

Nonlinear population dynamics are ubiquitous in animals

T. J. Clark¹ and Angela D. Luis

Nonlinear dynamics, where a change in the input is not proportional to a change in the output, are often found throughout nature, for example in biochemical kinetics. Because of the complex suite of interacting abiotic and biotic variables present in ecosystems, animal population dynamics are often thought to be driven in a nonlinear, state-dependent fashion. However, so far these have only been identified in model organisms and some natural systems. Here we show that nonlinear population dynamics are ubiquitous in nature. We use nonlinear forecasting to analyse 747 datasets of 228 species to find that insect population trends were highly nonlinear (74%), followed by mammals (58%), bony fish (49%) and birds (35%). This indicates that linear, equilibrium-based model assumptions may fail at predicting population dynamics across a wide range of animal taxa. We show that faster-reproducing animals are more likely to have nonlinear and high-dimensional dynamics, supporting past ecological theory. Lastly, only a third of time series were predictable beyond two years; therefore, the ability to predict animal population trends using these methods may be limited. Our results suggest that the complex dynamics necessary to cause regime shifts and other transitions may be inherent in a wide variety of animals.

Dynamic phenomena in nature are becoming increasingly recognized as nonlinear¹. That is, dynamic systems are state dependent and cannot be modelled by the sum of their non-interacting, additive parts. As an example, animal population dynamics are suspected to be nonlinear because of the complexity and interrelatedness of variables that influence populations and the regularity in which some animal populations fluctuate^{2–4}. The search for dynamic complexity (defined herein as nonlinearity and high dimensionality, that is, a high number of interacting processes that produce observed dynamics)⁵ in animals began with accounts of chaotic dynamics in model organisms and has since expanded to include some natural systems including marine ecosystems^{4,6–9}. However, this search has been hampered by the amount of data required to adequately analyse time series⁴, and how prevalent dynamic complexity is in animal population dynamics is unknown. This has led some to conclude that complex dynamics may be rare in animals^{9,10}.

Despite evidence for dynamic complexity, fluctuations in animal abundances are thought to be largely explained by low-dimensional, nonlinear dynamics¹¹. Complex dynamics were first explored with simple population models, such as the discrete logistic equation ($N_{t+1} = N_t e^{r(1 - \frac{N_t}{K})}$), which shows increasing nonlinearity as the intrinsic growth rate r increases^{5,7,12}. Prevailing theory has since suggested that the demography necessary for nonlinear dynamics may be present in smaller, faster-reproducing animals with non-overlapping generations^{7,8,12}. Nonlinear dynamics may occur for other reasons, including deterministic, multiplicative interactions between species or other abiotic/biotic variables^{4,13} and stochastic environmental noise^{13,14}. For example, Dixon et al.¹³ found that nonlinear episodic fluctuations in a larval damselfish population resulted from both deterministic (lunar phases) and stochastic processes (for example, wind direction). Similarly, Hsieh et al.⁴ found evidence that stochastic physical forcing combines with low-dimensional, nonlinear and deterministic biological variables to govern marine ecosystems in the North Pacific Ocean. Therefore, we hypothesize that nonlinear population dynamics are driven by: (1a) high intrinsic

growth rate; (1b) high dimensionality; (1c) stochastic noise; and (1d) that these dynamics are shared by animals of similar taxonomic classification.

In contrast to the quest for nonlinearity in ecology, there has been little research exploring the factors that cause higher dimensionality in animal time series. Some have suggested that marine ecosystems, especially fisheries, are potentially high-dimensional due to the added dimensions of human intervention via fishing effort (for example, fisheries management and market prices)^{6,15–17}. In contrast, some trophic cascade researchers have suggested that the paucity of documented cascades in terrestrial ecosystems are due to their reticulate, high-dimensional nature^{18,19}. Despite this, high-dimensional dynamics are thought to be rare in animal populations¹¹. This is supported in part by the pervasiveness of only a few strong interactions found embedded in many weak interactions in the analyses of natural food webs^{20,21}. Still we hypothesize that high-dimensional dynamics are more prevalent among: (2a) faster-reproducing animals, like insects and fish, which may be prone to numerous, complex relationships within their dynamic systems due to demography^{6,22}; (2b) lower-level consumers (for example, herbivores), which have been found to have many strong interactions compared with higher trophic levels^{23,24}; and (2c) that these dynamics are common across taxonomically similar species.

Despite prior research, it is unclear if the aforementioned factors drive complex population dynamics writ large. Furthermore, these complex dynamics are necessary for abrupt and unexpected ecological shifts, like transitions between equilibria, steep declines in abundance and local extinctions^{25–27}. If certain taxa are capable of producing complex dynamics, this may make the long-term prediction of their populations difficult, potentially limiting precautionary management strategies in the face of increasing perturbations^{4,6}. Broadly, more complex dynamics are thought to be more difficult to predict due to a variety of factors that make time series more fluctuating and random in nature^{28,29}. Therefore, we expect that the following factors will make time series more difficult to predict: (3a)

Wildlife Biology Program, Department of Ecosystem and Conservation Sciences, W.A. Franke College of Forestry and Conservation, University of Montana, Missoula, MT, USA. *e-mail: tyler3.clark@umontana.edu

faster-reproducing species, which due to demography often have rapidly fluctuating populations; (3b) nonlinearity, which is known to make animal populations more difficult to predict^{28,30}; and (3c) that taxonomically similar species will have related predictability (hypotheses for factors affecting predictability).

In the present study, we carried out a large-scale, cross-taxa test for dynamic complexity and predictability in animal population dynamics. We collected animal time series from a variety of sources, including global databases and grey literature (see Methods). We filtered these datasets according to a strict regimen (Supplementary Information, section 2), including only those with ≥ 30 time steps due to model constraints³⁰. This narrowed our final dataset to 747 time series consisting of 228 animal species, including 75 birds, 49 insects, 54 bony fish and 50 mammals.

We then used nonlinear forecasting^{14,31} to determine the dimensionality (E , a relative index for the number of strongly interacting variables), nonlinearity (θ , state-dependence) and predictability (ρ , forecast skill) of these time series. These innovative methods, known collectively as empirical dynamic modelling (EDM), use state-space population reconstruction through lagged coordinate embedding of time series (see Methods)^{14,31,32}. EDM was developed specifically for short and noisy time series³⁰ and can reliably identify and predict complex dynamics if present^{6,14,31}. These methods have been used previously to explore complex dynamics and predictability across taxa in large datasets for marine fisheries^{6,17} and in planktonic ecosystems³³. Yet, no study has conducted these tests across a wide range of animal taxa.

We used a two-step process that determined the best parameters using out-of-sample forecast skill (ρ): (1) simplex projection, a nearest neighbours algorithm³¹, to calculate E , the embedding dimension, which is the number of consecutive time lags needed to reconstruct the state-space and is an index of dimensionality³¹; and (2) sequentially weighted global linear maps (S-maps) with the best E from step 1 to calculate the local weighting parameter θ , where larger values of θ indicate higher nonlinearity¹⁴. Time series were then classified as nonlinear if there was a significant reduction in error in prediction from a linear to a nonlinear model³⁰ using a randomization procedure (see Methods)³⁴. Alternatively, time series that were not significant were classified as linear-stochastic, either because the dynamics were best modelled by a stable deterministic system or because they cannot be distinguished from a solely stochastic process (see Methods)⁶. We then chose the respective parameter (E and θ) that produced the highest forecast skill defined as the correlation between the actual and predicted time series (see Methods). Models of time series for E and θ that were not significantly correlated with the actual time series were categorized as 'not predictable' and excluded from the statistical analysis of dimensionality and nonlinearity (see Methods)⁶.

We then used E , θ and ρ as response variables in post-hoc analyses to test our hypotheses for factors affecting nonlinearity (hypotheses 1a–d), dimensionality (hypotheses 2a–c) and predictability (hypotheses 3a–c), respectively. We constructed a phylogenetic tree to explore if dynamic complexity was related to evolutionary relationships among species (hypotheses 1d, 2c and 3c). We collated life-history traits (for example, lifespan) for each species to capture differences in demographics^{35,36} from a slow-to-fast continuum, to test the hypotheses that life history and growth rate (see Methods) affect dynamic complexity (hypotheses 1b, 2a and 3a). We collected information at the trophic level for each species to test the hypothesis that trophic interactions affect dimensionality (hypothesis 2b). Lastly, we accounted for confounding differences between time series due to factors such as time series length, generation time and sampling methodology using nuisance variables (see Methods).

Results

Nonlinearity. We show that nonlinear dynamics are common across a wide variety of animal taxa and are related to life-history

strategies, supporting past evidence for some of these differences in ecological theory⁸ and literature for marine ecosystems^{4,6} (Fig. 1a). Our best model using the Akaike's information criterion (AIC) indicated that nonlinear dynamics broadly differed between taxonomic class, with 74% of insect species, 58% of mammals, 49% of bony fish and 35% of birds showing nonlinear dynamics (Fig. 1a). Time series with faster life-history traits (that is, principal components axis of shorter body length, minimum age of first reproduction and longevity; see Methods and Extended Data Fig. 1) were 0.9 times more likely to display nonlinear dynamics (supporting hypothesis 1a). This result lends support to the prediction that nonlinear dynamics occur in response to elevated growth rate, which has been theoretically and experimentally shown to cause a higher probability of transitions between dynamical regimes^{5,7,11,12,37}. Additionally, time series with less stochastic noise ($1 - \rho$ (refs. 6,28); more predictable) were 1.7 times more likely to be nonlinear, contrary to our hypothesis (1c; Extended Data Fig. 2). Lastly, time series with higher dimensionality were 0.9 times more likely to be nonlinear, supporting the hypothesis that an increasing number biotic/abiotic interactions increase nonlinearity (1b; Extended Data Fig. 2) (refs. 4,13).

The nonlinear signal appeared more nuanced when time series were classified by taxonomic order; however, we could not test for the significance of taxonomic order due to the small sample size (Fig. 1b and Extended Data Fig. 3). For example, rodents (61%) are moderately nonlinear, preliminarily suggesting that this may be the dynamic cause of known fluctuations in these animals³⁸ in contrast to relatively linear carnivores (40%). Hemipterans, primarily comprised of aphid time series (29 species; 88% of insect time series), were highly nonlinear (74%), potentially due to their parthenogenetic life history, which is characteristic of many aphid populations²². Similar to past research³³, there was slight support that nonlinearity was more probably detected in longer time series (Extended Data Fig. 2). We suggest that nonlinear dynamics in animals are primarily driven by their life history and their complex, interconnected relationship with other species, surrounding habitat and environment.

Dimensionality. Median dimensionality was similar across taxonomic classes and unrelated to phylogeny (Fig. 2). This supports a common theory that many population fluctuations can be explained by low-dimensional, nonlinear dynamics^{11,30}. However, we show that high-dimensional, nonlinear dynamics still pervade some animal population dynamics (Fig. 2). Still, we caution against overinterpretation of the values of dimensionality, since these are only heuristic values that describe the complexity of the system or the number of interacting processes that created these dynamics^{6,31,39}. Animals that were smaller and faster-reproducing had higher dimensionality (supporting hypothesis 2a; Extended Data Fig. 4), suggesting that these species may strongly interact with more abiotic/biotic variables. We were also more likely to find higher dimensionality in longer time series (Extended Data Fig. 4). Contrary to our hypotheses and previous theory²³, we found that dimensionality was not related to taxonomic class or trophic level (hypotheses 2b and 2c; Extended Data Fig. 4), indicating that animals across taxa, not only marine organisms, may be subject to high numbers of interactions.

Predictability. We found that 86% of time series were predictable 1 time step into the future. Predictability was related to taxonomic class (supporting hypothesis 3c), with insects (96%) being very predictable, followed by bony fish (78%), mammals (57%) and birds (55%). In contrast to our hypothesis (3b), we found a higher forecast skill for time series with nonlinear dynamics than those with linear-stochastic dynamics (or non-predictable dynamics) (Fig. 3 and Extended Data Fig. 5). Animals with faster life-history traits had dynamics that were more predictable (Extended Data Fig. 5), contrary to our hypothesis (3a) and past results with similar modelling

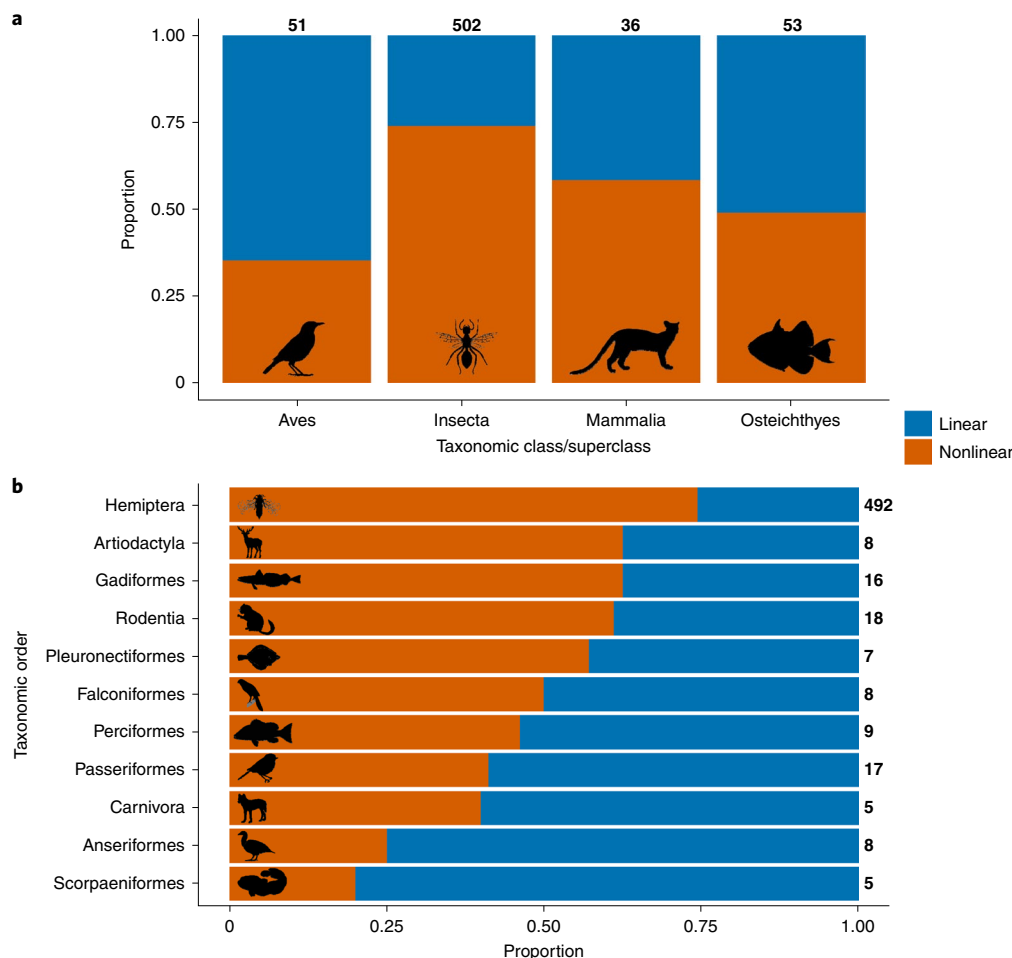


Fig. 1 | Animal time series with linear or nonlinear population dynamics. a, Animal time series arranged by taxonomic class/superclass. **b,** Animal time series arranged by taxonomic order, where sample size ≥ 5 . Only predictable time series were used. The numbers in bold show the sample size.

techniques⁴⁰. As the number of time steps into the future increased, predictability declined exponentially, with one-third of time series predictable three time steps into the future (Table 1). This indicates that medium- or long-term (>2 years) prediction of nonlinear population dynamics may be limited and that only short-term predictions are attainable with these methods^{6,14}.

As the number of time steps into the future increased, time series with nonlinear dynamics declined exponentially in forecast skill^{6,14}. Time series with linear-stochastic dynamics had a similar decline in forecast skill, yet overall they were not as predictable as nonlinear dynamics. Similarly, this has been found in a large-scale analysis of fisheries time series using nonlinear forecasting⁶; it may indicate poorly resolved attractors and, therefore, strong linear-stochastic noise (average $1 - \rho = 0.47$; Fig. 3). This suggests that no matter the species studied, the underlying systems may be hard to predict with these methods due to high-dimensional nonlinearity or linear stochasticity. In fact, these results mirror recent doubts over the feasibility of accurately predicting some complex ecological systems^{41,42}. Despite these difficulties, we suggest that calculating the dimensionality and nonlinearity of time series may facilitate a better understanding of animal population dynamics and help build better models for prediction^{4,14,30,31}. For example, if a system is governed deterministically by nonlinear, low-dimensional dynamics, a simple mechanistic model may capture the behaviour well; monitoring these few input variables in an ecosystem may then be realistic³⁰. However, if a system is governed deterministically by high-dimensional, nonlinear dynamics or stochastically by high-dimensional

linear dynamics, the aforementioned models may fail and require a statistical autoregressive model or state-space reconstruction^{30,31}. As such, we and others^{6,30} suggest using nonlinear forecasting to understand the complexity and predictability of a system, thereby building better-informed models.

Broader implications. We found that nonlinear dynamics were ubiquitous across a wide variety of animal taxa and related to faster-reproducing life-history traits. Previously, complex dynamic behaviours (for example, chaos) were met with scepticism and some have suggested that they are probably rare in animals^{9,10} or that only insects have the demography to allow these dynamics to emerge^{8,22}. Additionally, this bias is reflected in the fact that linear, equilibrium-based model assumptions are the standard for most models of animal population management and conservation. However, our results suggest that nonlinear dynamics may be more common than previously thought; as others have suggested^{16,28,29}, a methodological shift towards nonlinear, state-space models may be necessary for a wide variety of animals.

Prior research using similar datasets (for example, the Global Population Dynamics Database (GPDD)) have ignored observational uncertainty in animal population estimates, which has been shown to bias tests and estimates of ecological processes (for example, density dependence) leading to poor inference⁴³. To account for uncertainty, we have eliminated potentially noisy datasets using informed filtering criteria (see Supplementary Information). Furthermore, EDM has been shown to be robust against observational noise⁴⁴; regularization

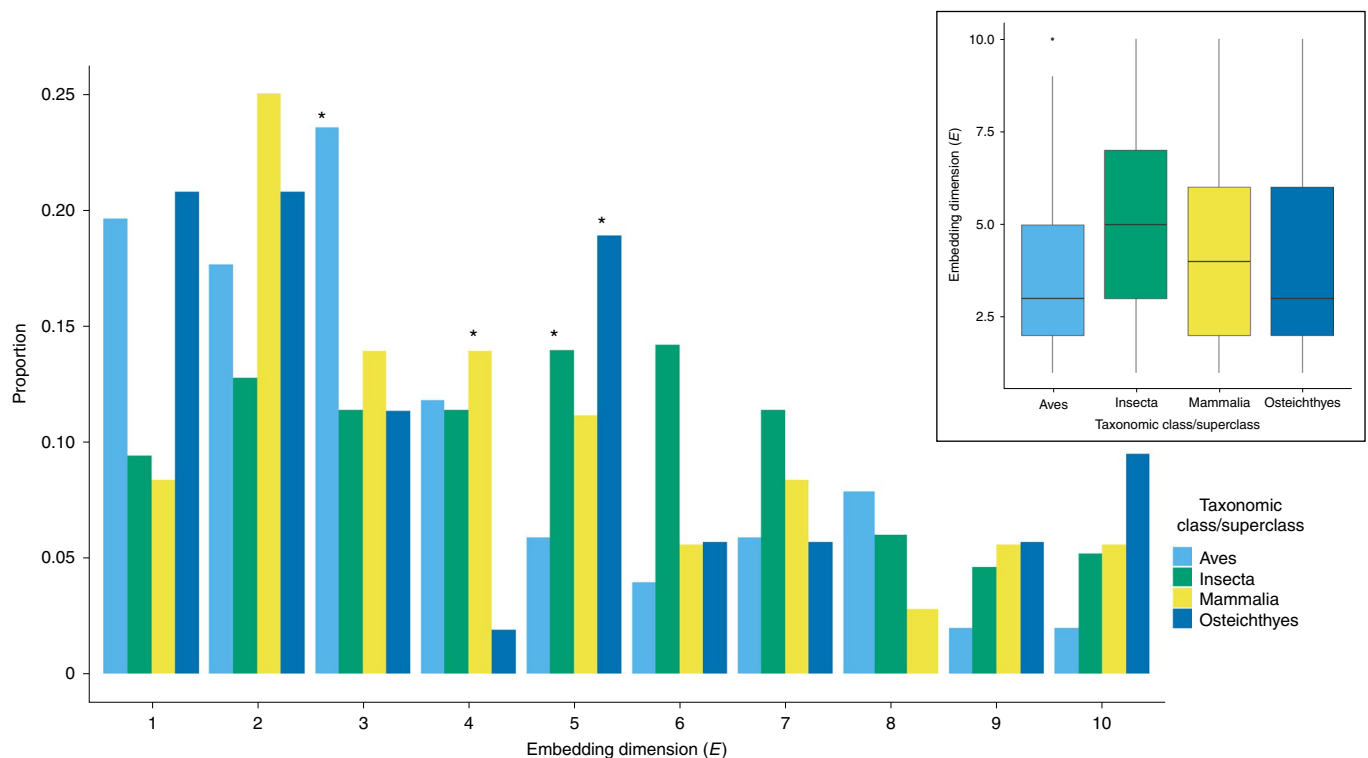


Fig. 2 | Dimensionality in animal time series by taxonomic class/superclass. Higher values of embedding dimension (E) indicate that the underlying dynamics are high-dimensional. The asterisks above the bars in the plot show the median dimensionality for each taxonomic class. The inset shows boxplots of embedding distribution by taxonomic class (centre line, median; box limits, upper and lower quartiles; whiskers, 1.5 \times interquartile range; dots, outliers).

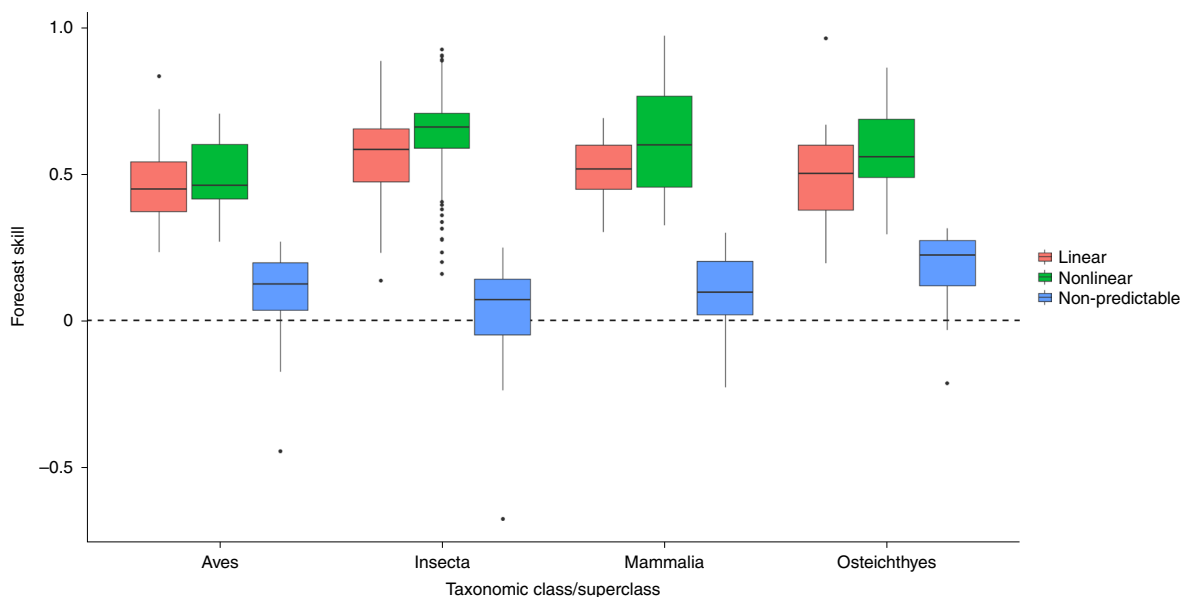


Fig. 3 | Forecast skill of animal time series by taxonomic class/superclass. The forecast skill p is the Pearson product-moment correlation coefficient of the leave-one-out cross-validation of a time series using the S-map procedure. Time series were categorized as predictable if the Pearson product-moment correlation coefficient was statistically significant ($P \leq 0.05$). Predictable time series were then categorized as linear-stochastic or nonlinear (centre line, median; box limits, upper and lower quartiles; whiskers, 1.5 \times interquartile range; dots, outliers).

schemes are being developed to extend these methods to overcome sensitivity to high process noise⁴⁵. It is difficult to separate the relative contribution of observational versus process error in time series datasets and estimates of observational error are often imprecise⁴³.

Therefore, future work should be done to determine which error terms (observational versus process) dominate in animal time series (for example, see Ahrestani et al.⁴⁶) and how to account for these in modelling frameworks to lead to strong ecological inference.

Table 1 | Proportion of animal time series with predictable forecasts as time steps in the future increase

Class/ superclass	Time step 1	Time step 2	Time step 3	Time step 4	Time step 5
Aves	0.548	0.140	0.086	0.054	0.075
Insecta	0.960	0.602	0.373	0.373	0.375
Mammalia	0.571	0.286	0.111	0.143	0.143
Osteichthyes	0.779	0.191	0.103	0.103	0.118
Total	0.859	0.481	0.290	0.290	0.331

Time series were categorized as 'predictable' if the Pearson product-moment correlation coefficient of the leave-one-out cross-validation was statistically significant ($P \leq 0.05$).

Sustained oscillations, aperiodic behaviour, alternative stable states and critical transitions have all been implicated as symptoms of nonlinear dynamics^{25–27}. For example, regime shifts, where dramatic changes occur in the structure and function of an ecosystem, are known to cause ecological catastrophes²⁷ like the dominance of macroalgae on coral reefs following climate-induced bleaching⁴⁷. Our findings indicate that the nonlinear dynamics that are necessary to cause catastrophic shifts may be present in a wide variety of animal taxa. These dynamics do not mean that regime shifts are inevitable; however, simulations and empirical data have shown that that elevated nonlinearity can indicate if regime shifts may occur^{4,6,25,27}. Additionally, nonlinear dynamics have been linked to slower, less dramatic transitions, such as local extinction in deteriorating environments⁴⁸. Yet, in the face of extreme perturbations, such as climate change, anthropogenic stress or inconsistent management (for example, resource management cycles⁴⁹), critical transitions are more likely to occur, often well before the bifurcation point^{26,27}. In consideration of our results and others^{4,6,49}, we suggest that a careful, precautionary and adaptive management strategy is essential for all animals.

Methods

Data collection. Population dynamics time series data were compiled from the GPDD, other databases and grey literature (Supplementary Information, section 1). Similar to past applications of GPDD data^{46,50}, we conducted a strict filtering criteria due to inconsistencies in time series, such as long strings of repeating, non-unique values that have been shown to limit the classification of nonlinearity and prediction of time series³³ (Supplementary Information, section 1 and Extended Data Fig. 6). This narrowed our final dataset to 747 time series consisting of 228 animal species, including 75 bird species, 49 insects, 54 bony fish and 50 mammals. Datasets were limited to ≥ 30 population counts per time series due to modelling constraints⁵⁰, with a median of 47.5 and range of 30–275 (Extended Data Fig. 7). Time series were first-differenced ($\Delta x = x_t - x_{t-1}$) and standardized to reduce short-term autocorrelation, preserve stationarity and allow the comparison of datasets²⁸.

We created a full phylogenetic tree for species in our dataset from a comprehensive supertree, the Open Tree of Life⁵¹, using the rotl⁵² package (v3.0.10). We collected species-specific life-history traits from Brook et al.³⁶ and independent sources (for example, <https://www.demogr.mpg.de/longevityrecords/0203.htm>, fishbase.org; audubon.org) for each of the 232 species. These traits included: (1) maximum body size (length in mm); (2) average age at first reproduction (months); (3) longevity (maximum age attained in the wild in months); and (4) fertility (number of young per year). Due to high multicollinearity between life-history traits, we used principal component analysis to produce predictors from the two principal components with the highest eigenvalues (explaining 97.1% of the variation in the dataset). Additionally, we categorized each species according to trophic level³⁶.

Nonlinear forecasting. We used two separate nonlinear forecasting models developed for short time series using state-space reconstruction through lagged coordinate embedding (Takens' theorem; for a short explanation, see <https://youtu.be/fevurdpIRYg>). Takens' theorem proves that the system dynamics of a single time series can be represented by substituting the time lags of the same time series. All nonlinear forecasting was conducted in R⁵³ (v3.5.0) using the package rEDM v.0.7.2⁵⁴. First, we used simplex projection to calculate the embedding dimension (E), which is the number of consecutive time lags necessary to reconstruct the system state-space using lagged coordinates. In a sense, E provides an index

of the dimensionality or the number of interacting variables of the underlying dynamic system for each time series³¹. Second, we classified the system dynamics of each time series as linear or nonlinear using sequentially weighted global linear maps (S-maps)¹⁴. S-maps are akin to linear autoregressive models where the model coefficients of the current state of the system from which the prediction is being made ('predictee') depend on where it is located in an E -dimensional embedding^{31,14}. When $\theta = 0$, this represents a global linear model for all predictees regardless of location in state-space, and the model reduces to an autoregressive model of order E . When $\theta > 0$, this represents a local nonlinear model where neighbours that are closer to the point to be predicted are given stronger weighting, creating a locally weighted map. Systems are considered to be nonlinear if greater weighting results in a better model than equally weighting all the data¹⁴. For simplex projection, we iterated embedding dimension (E) from 1–10, identified the best E , then applied it with S-maps and varied nonlinear tuning parameter (θ) from 0–8, respectively for each time series. Time series were then classified as nonlinear if the change in mean absolute error (MAE) from a linear to nonlinear model ($\Delta\text{MAE} = \text{MAE}_{\theta=0} - \text{MAE}_{\theta=\text{min}}$) was positive and significant at $P \leq 0.05$ (ref. ³⁰). Time series that failed this test were classified as linear-stochastic either because the dynamics were best modelled by a stable deterministic system or because they could not be distinguished from a solely stochastic process⁶. To determine the P value, we ran a randomization procedure where we calculated ΔMAE for the original time series, then generated 1,000 phase-randomized surrogates³⁴, which preserve the basic statistical properties of the time series like autocorrelation but introduce randomization, thereby creating a null distribution ΔMAE to compare our original ΔMAE against³⁰. We then chose the respective parameter (E and θ), which produced the highest forecast skill, defined as the Pearson's correlation coefficient between the actual and predicted time series via leave-one-out cross-validation^{14,31}. Models of time series for E and θ that did not have a significant Pearson's product-moment correlation coefficient at $P \leq 0.05$ were categorized as non-predictable and excluded from the statistical analysis of dimensionality and nonlinearity⁶. Although the variables E , θ and ρ are useful descriptions of the reconstruction of the original dynamic system, we caution against overinterpretation of these variables or the cause versus effect relationship between them³³.

To quantify the potential predictability and noise of our datasets, we analysed the nonlinear forecasts from S-maps as a function of how many time steps in the future ($t_f = 1$ –5) the forecast skill persisted^{6,14}. We then calculated the maximum ρ at $t_f = 1$ to understand the amount of deterministic signal and noise ($1 - \rho$) present in each time series. In sum, our final dataset included 747 time series of 228 species, with 642 time series of 163 species being assigned as significant at $P \leq 0.05$ and therefore predictable by the Pearson's product-moment correlation coefficient. For examples of these time series, see Extended Data Fig. 8.

Statistical analysis. Due to overrepresentation of well-studied species (see Supplementary Information section 1), we randomly chose a single time series for each species for subsequent analyses. Subsequent random sampling did not change our results. We analysed the importance of phylogenetic relationship in explaining system dimensionality, nonlinearity and model forecast skill by creating a covariance matrix of the evolutionary relationships between all species in our dataset. We analysed dimensionality and forecast skill using phylogenetic generalized least squares with expected covariance under a Brownian model, finding no significant phylogenetic relationship ($\lambda = 0.001$ and -0.055 , respectively)⁵⁵. Because nonlinearity was a binomial variable, we analysed the importance of phylogeny using the binaryPGLMM function in the ape⁵⁶ package (v5.0), finding no significant phylogenetic relationship ($s^2 = 0.2563$, $P = 0.1228$).

After removing phylogenetic covariance, we analysed the relation of taxonomic classification, trophic level and life history to dimensionality, the detection of nonlinear dynamics and model forecast skill using generalized linear models. The variable E , an ordinal categorical variable, was analysed using a cumulative logit model ($n = 157$). We attempted to analyse nonlinearity using a logit link function with a binomial distribution ($n = 157$). However, due to complete separation in the variable 'data type', we instead used Bayesian analysis with a non-informative Cauchy prior⁵⁶ for each coefficient using the package arm v.1.10-1 (ref. ³⁷). The variable ρ was analysed using an identity link function and Gaussian distribution ($n = 224$).

We conducted backwards stepwise model selection based on the AIC, deleting model terms until the lowest possible AIC was achieved. To analyse nonlinearity, we evaluated model selection using the Bayesian information criterion and AIC, which subsequently gave the same results. The maximal models considered were the following: (1) $E \approx \text{TC} + \text{PC1} + \text{PC2} + \text{trophic} + N + \text{CV} + G + \text{data type} + \text{data type} \times \text{CV}$; (2) $\text{nonlinear} \approx \text{TC} + \text{PC1} + \text{PC2} + E + \rho + N + \text{CV} + G + \text{data type} + \text{data type} \times \text{CV}$; (3) $\rho \approx \text{TC} + \text{PC1} + \text{PC2} + \text{predict category} + N + \text{CV} + G + \text{data type} + \text{data type} \times \text{CV}$. Predictor variables that were part of our hypotheses, in addition to E and ρ , were the following: TC was a categorical variable representing taxonomic class. PC1 represented the first principal component (72.2% of explained variation) of the principal component analysis of life-history traits, with positive values representing shorter body length, minimum age of first reproduction and longevity. PC2 represented the second principal component (24.9% of explained variation) of the principal component analysis, with positive values representing less fertility (Extended Data Fig. 1). 'Trophic' was a categorical

variable representing trophic level, including herbivores, detritivores, carnivores and omnivores. 'Predict category' was a categorical variable defining if a time series was classified as non-predictable, predictable linear or predictable nonlinear. Variables that were used to control for inconsistencies between our time series were the following: N was the number of values in a time series and was included as a control because nonlinear dynamics and dimensionality may be obscured when time series are too short^{6,30}. CV represented the coefficient of variation calculated from first-differenced time series to control for higher-level variability in our time series. G was equal to the number of generations monitored in each census (time series length/minimum age at first reproduction) to control for the temporal mismatch between generation length and sampling frequency³⁵. However, due to lack of data, we were unable to account for differences in spatial scales of sampling, which may cause differences in dependent variables between animal groups sampled. 'Data type' was a categorical variable with six categories (for example, breeding individuals, spawning biomass) to control for the different units and sampling methods used to collect the abundance data. Furthermore, the interaction between data type and CV was used to control for statistical artefacts in the time series (for example, harvested populations may be more variable)⁶. The model selection results are included in Supplementary Tables 1–3.

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data needed to reproduce the analysis can be found on Github (<https://doi.org/10.5281/zenodo.3470260>).

Code availability

The code needed to reproduce the analysis can be found on Github (<https://doi.org/10.5281/zenodo.3470260>).

Received: 29 October 2018; Accepted: 31 October 2019;

Published online: 09 December 2019

References

- Scott, A. C. *The Nonlinear Universe: Chaos, Emergence, Life* (Springer, 2007).
- Schaffer, W. M. & Kot, M. Do strange attractors govern ecological systems? *BioScience* **35**, 342–350 (1985).
- Pascual, M. & Ellner, S. P. Linking ecological patterns to environmental forcing via nonlinear time series. *Ecology* **81**, 2767–2780 (2000).
- Hsieh, C. H., Glaser, S. M., Lucas, A. J. & Sugihara, G. Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. *Nature* **435**, 336–340 (2005).
- May, R. M. & Oster, G. F. Bifurcations and dynamic complexity in simple ecological models. *Am. Nat.* **110**, 573–599 (1976).
- Glaser, S. M. et al. Complex dynamics may limit prediction in marine fisheries. *Fish. Fish. (Oxf)* **15**, 616–633 (2014).
- May, R. M. Simple mathematical models with very complicated dynamics. *Nature* **261**, 459–467 (1976).
- Turchin, P. & Taylor, A. D. Complex dynamics in ecological time series. *Ecology* **73**, 289–305 (1992).
- Hassell, M. P., Lawton, J. H. & May, R. M. Patterns of dynamical behaviour in single-species populations. *J. Anim. Ecol.* **45**, 471–486 (1976).
- Mueller, L. D. & Joshi, A. *Stability in Model Populations* (Princeton Univ. Press, 2000).
- Dennis, B., Desharnais, R. A., Cushing, J. M., Henson, S. M. & Costantino, R. F. Estimating chaos and complex dynamics in an insect population. *Ecol. Monogr.* **71**, 277–304 (2001).
- May, R. M. Biological populations with nonoverlapping generations: stable points, stable cycles, and chaos. *Science* **186**, 645–647 (1974).
- Dixon, P. A., Milicich, M. & Sugihara, G. Episodic fluctuations in larval supply. *Science* **283**, 1528–1530 (1999).
- Sugihara, G. Nonlinear forecasting for the classification of natural time series. *Philos. Trans. Royal Soc. A* **348**, 477–495 (1994).
- May, R. M., Beddington, J. R., Clark, C. W., Holt, S. J. & Laws, R. M. Management of multispecies fisheries. *Science* **205**, 267–277 (1979).
- Fogarty, M. J., Gamble, R. & Perretti, C. T. Dynamic complexity in exploited marine ecosystems. *Front. Ecol. Evol.* **4**, 68 (2016).
- Klein, E. S., Glaser, S. M., Jordaan, A., Kaufman, L. & Rosenberg, A. A. A complex past: historical and contemporary fisheries demonstrate nonlinear dynamics and a loss of determinism. *Mar. Ecol. Prog. Ser.* **557**, 237–246 (2016).
- Polis, G. A., Sears, A. L. W., Huxel, G. R., Strong, D. R. & Maron, J. When is a trophic cascade a trophic cascade? *Trends Ecol. Evol.* **15**, 473–475 (2000).
- Shurin, J. B. et al. A cross-ecosystem comparison of the strength of trophic cascades. *Ecol. Lett.* **5**, 785–791 (2002).
- McCann, K., Hastings, A. & Huxel, G. Weak trophic interactions and the balance of nature. *Nature* **395**, 794–798 (1998).
- Ziebarth, N. L., Abbott, K. C. & Ives, A. R. Weak population regulation in ecological time series. *Ecol. Lett.* **13**, 21–31 (2010).
- Zhou, X. et al. Detecting chaotic dynamics of insect populations from long-term survey data. *Ecol. Entomol.* **22**, 231–241 (1997).
- Leibold, M. A. & Chase, J. M. *Metacommunity Ecology* (Princeton Univ. Press, 2018).
- de Ruiter, P. C., Neutel, A. M. & Moore, J. C. Energetics, patterns of interaction strengths, and stability in real ecosystems. *Science* **269**, 1257–1260 (1995).
- Scheffer, M., Carpenter, S., Foley, J. A., Folke, C. & Walker, B. Catastrophic shifts in ecosystems. *Nature* **413**, 591–596 (2001).
- Scheffer, M. et al. Early-warning signals for critical transitions. *Nature* **461**, 53–59 (2009).
- Dakos, V., Glaser, S. M., Hsieh, C. & Sugihara, G. Elevated nonlinearity as an indicator of shifts in the dynamics of populations under stress. *J. R. Soc. Interface* **14**, 20160845 (2017).
- Bjornstad, O. N. & Grenfell, B. T. Noisy clockwork: time series analysis of population fluctuations in animals. *Science* **293**, 638–643 (2001).
- Coulson, T., Rohani, P. & Pascual, M. Skeletons, noise and population growth: the end of an old debate? *Trends Ecol. Evol.* **19**, 359–364 (2004).
- Hsieh, C. H., Anderson, C. & Sugihara, G. Extending nonlinear analysis to short ecological time series. *Am. Nat.* **171**, 71–80 (2008).
- Sugihara, G. & May, R. M. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* **344**, 734–741 (1990).
- Ye, H. et al. Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proc. Natl Acad. Sci. USA* **112**, E1569–E1576 (2015).
- Giron-Nava, A. et al. Quantitative argument for long-term ecological monitoring. *Mar. Ecol. Prog. Ser.* **572**, 269–274 (2017).
- Ebisuzaki, W. A method to estimate the statistical significance of a correlation when the data are serially correlated. *J. Clim.* **10**, 2147–2153 (1997).
- Herrando-Pérez, S., Delean, S., Brook, B. W. & Bradshaw, C. J. A. Strength of density feedback in census data increases from slow to fast life histories. *Ecol. Evol.* **2**, 1922–1934 (2012).
- Brook, B. W., Traill, L. W. & Bradshaw, C. J. Minimum viable population sizes and global extinction risk are unrelated. *Ecol. Lett.* **9**, 375–382 (2006).
- Costantino, R. F., Desharnais, R. A., Cushing, J. & Dennis, B. Chaotic dynamics in an insect population. *Science* **275**, 389–391 (1997).
- Turchin, P. & Ellner, S. P. Living on the edge of chaos: population dynamics of Fennoscandian voles. *Ecology* **81**, 3099–3116 (2000).
- Takens, F. Detecting strange attractors in turbulence. In *Proc. Symposium on Dynamical Systems and Turbulence* (eds Rand, D. & Young, L.) 366–381 (Springer, 1981).
- Ward, E. J., Holmes, E. E., Thorson, J. T. & Collen, B. Complexity is costly: a meta-analysis of parametric and non-parametric methods for short-term population forecasting. *Oikos* **123**, 652–661 (2014).
- Mouquet, N. et al. Predictive ecology in a changing world. *J. Appl. Ecol.* **52**, 1293–1310 (2015).
- Petchey, O. L. et al. The ecological forecast horizon, and examples of its uses and determinants. *Ecol. Lett.* **18**, 597–611 (2015).
- 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).
- Perretti, C. T., Sugihara, G. & Munch, S. B. Nonparametric forecasting outperforms parametric methods for a simulated multispecies system. *Ecology* **94**, 794–800 (2013).
- 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).
- 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).
- Graham, N. A., Jennings, S., MacNeil, M. A., Mouillot, D. & Wilson, S. K. Predicting climate-driven regime shifts versus rebound potential in coral reefs. *Nature* **518**, 94–97 (2015).
- Drake, J. M. & Griffen, B. D. Early warning signals of extinction in deteriorating environments. *Nature* **467**, 456–459 (2010).
- Fryxell, J. M., Packer, C., McCann, K., Solberg, E. J. & Saether, B. E. Resource management cycles and the sustainability of harvested wildlife populations. *Science* **328**, 903–906 (2010).
- Keith, D. et al. Temporal correlations in population trends: conservation implications from time-series analysis of diverse animal taxa. *Biol. Conserv.* **192**, 247–257 (2015).
- Hinchliff, C. E. et al. Synthesis of phylogeny and taxonomy into a comprehensive tree of life. *Proc. Natl Acad. Sci. USA* **112**, 12764–12769 (2015).
- Michonneau, F., Brown, J. W. & Winter, D. J. rot: an R package to interact with the Open Tree of Life data. *Methods Ecol. Evol.* **7**, 1476–1481 (2016).
- R Core Team. *R: A Language and Environment for Statistical Computing* (R Foundation for Statistical Computing, 2018).
- Ye, H. et al. rEDM: Applications of empirical dynamic modeling from time series. R package version 0.7.5 <https://github.com/ha0ye/rEDM> (2018).

55. Paradis, E., Claude, J. & Strimmer, K. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics*. **20**, 289–290 (2004).
56. Gelman, A., Jakulin, A., Pittau, M. G. & Su, Y.-S. A weakly informative default prior distribution for logistic and other regression models. *Ann. Appl. Stat.* **2**, 1360–1383 (2008).
57. Gelman, A. et al. arm: Data analysis using regression and multilevel/hierarchical models. R package version 1.10-1 <https://CRAN.R-project.org/package=arm> (2018).

Acknowledgements

Funding for this project was provided by the National Science Foundation (NSF) Graduate Research Fellowship under grant no. 366280 (T.J.C.), the Idaho Department of Fish and Game (T.J.C.) and NSF grant no. 1836793 (A.D.L.). We thank H. Wilhelm Martin for his superb assistance with the analysis and M. Hebblewhite and J. Maron for comments on earlier drafts.

Author contributions

T.J.C. collected and analysed the data and wrote the manuscript. A.D.L. supervised the project and edited the manuscript.

Additional information

Extended data is available for this paper at <https://doi.org/10.1038/s41559-019-1052-6>.

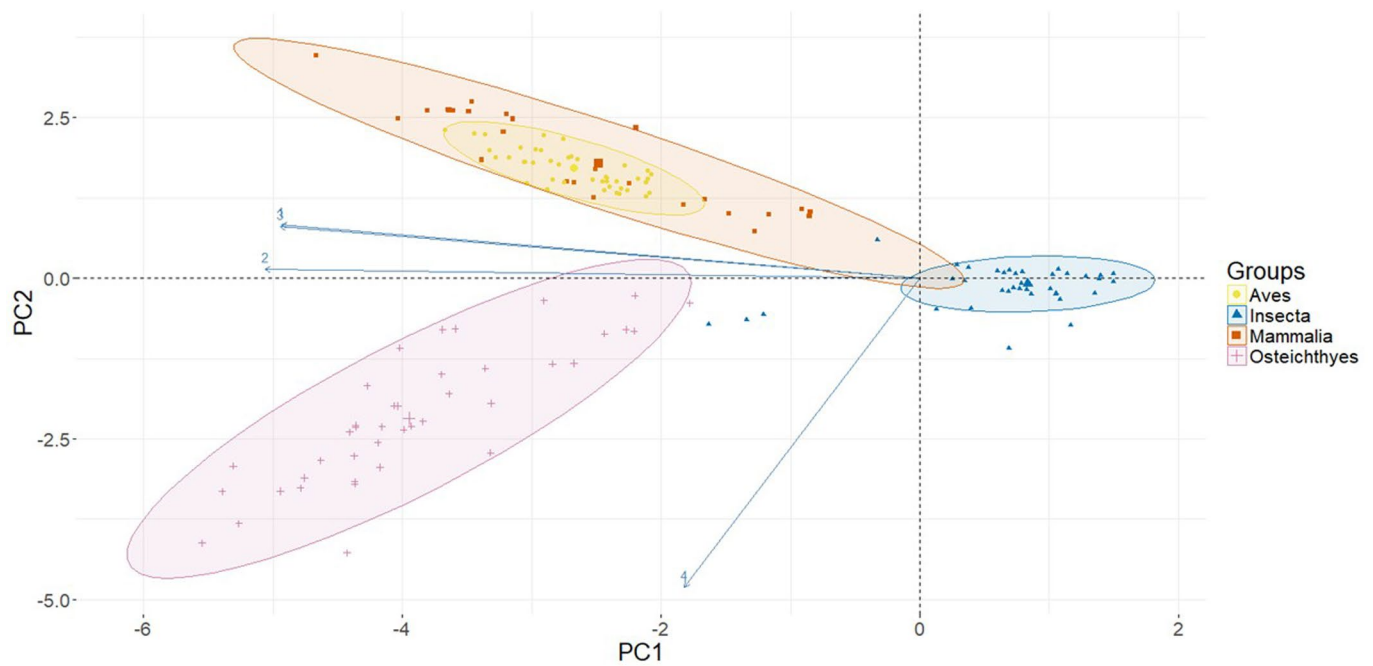
Supplementary information is available for this paper at <https://doi.org/10.1038/s41559-019-1052-6>.

Correspondence and requests for materials should be addressed to T.J.C.

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

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

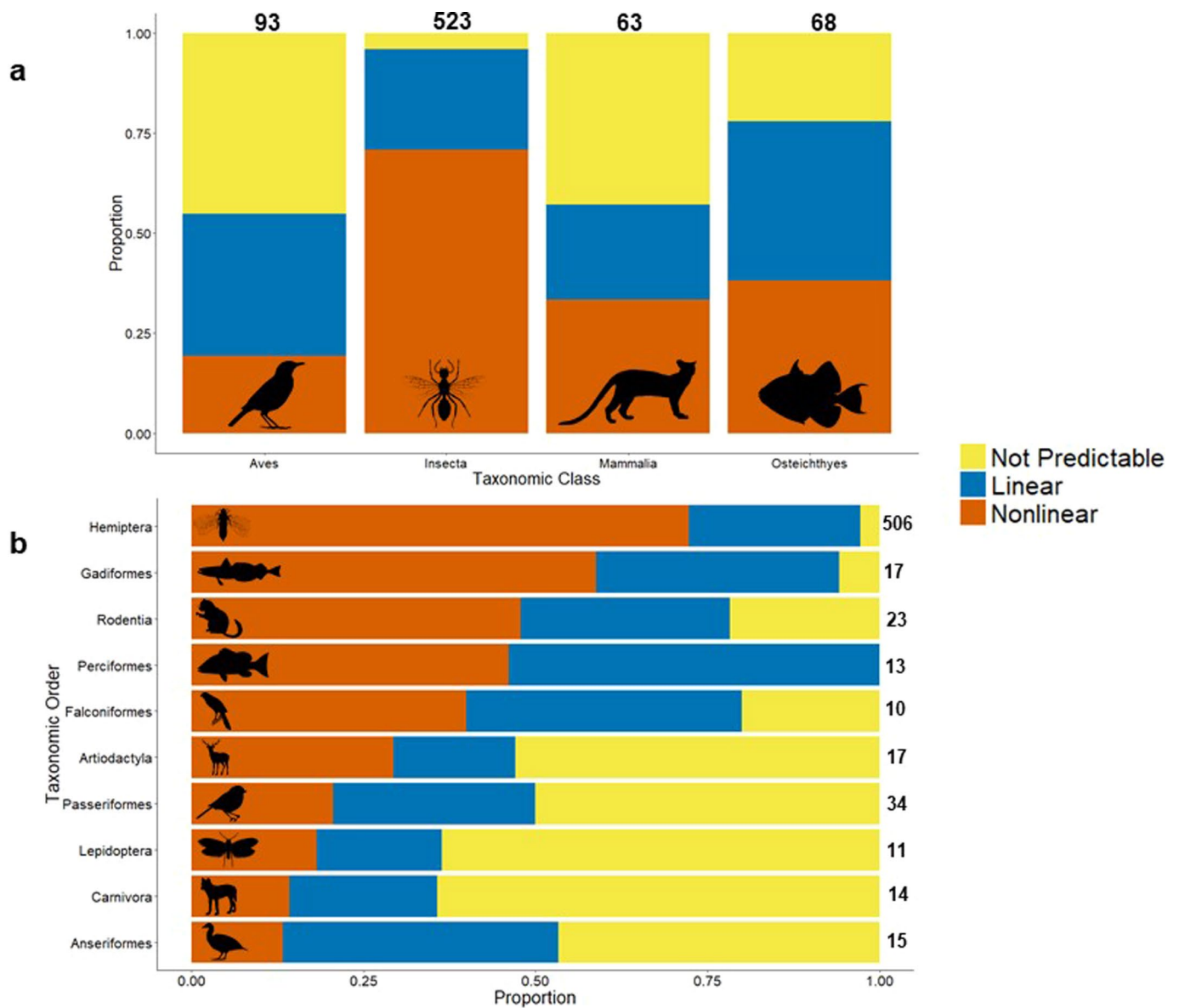
© The Author(s), under exclusive licence to Springer Nature Limited 2019



Extended Data Fig. 1 | Principal components analysis of life-history traits. PC1 explains 72.2% of variation in our data (axes 1, 2, and 3), representing body length (mm), minimum age at first reproduction (months), and lifespan (months), respectively. PC2 explains 24.9% of variation in our data (axis 4), representing fertility (# of young per year). Colored ellipses represent 95% probability that data for each taxonomic classification fall within the ellipse.

Final Model Intercepts	Coefficient Estimates (log odds)	Standard Errors
Taxonomic Class - Aves	-0.164	0.562
Taxonomic Class - Insecta	0.691	0.774
Taxonomic Class - Mammalia	0.442	0.509
Taxonomic Class - Osteichthyes	0.054*	0.466
PC1	-0.113	0.216
E	-0.067	0.064
ρ	0.734*	0.200
CV	-0.718	1.103

Extended Data Fig. 2 | Final model results of nonlinearity. E is the embedding dimensionality, ρ is the forecast skill, and CV is the coefficient of variation of a time-series. An asterisk indicates coefficients that were significant at $P \leq 0.05$.

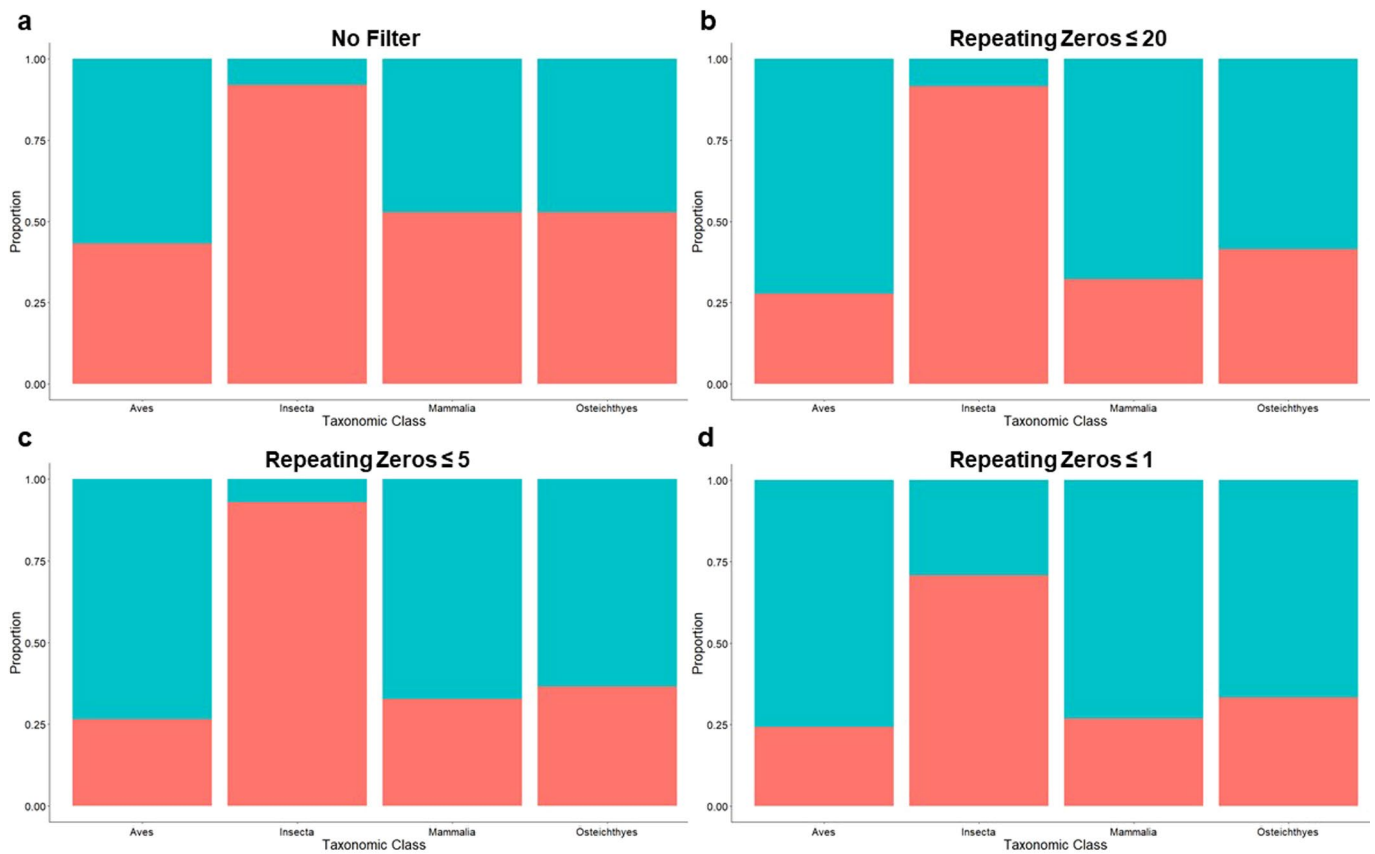


Final Model Intercepts	Coefficient Estimates	Standard Errors
PC1	0.090	0.065
N	0.383*	0.149

Extended Data Fig. 4 | Final model results of dimensionality (E). PC1 is the first principal component of life history traits, representing a combination of body length (mm), minimum age of first reproduction (months), and longevity (months) of animals (positive coefficient estimates = faster life histories; Extended Data Fig. 1). N is the time-series length. An asterisk indicates coefficients that were significant at $P \leq 0.05$.

Final Model Intercepts	Coefficient Estimates	Standard Errors
Taxonomic Class - Aves	0.661*	0.132
Taxonomic Class - Insecta	-0.226	0.172
Taxonomic Class - Mammalia	0.330*	0.095
Taxonomic Class - Osteichthyes	0.308*	0.101
PC1	0.221*	0.049
Linear	0.480*	0.095
Nonlinear	0.309*	0.087

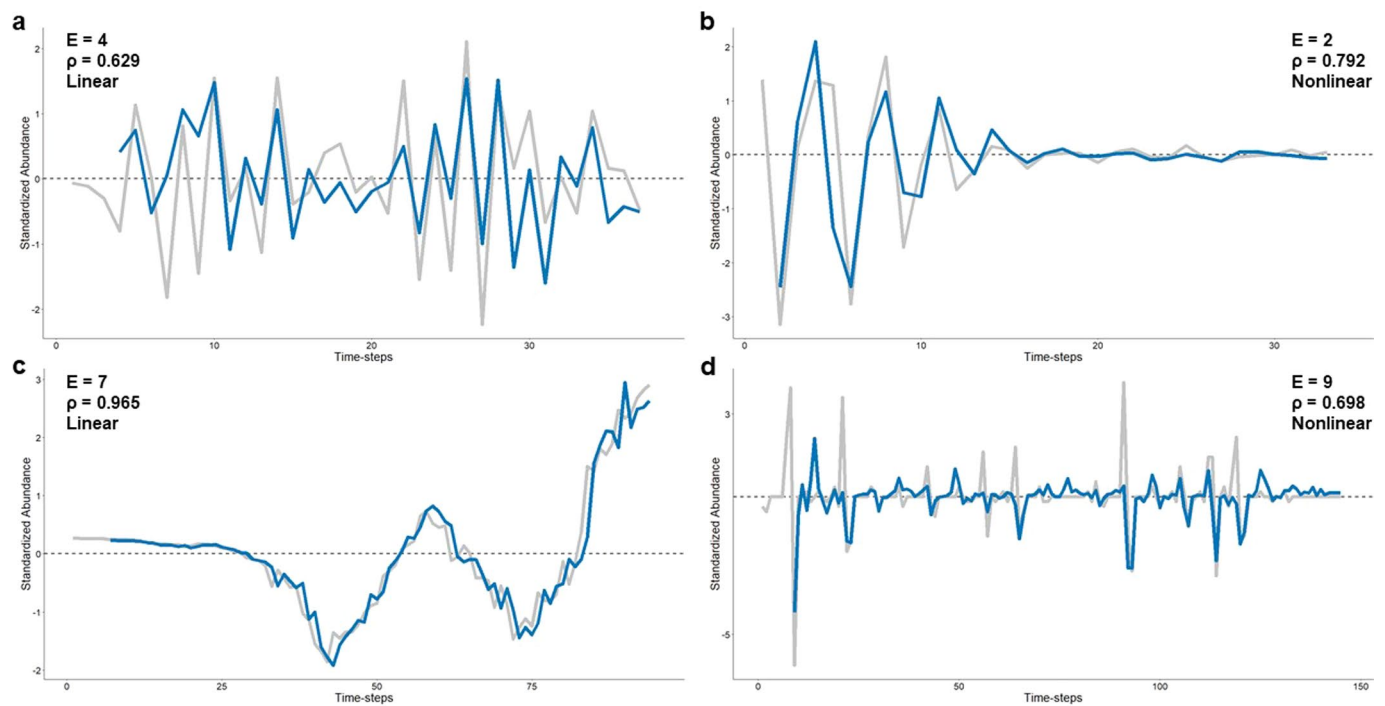
Extended Data Fig. 5 | Final model results of forecast skill (ρ). PC1 is the first principal component of life history traits, representing a combination of body length (mm), minimum age of first reproduction (months), and longevity (months) of animals. Linear, nonlinear, and not predictable represent the categorization of population dynamics. CV is the coefficient of variation of a time-series. An asterisk indicates coefficients that were significant at $P \leq 0.05$.



Extended Data Fig. 6 | Repeating zeroes in datasets do not change likelihood of nonlinearity. Proportion of linearity/nonlinearity in animal time-series, arranged by taxonomic class. Due to some time-series having long sequences of zeroes, we filtered out time-series with strings of zeroes. **a**, Time-series with no filtering. **b**, Strings of zeroes > 20 filtered. **c**, Strings of zeroes > 5 filtered. **d**, Strings of zeroes > 1 filtered.

Taxonomic Class	Median Time-series Length	# Datasets	# Predictable Datasets	# Species	# Predictable Species
Aves	44	93	51	75	44
Insecta	140	523	502	49	44
Mammalia	40	63	36	50	29
Osteichthyes	96	68	53	54	46
Total	47.5	747	642	228	163

Extended Data Fig. 7 | Summary statistics for animal time-series by taxonomic class. Median time-series length represents median number of time-series data for the final dataset. Predictable datasets were categorized if the Pearson correlation coefficient of out-of-sample prediction was significant at $P \leq 0.05$.



Extended Data Fig. 8 | Examples of predictable time-series. **a**, Standardized abundance of woodcock (*Scolopax minor*) over time. **b**, Standardized abundance of grey red-backed voles (*Myodes rufocanus*) over time. **c**, Standardized abundance of dover soles (*Solea solea*) over time. **d**, Standardized abundance of woolly beech aphids (*Phyllaphis fagi*) over time. Grey lines represent the observed abundance, blue lines represent predicted abundance.

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see [Authors & Referees](#) and the [Editorial Policy Checklist](#).

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a Confirmed

- ☐ ☒ The exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement
- ☐ ☒ A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
- ☐ ☒ The statistical test(s) used AND whether they are one- or two-sided
Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- ☐ ☒ A description of all covariates tested
- ☐ ☒ A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
- ☐ ☒ A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
- ☒ ☐ For null hypothesis testing, the test statistic (e.g. F , t , r) with confidence intervals, effect sizes, degrees of freedom and P value noted
Give P values as exact values whenever suitable.
- ☒ ☐ For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
- ☐ ☒ For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
- ☒ ☐ Estimates of effect sizes (e.g. Cohen's d , Pearson's r), indicating how they were calculated

Our web collection on [statistics for biologists](#) contains articles on many of the points above.

Software and code

Policy information about [availability of computer code](#)

Data collection

Data were primarily collected from the Global Population Dynamics Database (NERC Centre for Population Biology, Imperial College (2010). <http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html>.) The raw dataset, summary of results, and citations are available at: <http://doi.org/10.5281/zenodo.3470260>

Data analysis

"rEDM" package v0.6.18 (Ye et al. 2018; <https://github.com/ha0ye/rEDM>) of R v3.4.2 (R Core Team 2018) was primarily used to analyze the data. R code is available at: <http://doi.org/10.5281/zenodo.3470260>

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research [guidelines for submitting code & software](#) for further information.

Data

Policy information about [availability of data](#)

All manuscripts must include a [data availability statement](#). This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Code and data to reproduce the analysis can be found on Github: <http://doi.org/10.5281/zenodo.3470260>.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☐ Life sciences ☐ Behavioural & social sciences ☒ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see [nature.com/documents/nr-reporting-summary-flat.pdf](https://www.nature.com/documents/nr-reporting-summary-flat.pdf)

Ecological, evolutionary & environmental sciences study design

All studies must disclose on these points even when the disclosure is negative.

Study description	We studied the presence of nonlinearity, dimensionality, and predictability across a large dataset of animal population time-series. Our study design was not experimental.
Research sample	Data were primarily collected from the Global Population Dynamics Database (NERC Centre for Population Biology, Imperial College (2010). http://www.sw.ic.ac.uk/cpb/cpb/gpdd.html .) The raw dataset, summary of results, and citations are available at: http://doi.org/10.5281/zenodo.3470260 .
Sampling strategy	Time-series were chosen from knowledge of past research using datasets (e.g., GPDD), and by conducting a literature search with keywords such as "time-series". We collected time-series until we subjectively decided that we had enough time-series of species (~50) for each taxonomic classification to conduct a robust statistical analysis.
Data collection	Data was collected by T.C. as aforementioned using existing databases and literature searches.
Timing and spatial scale	Not applicable.
Data exclusions	Data was filtered from our full set because of inconsistencies and ambiguities, such as long repeating non-unique strings of values for population data. Additionally, we only conducted statistical analysis for one randomly assigned time-series for each species, due to multiple time-series for certain species like aphids and moths. The filtering process for our dataset is described in full in Methods and Supplementary Information Methods.
Reproducibility	Time-series observations were not replicated. The analysis can be reproduced using datasets and code made publicly available by the author.
Randomization	Any manipulative experiments were not performed and thus randomization is not relevant to our study.
Blinding	Any manipulative experiments were not performed and thus randomization is not relevant to our study.
Did the study involve field work?	<input type="checkbox"/> Yes <input checked="" type="checkbox"/> No

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> Antibodies
<input checked="" type="checkbox"/>	<input type="checkbox"/> Eukaryotic cell lines
<input checked="" type="checkbox"/>	<input type="checkbox"/> Palaeontology
<input checked="" type="checkbox"/>	<input type="checkbox"/> Animals and other organisms
<input checked="" type="checkbox"/>	<input type="checkbox"/> Human research participants
<input checked="" type="checkbox"/>	<input type="checkbox"/> Clinical data

Methods

n/a	Involved in the study
<input checked="" type="checkbox"/>	<input type="checkbox"/> ChIP-seq
<input checked="" type="checkbox"/>	<input type="checkbox"/> Flow cytometry
<input checked="" type="checkbox"/>	<input type="checkbox"/> MRI-based neuroimaging