalized by time series length and noise level. Results for the test dataset and validation dataset #1 are shown. DLE = direct Lyapunov exponent, JLE = Jacobian-based Lyapunov exponent, RQA = recurrence quantification analysis, PE = permutation entropy, HVG = horizontal visibility graphs, CDT = chaos decision tree.

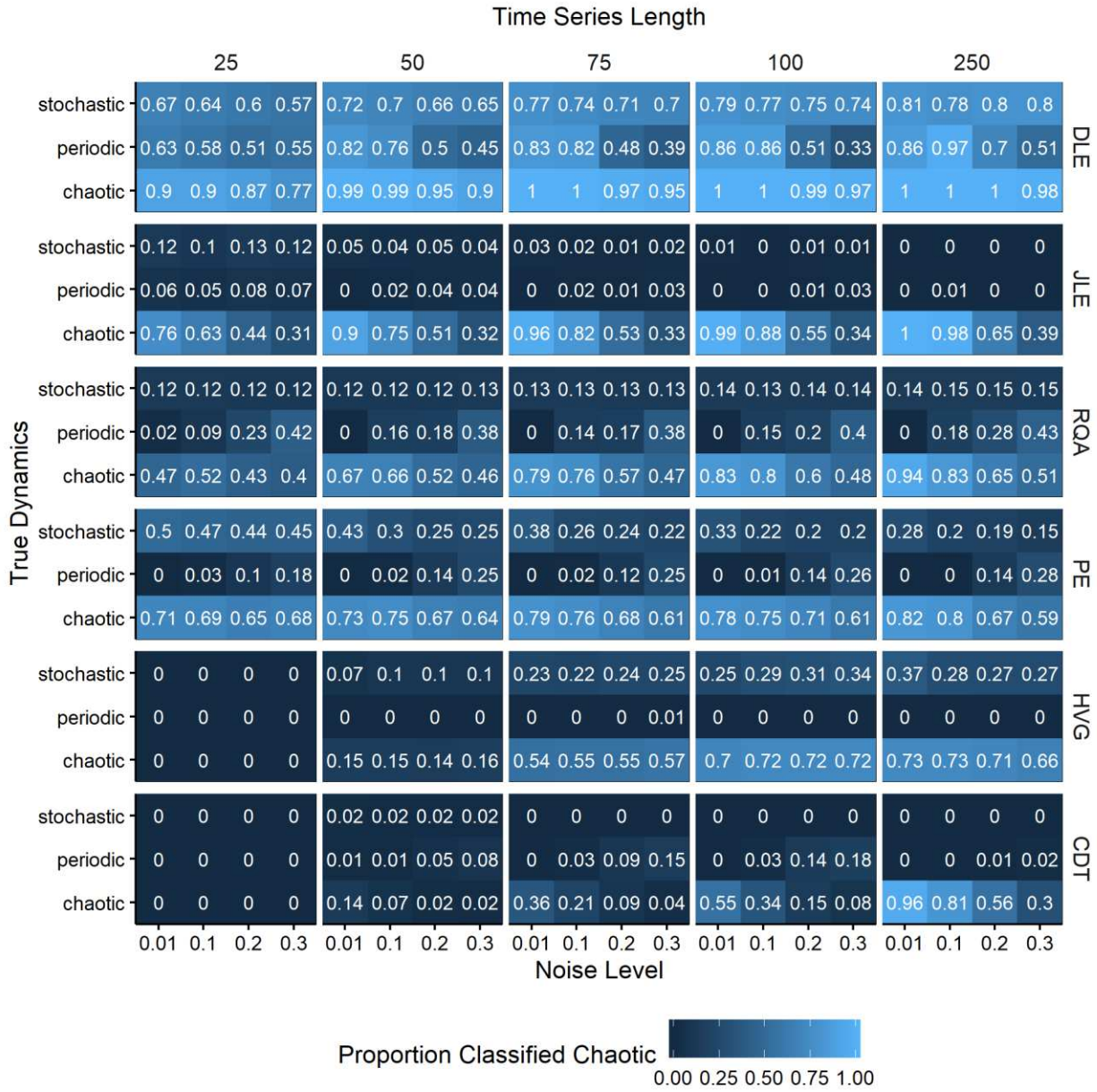


Fig. S2. Proportion of simulated time series from the *test* dataset classified as chaotic for all replicates of all models within each dynamical regime, for different levels of observation noise and time series length, for each chaos detection method. DLE = direct Lyapunov exponent, JLE = Jacobian-based Lyapunov exponent, RQA = recurrence quantification analysis, PE = permutation entropy, HVG = horizontal visibility graphs, CDT = chaos decision tree.

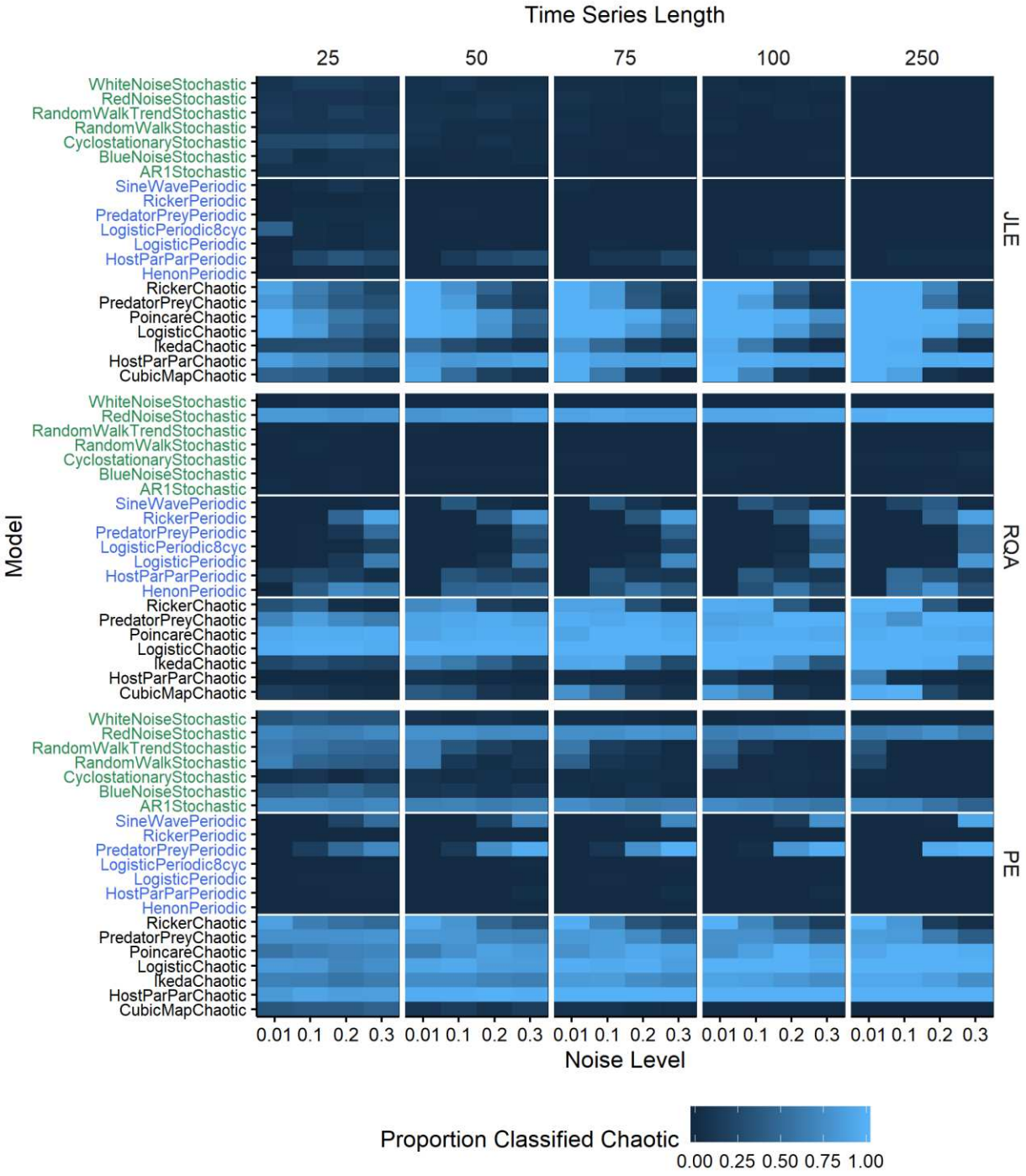


Fig. S3. Proportion of simulated time series from the *test* dataset classified as chaotic for all replicates of each individual model, for different levels of observation noise and time series length. Results for the 3 most reliable chaos detection methods are shown. JLE = Jacobian-based Lyapunov exponent, RQA = recurrence quantification analysis, PE = permutation entropy.

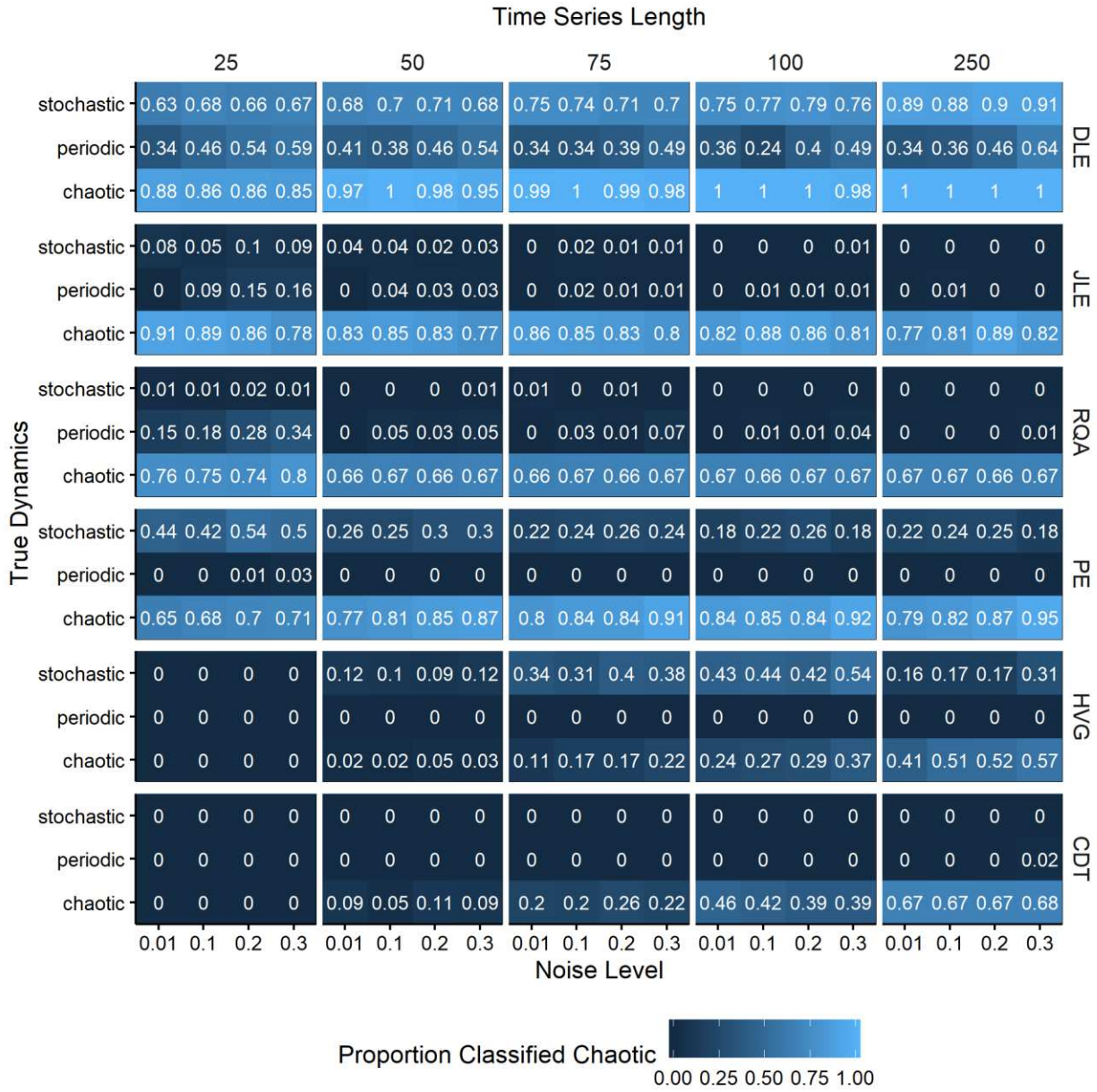


Fig. S4. Proportion of simulated time series *validation* dataset #1 classified as chaotic for all replicates of all models within each dynamical regime, for different levels of observation noise and time series length, for each chaos detection method. DLE = direct Lyapunov exponent, JLE = Jacobian-based Lyapunov exponent, RQA = recurrence quantification analysis, PE = permutation entropy, HVG = horizontal visibility graphs, CDT = chaos decision tree.

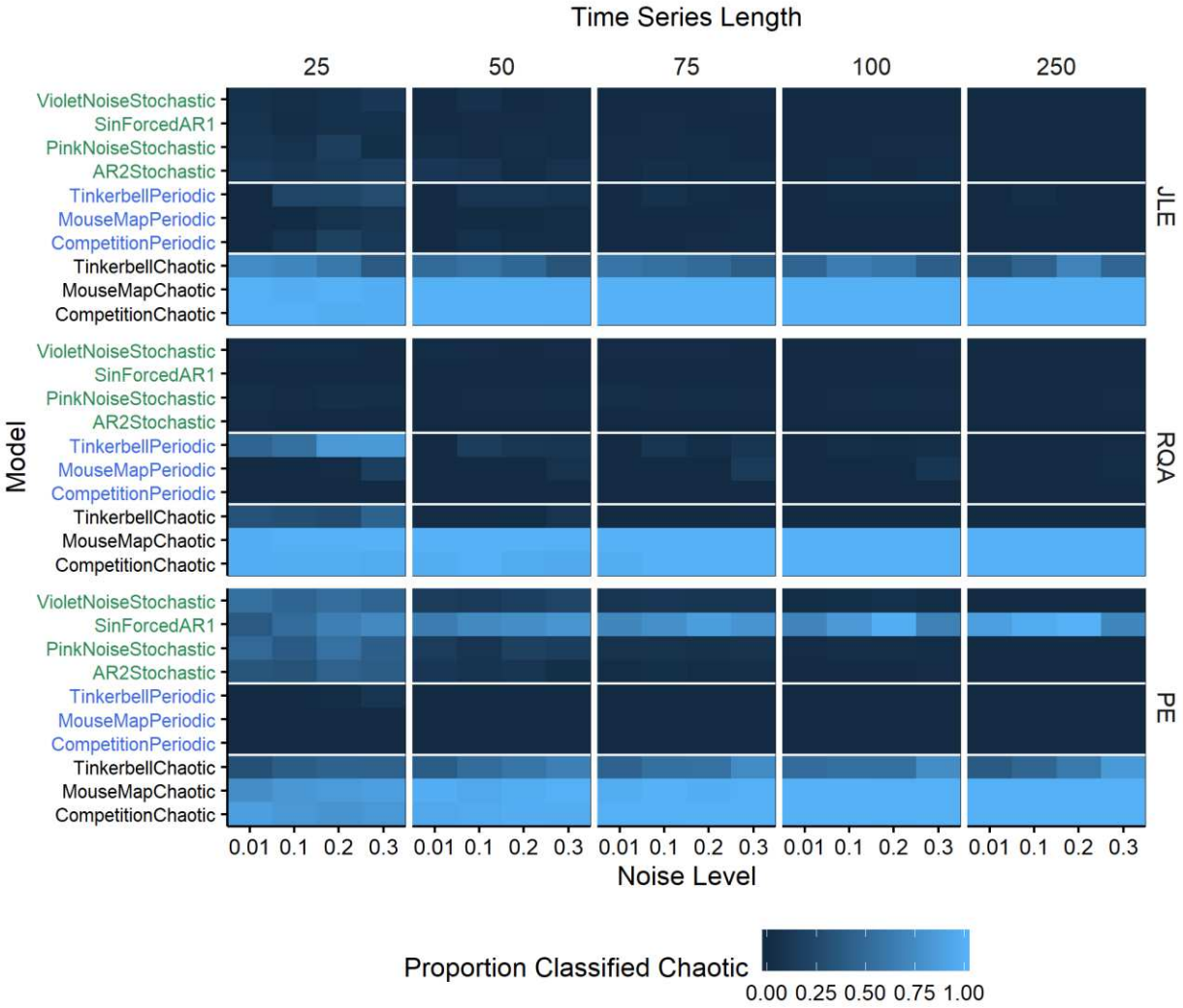


Fig. S5. Proportion of simulated time series from *validation* dataset #1 classified as chaotic for all replicates of each individual model, for different levels of observation noise and time series length. Results for the 3 most reliable methods are shown. JLE = Jacobian-based Lyapunov exponent, RQA = recurrence quantification analysis, PE = permutation entropy.

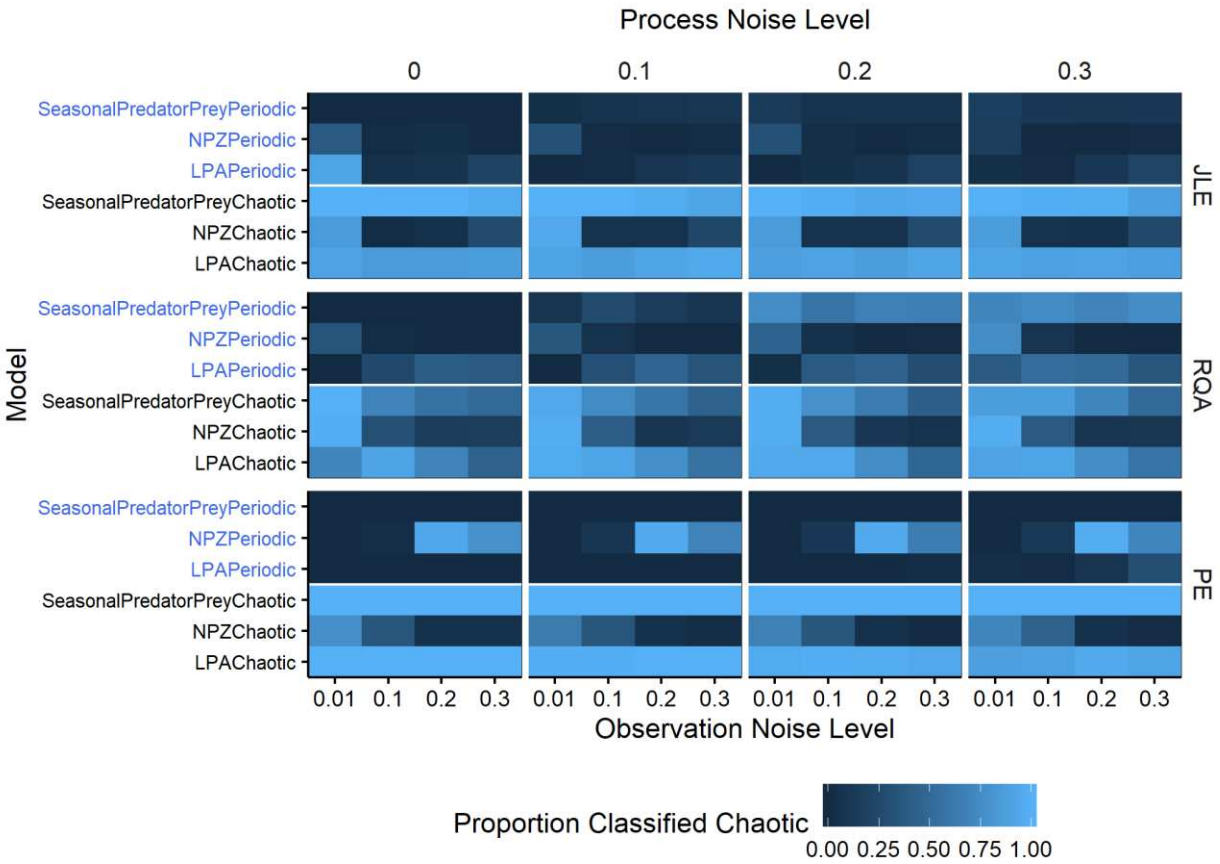
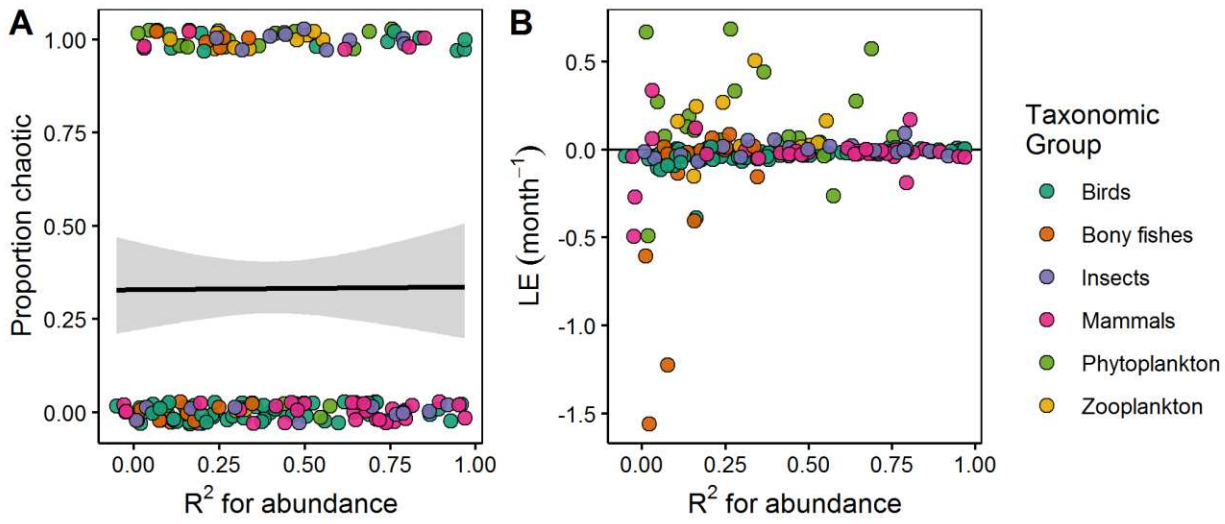


Fig. S6. Proportion of simulated times series from *validation* dataset #2 classified as chaotic for all replicates of each individual model, for different levels of observation noise and process noise. Results for the 3 most reliable methods are shown. JLE = Jacobian-based Lyapunov exponent, RQA = recurrence quantification analysis, PE = permutation entropy.



1056

1057 **Fig. S7.** Chaotic dynamics in relation to predictability. (A) Proportion of time series classified as
1058 chaotic using the Jacobian method and (B) values of the Lyapunov exponent (LE) with color
1059 indicating taxonomic group. In (A), line is logistic regression and 95% confidence interval, and
1060 points are vertically offset by a random distance to reduce overplotting.

1061

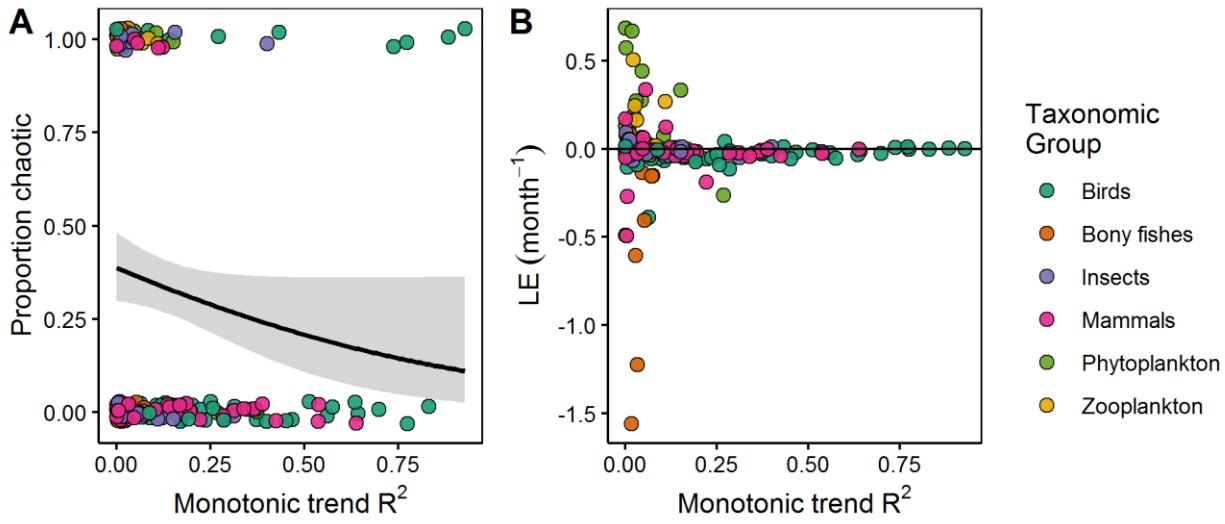


Fig. S8. Chaotic dynamics in relation to monotonic trend, as measured by the squared Spearman rank correlation coefficient. (A) Proportion of time series classified as chaotic using the Jacobian method and (B) values of the Lyapunov exponent (LE) with color indicating taxonomic group. In (A), line is logistic regression and 95% confidence interval, and points are vertically offset by a random distance to reduce overplotting.

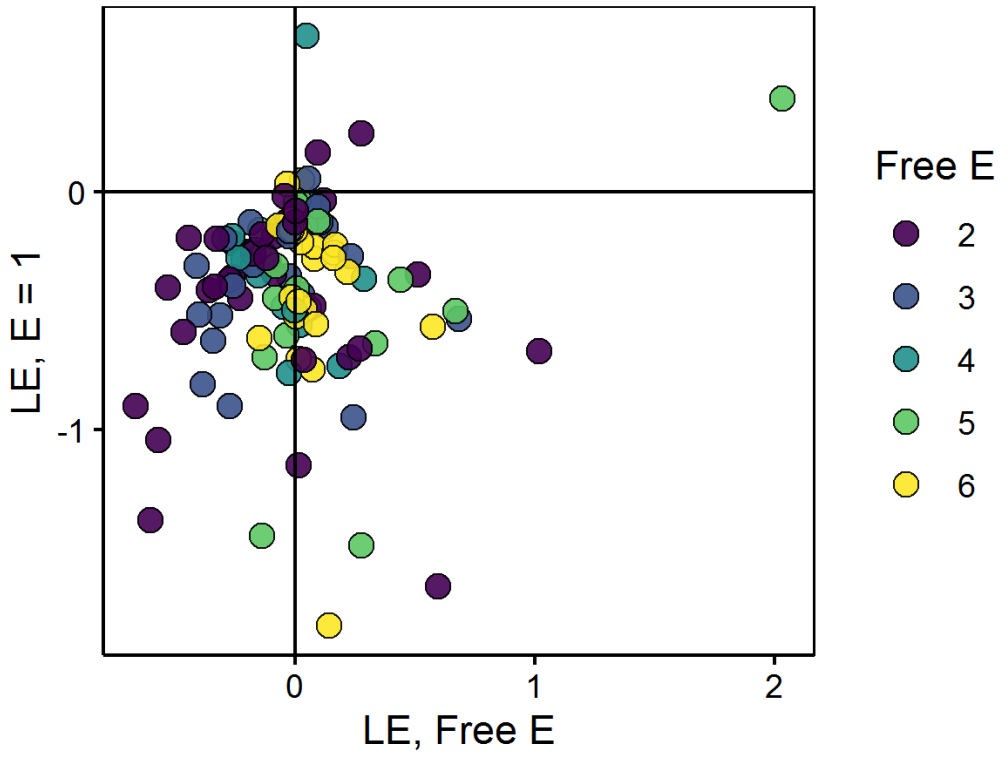
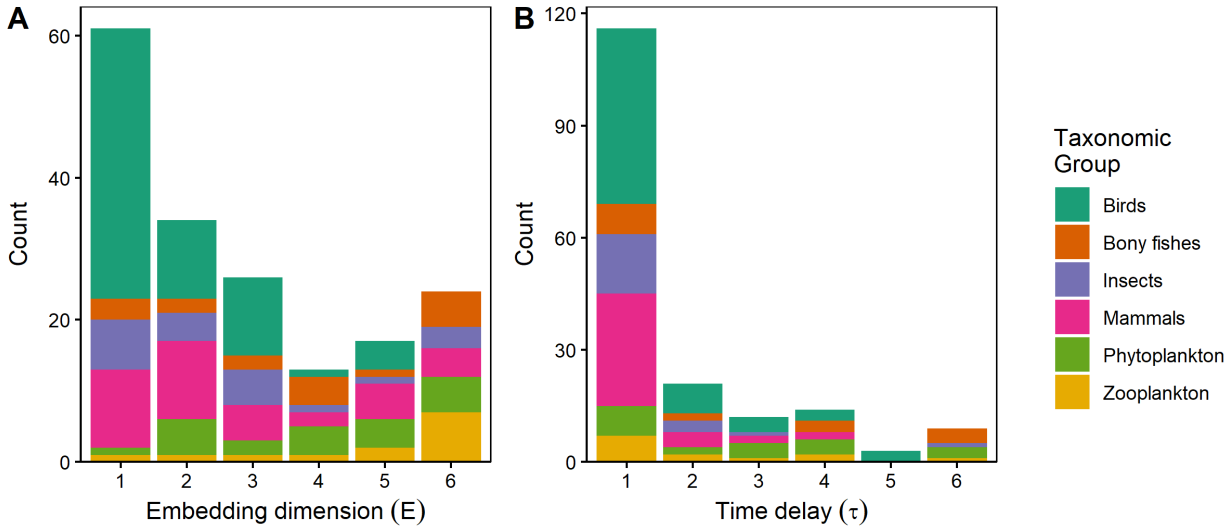


Fig. S9. Lyapunov exponent (LE, timestep^{-1}) with unconstrained embedding dimension (Free E) and with embedding dimension fixed to 1 (E=1), excluding time series with Free E = 1, which would fall along 1:1 diagonal.



1076

1077 **Fig. S10.** Distribution of embedding dimension (E) and time delay ($\boldsymbol{\tau}$) values by taxonomic
1078 class.

1079

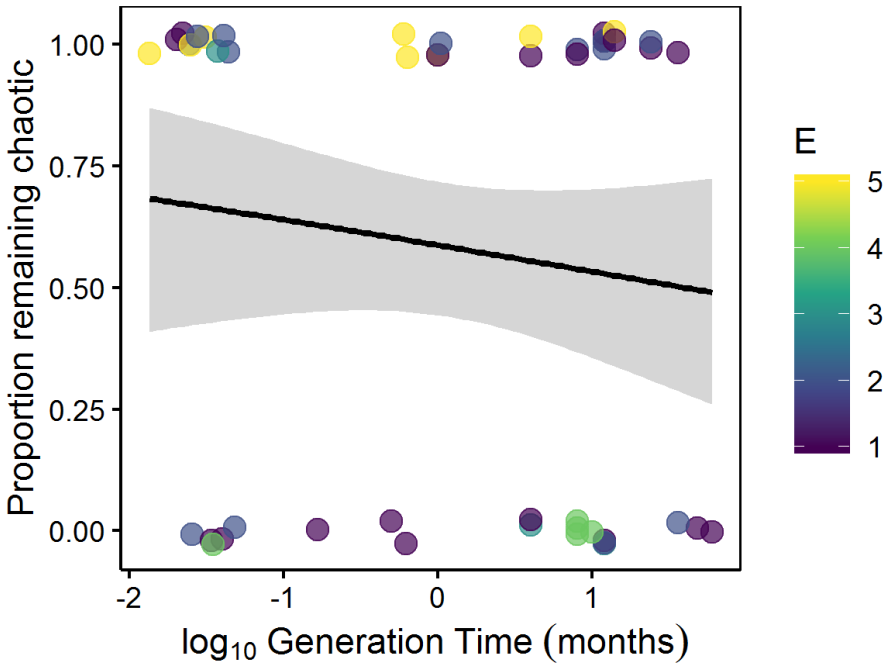


Fig. S11. Proportion of chaotic series that remained classified as chaotic using the Jacobian method when time series were truncated to the last 30 time points, in relation to generation time. Line is logistic regression and 95% confidence interval, and points are vertically offset by a random distance to reduce overplotting.

Table S1. Models used to generate dynamics with a known embedding dimension *E* and time delay $\boldsymbol{\tau}$. Models were generated with random initial conditions and the first 500 points were removed to avoid transients.

Model	**Parameters**	**Known *E***	**Known τ**
$x_t = x_{t-1}e^{r-x_{t-1}}$	$r = 3$	1	1
$x_t = x_{t-2}e^{r-x_{t-2}}$	$r = 3$	1	2
$x_t = x_{t-3}e^{r-x_{t-3}}$	$r = 3$	1	3
$x_t = x_{t-1}e^{r-x_{t-1}-x_{t-2}}$	$r = 3.25$	2	1
$x_t = x_{t-2}e^{r-x_{t-2}-x_{t-4}}$	$r = 3.25$	2	2
$x_t = x_{t-3}e^{r-x_{t-3}-x_{t-6}}$	$r = 3.25$	2	3
$x_t = x_{t-1}e^{r-x_{t-1}-x_{t-2}-x_{t-3}}$	$r = 3.25$	3	1
$x_t = x_{t-2}e^{r-x_{t-2}-x_{t-4}-x_{t-6}}$	$r = 3.25$	3	2
$x_t = x_{t-3}e^{r-x_{t-3}-x_{t-6}-x_{t-9}}$	$r = 3.25$	3	3

Table S2. Model formulations used for s-map regression.

Name	Model
First difference, abundance	$x_t = x_{t-\tau} + f(x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-E\tau})$ $x_t - x_{t-\tau} = f(x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-E\tau})$
Growth rate, abundance	$x_t = x_{t-\tau} e^{f(x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-E\tau})}$ $\ln\left(\frac{x_t}{x_{t-\tau}}\right) = f(x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-E\tau})$
Growth rate, log abundance	$x_t = x_{t-\tau} e^{f(\ln x_{t-\tau}, \ln x_{t-2\tau}, \dots, \ln x_{t-E\tau})}$ $\ln\left(\frac{x_t}{x_{t-\tau}}\right) = f(\ln x_{t-\tau}, \ln x_{t-2\tau}, \dots, \ln x_{t-E\tau})$

Table S3. Models used to generate periodic dynamics (test dataset).

Name	Model	Parameters	Reference
Logistic Map (8-cycle)	$x_{t+1} = rx_t(1 - x_t)$	$r = 3.55$	[22,37]
Logistic Map (3-cycle)	$x_{t+1} = rx_t(1 - x_t)$	$r = 3.828427$	[22,37]
Ricker Map (2-cycle)	$x_{t+1} = x_t e^{r(1-x_t)}$	$r = 2.2$	[125]
Henon Map (4-cycle)	$x_{t+1} = 1 - ax_t^2 + bx_{t-1}$	$a = 0.95$ $b = 0.3$	[22,126]
Sine Wave (12-cycle)	$x_t = a\cos\left(\frac{2\pi}{b}t\right) + a$	$a = 1$ $b = 12$	
Predator-Prey (5-cycle)	$x_{t+1} = x_t + \tau x_t\left(a - x_t - \frac{by_t}{(1+\alpha x_t)(1+\beta y_t)}\right)$ $y_{t+1} = y_t + \tau y_t\left(-c + \frac{dx_t}{(1+\alpha x_t)(1+\beta y_t)}\right)$	$a = 2, b = 2$ $c = 2, d = 1.85$ $\alpha = 0.1, \beta = 0.1$ $\tau = 1.1$	[127]
Host-Parasitoid-Parasitoid (6-cycle)	$x_{t+1} = x_t e^{r\left(1-\frac{x_t}{K}\right) - ay_t^{-m+1} - bz_t^{-n+1}}$ $y_{t+1} = x_t\left(1 - e^{-ay_t^{-m+1} - bz_t^{-n+1}}\right)\frac{ay_t^{-m+1}}{ay_t^{-m+1} + bz_t^{-n+1}}$ $z_{t+1} = x_t\left(1 - e^{-ay_t^{-m+1} - bz_t^{-n+1}}\right)\frac{bz_t^{-n+1}}{ay_t^{-m+1} + bz_t^{-n+1}}$	$r = 2.5, K = 20$ $a = 0.9, b = 1.12$ $m = 0.7, n = 0.4$	[128]

Table S4. Models used to generate chaotic dynamics (test dataset).

Name	Model	Parameters	Reference
Logistic Map	$x_{t+1} = rx_t(1 - x_t)$	$r = 3.9$	[22,37]
Ricker Map	$x_{t+1} = x_t e^{r(1-x_t)}$	$r = 3.4$	[125]
Cubic Map	$x_{t+1} = f\cos(2\pi\theta_t) - Ax_t + x^3$ $\theta_{t+1} = \theta_t + \omega(\mathrm{mod}\ 1)$	$f = -0.8$ $A = 1.5$ $\omega = \frac{\sqrt{5}-1}{2}$	[22,129]
Ikeda Map	$x_{t+1} = 1 + u(x_t\cos\theta_t - y_t\sin\theta_t)$ $y_{t+1} = u(x_t\cos\theta_t - y_t\sin\theta_t)$ $\theta_t = 0.4 - \frac{6}{l + x_t^2 + y_t^2}$	$u = 0.9$	[22,130,131]
Poincare Oscillator	$x_{t+1} = \frac{1}{2\pi}\cos^{-1}\frac{\cos(2\pi x_t)+b}{\sqrt{l+b^2+2b\cos(2\pi x_t)}}(\mathrm{mod}\ 1)$	$b = 1.13$ $\tau = 0.65$	[132]
Predator-Prey	$x_{t+1} = x_t + \tau x_t\left(a - x_t - \frac{by_t}{(1+\alpha x_t)(1+\beta y_t)}\right)$ $y_{t+1} = y_t + \tau y_t\left(-c + \frac{dx_t}{(1+\alpha x_t)(1+\beta y_t)}\right)$	$a = 2, b = 2$ $c = 2, d = 1.85$ $\alpha = 0.1, \beta = 0.1$ $\tau = 1.27$	[127]
Host-Parasitoid-Parasitoid	$x_{t+1} = x_t e^{r\left(1-\frac{x_t}{K}\right) - ay_t^{-m+1} - bz_t^{-n+1}}$ $y_{t+1} = x_t\left(1 - e^{-ay_t^{-m+1} - bz_t^{-n+1}}\right)\frac{ay_t^{-m+1}}{ay_t^{-m+1} + bz_t^{-n+1}}$ $z_{t+1} = x_t\left(1 - e^{-ay_t^{-m+1} - bz_t^{-n+1}}\right)\frac{bz_t^{-n+1}}{ay_t^{-m+1} + bz_t^{-n+1}}$	$r = 3.4, K = 20$ $a = 0.9, b = 1.12$ $m = 0.7, n = 0.4$	[128]

Table S5. Models used to generate stochastic dynamics (test dataset).

Name	Model	Parameters	Reference
AR(1)	$x_{t+1} = c + \phi x_t + \epsilon_t$	$c = 8$ $\phi = 0.8$ $\epsilon \sim N(0,1)$	
Cyclostationary	$x_{t+1} = a_1 x_t + a_2 x_{t-1} + \epsilon_t$	$a_1 = 2\cos\left(\frac{2\pi}{10}\right) e^{-1/50}$ $a_2 = -e^{-1/25}$ $\epsilon \sim N(0,1)$	22,133
Random Walk	$x_{t+1} = x_t + \epsilon_t$	$\epsilon \sim N(0,1)$	22
Random Walk with Trend	$x_{t+1} = x_t + b + \epsilon_t$	$b = 0.1$ $\epsilon \sim N(0,1)$	22
White Noise	$x_t = \epsilon_t$	$\epsilon \sim N(0,1)$	
Red Noise	PSD proportional to $\frac{1}{f^2}$		134
Blue Noise	PSD proportional to f		134

Table S6. Models used to generate periodic dynamics (validation dataset #1).

Name	Model	Parameters	Reference
Three-species competition (4-cycle)	$x_{t+1} = x_t e^{r(1-x_t-ay_t-bz_t)}$ $y_{t+1} = y_t e^{r(1-y_t-az_t-bx_t)}$ $z_{t+1} = z_t e^{r(1-z_t-ax_t-by_t)}$	$r = 2.6$ $a = 0.65$ $b = 0.6$	[135]
Mouse map (2-cycle)	$x_{t+1} = e^{-\alpha {x_t}^2} + \beta$	$\alpha = 6.2$ $\beta = 0$	[136]
Tinkerbell Map (9-cycle)	${x_{t+1}} = {x_t}^2 - {y_t}^2 + ax_t + by_t$ $y_t = 2x_t y_t + cx_t + dy_t$	$a = 0.9, b = -0.5$ $c = 1.8, d = 0.5$	[137]

Table S7. Models used to generate chaotic dynamics (validation dataset #1).

Name	Model	Parameters	Reference
Three-species competition	$x_{t+1} = x_t e^{r(1-x_t-ay_t-bz_t)}$ $y_{t+1} = y_t e^{r(1-y_t-az_t-bx_t)}$ $z_{t+1} = z_t e^{r(1-z_t-ax_t-by_t)}$	$r = 3$ $a = 0.65$ $b = 0.6$	[135]
Mouse map	$x_{t+1} = e^{-\alpha {x_t}^2} + \beta$	$\alpha = 6.2$ $\beta = -0.5$	[136]
Tinkerbell Map	$x_{t+1} = {x_t}^2 - {y_t}^2 + ax_t + by_t$ $y_t = 2x_t y_t + cx_t + dy_t$	$a = 0.9, b = -0.5$ $c = 2.15, d = 0.5$	[137]

Table S8. Models used to generate stochastic dynamics (validation dataset #1).

Name	Model	Parameters	Reference
AR(2)	$x_{t+1} = \phi_1 x_t + \phi_2 x_{t-1} + \epsilon_t$	$\phi_1 = 0.9,$ $\phi_2 = -0.1,$ $\epsilon \sim N(0,1)$	
Seasonally-forced AR(1)	$x_{t+1} = b\left(x_t - A\sin\left(\frac{2\pi}{12}t\right)\right) + \epsilon_t$	$A = 2, b = 0.8$ $\epsilon \sim N(0,1)$	
Pink Noise	PSD proportional to $\frac{1}{f}$		[134]
Violet Noise	PSD proportional to f^2		[134]

Table S9. Models used to generate periodic dynamics with process noise (validation dataset #2).

Name	Model	Parameters	Reference
Seasonal Predator Prey	$\dot{x} = r(1 - A\ \sin 2\pi t))x - rx^2 - \frac{gx^2}{x^2 + h^2} - \frac{axy}{x + d}$ $\dot{y} = s(1 - A\ \sin 2\pi t))y - \frac{sy^2}{x}$	$A = 2.2$ $s = 1.5$ $g\ = 2$ $h\ = 0.13$ $a\ = 7.5$ $d\ = 0.06$ $r \sim N(7, \sigma)$	[138]
Larvae-Pupae-Adult (LPA)	$L_{t+1} = bA_t e^{-c_{el}L_t - c_{ea}A_t} e^{\xi_{1,t}}$ $P_{t+1} = (1 - \mu_l)L_t e^{\xi_{2,t}}$ $\mathrm{A_{t+1}} = [P_t e^{-c_{pa}A_t} + A_t(1 - \mu_a)]e^{\xi_{3,t}}$	$b = 6.598$ $c_{el} = 0.01209$ $c_{ea} = 0.01155$ $c_{pa} = 1$ $\mu_l = 0.2055$ $\mu_a = 0.96$ $\xi_{i,t} \sim N(-\frac{\sigma^2}{2}, \sigma)$	[139]
Nutrient-Phytoplankton-Zooplankton (NPZ)	$\dot{N} = -\frac{v_m\left(1 - A\sin\left(\frac{2\pi}{365}t\right)\right)NP}{k_s + N} e^{kh} + \gamma R_m Z(1 - e^{-\Lambda \mathrm{P}}) + mP + gZ$ $\dot{P} = \frac{v_m\left(1 - A\sin\left(\frac{2\pi}{365}t\right)\right)NP}{k_s + N} e^{kh} - R_m Z(1 - e^{-\Lambda \mathrm{P}}) - mP$ $\dot{Z} = (1 - \gamma)R_m Z(1 - e^{-\Lambda \mathrm{P}}) - gZ$	$v_m = 2$ $k_s = 0.1$ $k\ = 0.06$ $\Lambda\ = 0.2$ $\gamma\ = 0.3$ $m\ = 0.1$ $g\ = 0.2$ $A\ = 0$ $h\ = -35$ $R_m \sim N(0.5, \sigma)$	[140]

Table S10. Models used to generate chaotic dynamics with process noise (validation dataset #2).

Name	Model	Parameters	Reference
Seasonal Predator Prey	$\dot{x} = r(1 - A\ \sin 2\pi t))x - rx^2 - \frac{gx^2}{x^2 + h^2} - \frac{axy}{x + d}$ $\dot{y} = s(1 - A\ \sin 2\pi t))y - \frac{sy^2}{x}$	$A = 1$ $s = 1.25$ $g\ = 0$ $h\ = 0.08$ $a\ = 710$ $d\ = 0.04$ $r \sim N(6, \sigma)$	[138]
Larvae-Pupae-Adult (LPA)	$L_{t+1} = bA_t e^{-c_{el}L_t - c_{ea}A_t} e^{\xi_{1,t}}$ $P_{t+1} = (1 - \mu_l)L_t e^{\xi_{2,t}}$ $\mathrm{A_{t+1}} = [P_t e^{-c_{pa}A_t} + A_t(1 - \mu_a)]e^{\xi_{3,t}}$	$b = 6.598$ $c_{el} = 0.01209$ $c_{ea} = 0.01155$ $c_{pa} = 0.35$ $\mu_l = 0.2055$ $\mu_a = 0.96$ $\xi_{i,t} \sim N(-\frac{\sigma^2}{2}, \sigma)$	[139]
Nutrient-Phytoplankton-Zooplankton (NPZ)	$\dot{N} = -\frac{v_m\left(1 - A\sin\left(\frac{2\pi}{365}t\right)\right)NP}{k_s + N} e^{kh} + \gamma R_m Z(1 - e^{-\Lambda \mathrm{P}}) + mP + gZ$ $\dot{P} = \frac{v_m\left(1 - A\sin\left(\frac{2\pi}{365}t\right)\right)NP}{k_s + N} e^{kh} - R_m Z(1 - e^{-\Lambda \mathrm{P}}) - mP$ $\dot{Z} = (1 - \gamma)R_m Z(1 - e^{-\Lambda \mathrm{P}}) - gZ$	$v_m = 2$ $k_s = 0.1$ $k\ = 0.06$ $\Lambda\ = 0.3$ $\gamma\ = 0.7$ $m\ = 0.1$ $g\ = 0.2$ $A\ = 1$ $h\ = 0$ $R_m \sim N(4, \sigma)$	[140]

Table S11. False negative rates (FNR) and false positive rates (FPR) for 6 chaos detection methods across simulated datasets. Values in italics indicate misclassification rates >0.5.

Chaos detection method	Test		Validation #1		Validation #2	
	FNR	**FPR**	**FNR**	**FPR**	**FNR**	**FPR**
Direct LE	0.04	*0.68*	0.04	*0.61*	0.23	*0.68*
Jacobian LE	0.35	0.03	0.16	0.03	0.28	0.10
Recurrence quantification analysis	0.38	0.16	0.32	0.03	0.38	0.28
Permutation entropy	0.30	0.20	0.18	0.16	0.25	0.15
Horizontal visibility algorithm	*0.57*	0.09	*0.80*	0.13	*0.54*	0.05
Chaos decision tree	*0.77*	0.02	*0.72*	0.001	*0.65*	0.09

Table S12. 1-d models fit to the empirical GPDD dataset, including the average R^2 and Lyapunov exponent (LE) across all time series and the proportion of time series classified as chaotic. The HLM II model extends the model of [16] to allow for adult survival analogous to what [17] did with the Ricker model[125].

Source	Model	Average R^2	Average LE	Proportion chaotic
[10]	$\ln(n_{t+1}) = a + b\ln(n_t) + c[\ln(n_t)]^2$	0.22	-0.64	0.06
[125]	$n_{t+1} = n_t \exp[a - bn_t]$	0.17	-0.44	0.03
[16]	$n_{t+1} = n_t[a + bn_t]^{-c}$	0.24	-0.70	0.01
[17]	$n_{t+1} = n_t \exp[a - bn_t] + sn_t$	0.24	-0.41	0.06
HLM II	$n_{t+1} = n_t[a + bn_t]^{-c} + sn_t$	0.25	-0.47	0.06