OIKOS

Research

Unifying ecosystem responses to disturbance into a single statistical framework

Nathan P. Lemoine

N. P. Lemoine (https://orcid.org/0000-0003-3436-7196) ✉ (nathan.lemoine@marquette.edu), Dept of Biological Sciences, Marquette Univ., Milwaukee, WI, USA, and: Dept of Zoology, Milwaukee Public Museum, Milwaukee, WI, USA.

Oikos
130: 408–421, 2021
doi: 10.1111/oik.07752

Subject Editor: Pedro Peres-Neto
Editor-in-Chief: Dries Bonte
Accepted 24 November 2020

Natural ecosystems are currently experiencing unprecedented rates of anthropogenic disturbance. Given the potential ramifications of more frequent disturbances, it is imperative that we accurately quantify ecosystem responses to severe disturbance. Specifically, ecologists and managers need estimates of resistance and recovery from disturbance that are free of observation error, not biased by temporal stochasticity and that standardize disturbance magnitude among many disparate ecosystems relative to normal interannual variability. Here, I propose a statistical framework that estimates all four components of ecosystem responses to disturbance (resistance, recovery, elasticity and return time), while resolving all of the issues described above. Coupling autoregressive time series with exogenous predictors (ARX) models with impulse response functions (IRFs) allows researchers to statistically subject all ecosystems to similar levels of disturbance, estimate lag effects and obtain standardized estimates of resistance to and recovery from disturbance that are free from observation error and stochastic processes inherent in raw data.

Keywords: autoregression, climate change, statistics, stress, times series

Introduction

Natural ecosystems are currently experiencing unprecedented rates of disturbance due to human activity. Such disturbances immediately impact ecosystem function and can also often impair ecosystem function long after the disturbance itself has abated (Smith 2011). More frequent and severe droughts, for example, cause long-term shifts in plant community composition, widespread tree mortality and catastrophic declines in primary production (Ciais et al. 2005, Anderegg et al. 2015, Knapp et al. 2015a). Heat waves often co-occur with drought (Perkins et al. 2012), exacerbating soil water loss in terrestrial ecosystems and causing extensive mortality of foundational, habitat-forming species in marine systems (Ciais et al. 2005, Le Nohaïc et al. 2017, Smale et al. 2019). Other pulse disturbances, such as hurricanes or abnormal frost events, can also initiate rapid and lasting changes in ecosystem function (Lodge and McDowell 1991, Lirman et al. 2011). Hence, mitigating the consequences of disturbance with preventative or restorative actions, or forecasting how ecosystems will respond to future global change, are important goals in ecological research. However, achieving these

goals first requires that we improve our ability to accurately quantify ecosystem change following disturbance, predict trajectories of recovery and identify the abiotic and biotic constraints that dictate ecosystem responses to disturbance (Nimmo et al. 2015).

Ecosystem change following disturbance consists of four quantifiable components (Hodgson et al. 2015), hereafter referred to collectively as the ecosystem response to disturbance. First resistance (a.k.a. sensitivity) to disturbance describes the degree to which an ecosystem is instantaneously impacted by disturbance. Recovery from disturbance refers to the ability of an ecosystem to return to a stable state immediately following disturbance. Elasticity and return time refer to the rate and duration of time, respectively, required for an ecosystem to return to stability. The related concept of lag effects more broadly describes whether an altered ecosystem state persists after the disturbance has been removed (Sala et al. 2012); recovery, elasticity and return time are all quantifiable aspects of a lag effect. Thus, lag effects exist when an ecosystem does not recovery fully following disturbance and exhibits a non-zero return time. Although these four components have already been extensively studied in various ecosystems and with regard to different types of disturbance, they remain difficult to quantify in a standardized and accurate manner that enables syntheses and comparisons among ecosystems.

The most significant obstacles to synthesizing information are the inconsistent terminology and diverse methods that ecologists use to measure ecosystem responses. The terms resistance and sensitivity have been used interchangeably and possess at least five mathematical formulations in the drought literature alone (Table 1). Second, comparing estimates of resistance among sites or years is difficult even when using a single metric due to analytical flaws. Consider the example of estimating the instantaneous impact of extreme drought on aboveground net primary production (ANPP). Calculating resistance as the ratio of ANPP during drought to ANPP of the previous year (Lloret et al. 2011, Gazol et al. 2017, 2018, Stuart-Haëntjens et al. 2018) incorporates both temporal stochasticity and observation errors that can lead to inaccurate results. Moreover, using this metric, the context of drought resistance varies from site-to-site and year-to-year, rending impossible any inter- or even intra-site comparisons of ecosystem resistance to drought. Attempts to describe ecosystem recovery are even more inconsistent in both terminology and methods (Table 1), and suffer from many of the same analytical flaws. Here, I propose a statistical framework based on econometric techniques that simultaneously estimates all four components of ecosystem responses to disturbance using accepted, standardized language while resolving statistical problems with previous methods. Such a method is sorely needed, as most studies of ecosystem responses to disturbance analyze time series data but fail to use the appropriate autoregressive time series approaches (Kannenberg et al. 2020).

Problems with prior methods

The methods outlined in Table 1 consist of either ratio-based or regression-based procedures. Both methodologies possess

Table 1. Definitions and mathematical equations used to calculate ecosystem resistance, resilience, recovery and legacy effects following an extreme stress event.

Method	Name	Equation	Units	Citation
Reduction during stress				
1	Sensitivity	$\frac{x_t - x_{t-1}}{\text{ppt}_t - \text{ppt}_{t-1}}$	Change in primary production per mm change in rainfall	Wilcox et al. 2017
2	Sensitivity	$\frac{\Delta x}{\Delta \text{ppt}}$	Slope of the primary production – precipitation relationship	Huxman et al. 2004b, Knapp et al. 2015b
3	Sensitivity	$100 \times \left(\frac{x_t - \overline{x}}{\overline{x}} \right)$	Percent decline from long-term mean	Griffin-Nolan et al. 2018
4	Resistance	$\frac{x_t}{x_{t-1}}$	Proportion decline from pre-drought year	Lloret et al. 2011, Gazol et al. 2017, 2018
5	Resistance	$\ln\left(\frac{x_t}{x_{t-1}} \right)$	Log proportion decline from pre-drought year	Stuart-Haëntjens et al. 2018
Return following stress				
6	Recovery	$\frac{x_{t+1}}{x_t}$	Proportion increase in post-drought year	Lloret et al. 2011, Gazol et al. 2017, 2018
7	Resilience	$\frac{x_{t+1}}{x_{t-1}}$	Proportion decrease in post-drought year from pre-drought	Lloret et al. 2011
8	Resilience	$\ln\left(\frac{x_{t+1}}{x_{t-1}} \right)$	Log proportion decrease in post-drought year from pre-drought	Stuart-Haëntjens et al. 2018
9	Legacy effects	$100 \times \left(\frac{x_{t_1} - \overline{x}}{x_{t+1}} \right)$	Percent decrease in post-drought year from long-term mean	Griffin-Nolan et al. 2018
10	Legacy effects	$x_{t+1} + \hat{x}_{t+1}$	Observed – predicted for post-drought year	Sala et al. 2012, Anderegg et al. 2015

statistical and logical issues that should discourage their future application.

Ratio-based methods

The first problem with ratio-based methods is that, by working with raw observations, estimates of resistance and recovery include observation error. Sampling errors, spatial heterogeneity or any number of other processes can cause significant variation in the data and potentially yield inaccurate estimates of resistance and recovery. To illustrate the extent to which sampling error induces inaccuracies, I conducted a simple simulation experiment to estimate the resistance and recovery of ANPP from drought. For each of the 10 000 simulations, I assumed a true average ANPP of $x = 500$ g m^{-2} for ten time steps (i.e. years), similar to levels observed at the Konza Prairie Biological Station (Griffin-Nolan et al. 2018). I then imposed drought during the second time step using a resistance value randomly chosen from a Uniform(0.5,1) distribution, where $x_2 = x_1 \times$ resistance. I allowed ANPP to recover at the next time step with a recovery value randomly chosen from a Uniform(1,2) distribution, such that $x_3 = x_2 \times$ recovery. Next, I imposed a sampling error of $\pm$ 40 g m^{-2} to the true values, similar in magnitude to sampling uncertainty at KPBS (Griffin-Nolan et al. 2018), to create a new variable y of observed ANPP values: $y \sim N(x, 40^2)$. From the observed data y, I calculated resistance as y_2/y_1 and recovery as y_3/y_2 (Table 1). I repeated this simulation for a low productivity site, setting mean ANPP to 100 g m^{-2} and sampling error to $\pm$ 15 g m^{-2}, representative of the semi-arid shortgrass steppe in northern Colorado (Griffin-Nolan et al. 2018).

As expected, estimates of resistance and recovery using observed data (i.e. with noise) were highly variable. Although the observed values for resistance and recovery were both centered around the true value, any single observation of resistance and recovery could deviate significantly from the true value (Fig. 1). For example, observed resistance varied from approximately 0.3–0.75 for a true value of 0.5 at the highly productive site (Fig. 1A). At the low productivity site, errors were magnified; observed resistance varied from approximately 0.2–0.9 for a true value of 0.5 (Fig. 1A). The same patterns held for recovery, wherein sampling error resulted in large deviations from the true value and errors were larger for the less productive site (Fig. 1B). Large errors at the less productive site result from difficulties in using ratios; small deviations from the true value cause a larger proportional difference when the denominator (e.g. y_1 for resistance, y_2 for recovery) is small. These simulations also assumed that the true value of ecosystem function during normal years was constant, but temporal variability in ecosystem properties would compound inaccuracies arising from sample errors. Ecosystem resistance to, or recovery from, drought might therefore be overestimated if the pre-drought year is dry (thereby reducing the decline in ANPP caused by drought and seemingly increasing resistance) or the post-drought year is abnormally wet (thereby increasing the magnitude of recovery), further magnifying the lack of precision.

Finally, it is difficult to place estimates of ecosystem responses to disturbance from ratio-based methods into a proper context that enables cross-site comparisons. Continuing the drought example, the degree of disturbance induced by a given rainfall reduction varies with regional climate conditions. In other words, a 200 mm reduction in annual rainfall imposes a much stronger meterological drought in the arid shortgrass steppe than it does in mesic tallgrass prairies (Knapp et al. 2015b). However, most ratio-based methods do not incorporate such context-specificity (but see Wilcox et al. 2017). The lack of context-specificity

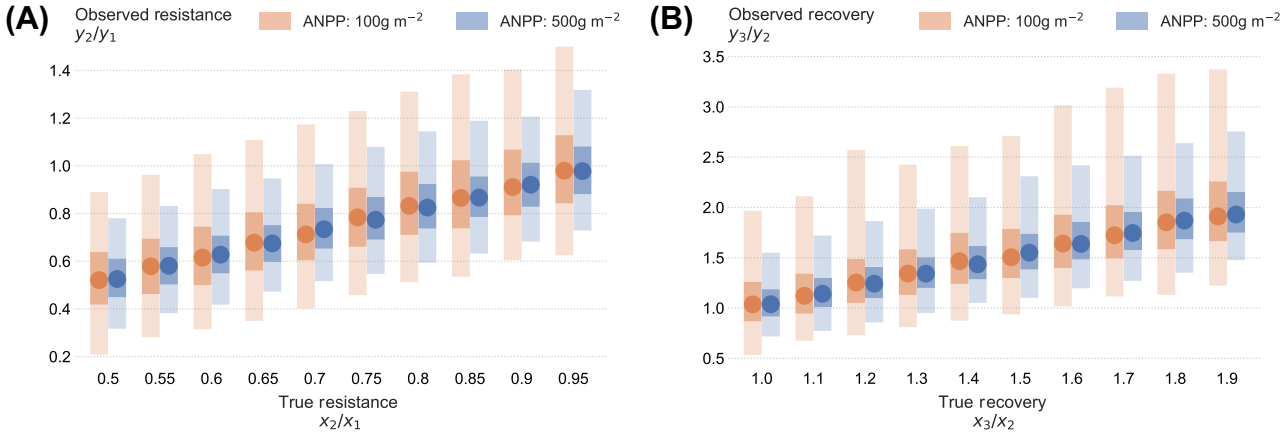

Figure 1. Relationship between (A) true resistance and observed resistance and (B) true recovery and observed recovery using simulated data. I simulated a time series of 10 observations for a given value of ANPP (100 or 500 g m^{-2}). For each simulation, I chose a random resistance value from a Uniform(0.5,1) distribution and a random recovery value from a Uniform(1,2) distribution. I set x_2 as the drought year, with a true ANPP equal to ANPP $\times$ resistance, and I set x_3 as the recovery year with a true ANPP equal to $x_2 \times$ recovery. I then introduced a sampling error of 10 or 40 g m^{-2} for the low and high-productivity simulations, respectively. I calculated the observed resistance and recovery values from the sampled data. Points show the median observed value, dark inner bars show the 50% quantiles, and light outer bars show the 95% quantiles.

inhibits accurate cross-site or cross-time comparisons of ecosystem responses to disturbance.

Regression-based methods

Perhaps the most popular method for identifying lag effects of disturbance is to calculate the predicted ecosystem state using regression and then estimating the degree of recovery as the residual error of the observed recovery from the predicted recovery. For example, to estimate the recovery of ANPP following drought, one would first regress ANPP against annual precipitation. The regression equation provides the predicted ANPP in the year following drought based on rainfall, and the residual of the post-drought year (observed-predicted) constitutes the lag effect. The shortgrass steppe of Colorado experienced an extreme drought in 2012 (Fig. 2A – red dot). Based on the ANPP–precipitation relationship, we can estimate the predicted value of ANPP in 2013 (Fig. 2B). The lag effect is then the observed ANPP in 2013 (Fig. 2A–B – green dot) minus the predicted value in 2013 (Fig. 2B).

However, identifying lag effects and quantifying recovery using this method suffers from logical and statistical issues. Statistically, observations will almost never fall exactly along the regression line; by definition of the sum-to-zero property of residuals, half the points will be above and half will be below the line. Thus, every observation will exhibit a lag effect to some degree. To circumvent this problem, some researchers consider only 'significant' lag effects, wherein the observation falls outside of the 95% CI of the regression line (Griffin-Nolan et al. 2018, Fig. 2B). Yet this method is akin to testing whether a single observation is exactly equal to the mean value, and says more about certainty of the mean value than the presence of any lag effect. The recovery point might fall outside the 95% CI due to observation error, and the likelihood that a point falls outside the 95% CI of the mean increases rapidly as sample size increases because the width of the 95% CI decreases proportionally to the inverse square-root of sample size (Fig. 2C). For long time series, over 80% of observations would be considered significant when compared to the 95% CI of the mean, even when data were generated without autoregressive lags (Fig. 2C). Comparing the

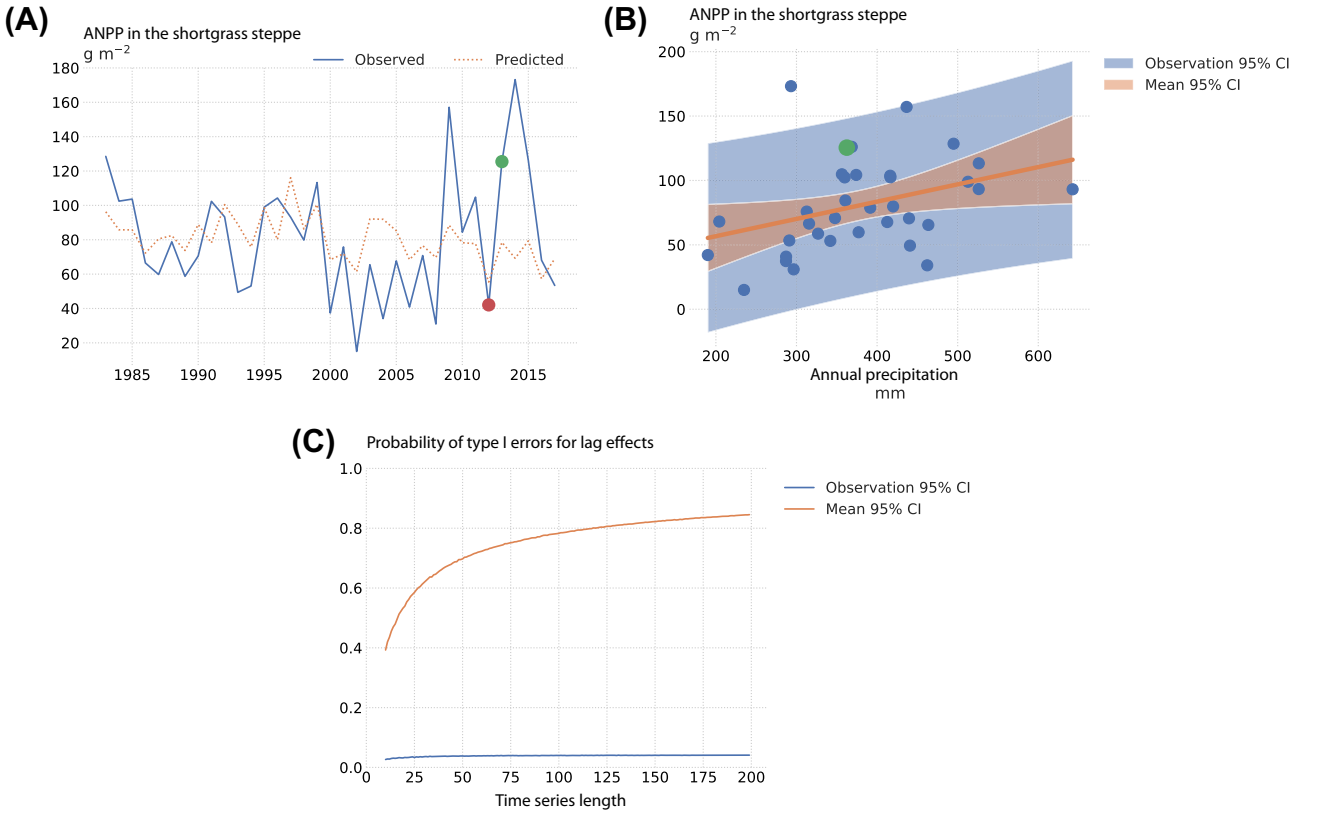

Figure 2. (A) Time series of ANPP at the Central Plains Experimental Range in the shortgrass steppe of Colorado. The orange dotted line shows the predicted ANPP based on the ANPP–precipitation relationship. Red dot shows the drought of 2012, green dot shows the year following the extreme drought. (B) The ANPP–precipitation relationship for the Central Plains Experimental Range. The orange envelope is the 95% CI of the mean, the blue envelope is the 95% CI of a prediction point. (C) Using the mean 95% CI (orange line) to statistically test for legacy effects results in high false positive rates as sample size increases and uncertainty about the mean decreases, while using the observation 95% CI (blue line) avoids this complication. Lines were generated by simulating 10 000 precipitation time series, then using a simulating primary production–precipitation relationship to estimate primary production in the absence of legacy effects. Type I error rates are the proportion of observations in a simulated time series that would be considered to possess significant legacy effects, despite being simulated without legacy effects.

presence of lag effects among sites might therefore simply be reporting differences in time series length. One solution is to use the observation CI instead of the mean CI. The mean CI is the most commonly plotted envelope and denotes the 95% CI of the regression line, equivalent to the 95% CI of the mean at any given point. The observation CI is the 95% of the individual observations at any given point and provides the envelope that is likely to contain a single observation. The mean and observation CIs are analogous to the standard error of the mean and standard deviation of the data, respectively. Using the observation CI opposed to the mean CI alleviates potential type I errors because the width of the observation CI depends only on residual error, not sample size (Fig. 2C).

Logically, quantifying recovery as the deviation from the predicted line assumes that the scatter surrounding the regression is caused entirely by lag effects. Is this true for all points, or only the single point in which the ecologist is interested? If lag effects are the only cause of scatter, incorporating autoregressive parameters into regression models should perfectly fit the data. Autoregressive parameters do improve fit, but they do not model data perfectly and often do not improve prediction accuracy for a single observation (Oesterheld et al. 2001). Alternatively, if lag effects apply only during the year following disturbance, do sources of variability present in other years not occur in the recovery year? If true, then lag effects should rarely switch signs, yet ANPP in the shortgrass steppe shows a small negative lag effect following the 2002 drought and a strong positive lag effect following the 2012 drought (thus impairing intra-site comparisons of lag effects, Fig. 2A). Either lag effects switch from positive to negative on a regular basis, or unaccounted sources of variation are influencing estimates of lag effects and recovery from raw data. In reality, many factors likely contribute to an imperfect primary production–precipitation relationship in all years, including the within-year distribution of rainfall event size and timing (Heisler-White et al. 2008), observation error, stochasticity in community composition and potential lag effects. Ecologists need a method that accurately separates the signal from the noise when assessing ecosystem responses to disturbance.

A statistical framework for ecosystem responses to disturbance

Here, I demonstrate that autoregressive models with exogenous predictors (ARX) and impulse response functions (IRFs) allow researchers to rigorously identify the presence of lag effects, subject all ecosystems to a similar level of disturbance, and obtain estimates of resistance, recovery, elasticity and return time that are free from stochastic processes inherent in raw data. ARX models and IRFs are particularly attractive because they are not only powerful, but simple to implement using any statistical software. ARX models are simply an autoregression model with an extra parameter and have already been used, if not identified by name, by ecologists examining lag effects of drought (Sala et al. 2012) or as autoregressive population models that include interspecific effects (Hansen et al. 1999). IRFs themselves require nothing more than a few basic calculations based on parameters from a fitted ARX model (Table 2). Thus, the IRF technique can be implemented by any ecologist using freely available software.

Impulse response functions

In econometrics, impulse response functions are simple calculations following time series analyses that describe the trajectory of dynamic systems following stress. They are particularly useful in systems that are costly or impossible to manipulate experimentally, such as financial markets. Indeed, economists have widely implemented IRFs to understand the resistance and recovery of financial markets to instantaneous 'shocks' (Creal and Wu 2017, Gambetti and Musso 2017). For example, Senbet (2016) used IRFs to visualize the consequences of higher federal interest rates on unemployment, consumption and other indicators of economic health. IRFs can also be used to understand how disturbances of different frequencies or press disturbances impact system dynamics. In medical studies, IRFs describe how the human body responds to elevated or depressed hormone activity (Schultz et al. 2015, Chang et al. 2017). Earth system modelers use IRFs to understand how global temperature or CO_2 concentrations respond to various disturbances, such as changes in oceanographic processes or vehicular emissions (Thompson and Randerson 1999, Joos et al. 2013, Millar et al. 2017, Zeng et al. 2017).

However, IRFs are currently only defined for a few autoregressive models that generally are not structured to test hypotheses about exogenous disturbances. In econometrics, IRFs exist only for univariate autoregressive (AR) models and multivariate vector autoregressive (VAR) models (Lutkepohl and Kratzig 2004, Bisgaard and Kulahci 2011, Box et al. 2015). AR models would allow ecologists to asses, for example, how systems recover from a shock to primary production, but there would be no link to an exogenous driver of production like precipitation. VARs include more than one process, but the processes are treated as a multivariate problem (Lutkepohl and Kratzig 2004, Bisgaard and Kulahci

Table 2. Analytical solutions for the simple $x^* = [\alpha, 0, 0, 0, \ldots, 0]$ presented here and depicted graphically in Fig. 3. IRF is the equation needed to graph the curve, or calculate the ecosystem state at any given time. Proofs are given in the Supporting information, as well as solutions for ARX(2) models. Note that these solutions assume that y_0 is the initial time step when $x_0^* = \alpha$ (not y_1, see x-axis of Fig. 3).

Model	IRF	Resistance	Recovery	Elasticity	Return time
ARX(0)	$y^*_t = \beta x^*_t$	$\beta\alpha$	n/a	n/a	n/a
ARX(1)	$y^*_t = \varphi_1{}^t \beta\alpha$	$\beta\alpha$	$\varphi_1\beta\alpha$	$\log(\varphi_1)$	$\log(0.05)/\log(\varphi_1)$

2011). For example, ANPP would be driven by lag effects of ANPP, current precipitation and lag effects of precipitation. However, precipitation would also be driven be lag effects of precipitation, current ANPP and lagged ANPP. Such bidirectional pathways do not make sense for most disturbance processes operating at small scales (e.g. coral cover and ocean temperatures). ARX models, on the other hand, include an independent exogenous prediction, but IRFs for ARX models have never been defined nor, to my knowledge, applied to disturbance problems in any field.

Identifying ARX–IRFs requires fitting ARX(p) models with time series data to first identify whether lag effects are present. ARX(p) models modify autoregressive AR models of order p by including one or more exogenous predictions:

$$y_t = \beta x_t + \varphi_1 y_{t-1} + \varphi_2 y_{t-2} + \ldots + \varphi_p y_{t-p} + \varepsilon_t$$

This model states that ecosystem state at time t (y_t) depends on contemporary exogenous values (x_t, e.g. annual precipitation, sea-surface temperature anomaly, etc.), previous ecosystem states up to p time steps in the past (y_{t-p}, i.e. lag effects), and error from both unmeasured processes and sampling issues (ε_t). The appropriate order p can be chosen via information theoretic methods (e.g. AIC, BIC) or via χ^2 likelihood ratio tests comparing successively lower orders (e.g. ARX(2) versus ARX(1), ARX(1) versus ARX(0), etc.), with p-values corrected for multiple comparisons. The lowest order model, ARX(0), is simply a linear regression of ecosystem state against the exogenous variable with no intercept if the response data have been standardized prior to regression (the intercept is the mean, and standardization of the response makes the mean equal to 0). Both the exogenous predictor x and ecosystem state y should be standardized to N(0,1), especially if the objective is to compare resistance and recovery from disturbance among different sites or ecosystems.

Once the appropriate ARX(p) model has been identified, the next step is to derive the IRF. IRFs use the fitted ARX(p) parameters to model the trajectory of ecosystem state through time following either a single or repeated disturbance. In other words, IRFs use the fitted model to predict ecosystem state in whatever time series of the exogenous predictor the researcher chooses. For disturbance ecology, we are often interested in how a disturbance affects ecosystems immediately, and how ecosystems recover from disturbance. This could be expressed by a new exogenous time series that has a disturbance in the first step, and allows the exogenous predictor to recover to the mean value for all subsequent time steps (as is common with IRFs for other models, (Lutkepohl and Kratzig 2004, Bisgaard and Kulahci 2011, Box et al. 2015)). Such a time step could be expressed as:

$$x^* = \left[\alpha, 0, 0, 0, \ldots, 0\right]$$

where α denotes the disturbance intensity at the initial time point. It is critically important that the exogenous predictor x be standardized prior to model fitting, such that α is the stress unit in standard deviations (for example, if the predictor is rainfall, then $\alpha = -2$ is a 2 standard deviation reduction in rainfall) and the mean is equal to 0. In this way, ecologists can statistically subject disparate ecosystems to the same level of relative disturbance (e.g. a 2σ decline in precipitation) to estimate resistance using the same x^* series for each ecosystem and calculating the resulting IRF.

IRFs are essentially functions, or curves, that describe how the system changes through time for a given sequence of disturbances. For the single, initial disturbance x^* used here, IRFs contain all the information needed to quantify resistance, recovery, elasticity and return time (Fig. 3). Following the definitions of (Hodgson et al. 2015), resistance measures the instantaneous impact of disturbance when $x^* = \alpha$. Recovery is the extent to which ecosystem state remains suppressed (or elevated) once the exogenous predictor returns to average conditions ($x^* = 0$). Elasticity is the rate of return to average conditions, and return time is the amount of time required to achieve average conditions once the exogenous predictor returns to normal. For the single pulse x^* presented here, the IRF has simple analytical solutions for ARX(0) and ARX(1) models (Table 2). More complicated x^* functions, such as multiple pulses, or higher order ARX($p > 1$) models, can be solved computationally (Supporting information). The use of IRFs and x^* provides three benefits: First, observation error is removed via the use of a statistical model. Second, temporal stochasticity is eliminated via the use of x^*, wherein the exogenous predictor can stabilize at average values (or take any form the researcher desires). Third, by presenting both the disturbance and ecosystem response in units of standard deviations relative to normal conditions, disturbance magnitudes can be identical across ecosystems, and each of the four responses is directly comparable, facilitating cross-system comparisons.

Behavior of ARX models and IRFs

In comparison to the regression-based method of assessing significant lag effects (Fig. 2), ARX models have a very low probability of incorrectly identifying a lag effect when none are present. This can be illustrated with a simulation experiment. I estimated the parameters (means and covariances) for the ANPP–precipitation relationship of the shortgrass steppe in Colorado using linear regression. Then, I randomly drew intercept and slope parameters from a multivariate normal distribution of the parameter variance–covariance matrix. Using the simulated parameters, I calculated the time series of ANPP based solely on the precipitation pattern (no lag effects) and added noise to each estimate based on the residual error from the original regression. I then standardized both precipitation and ANPP prior to fitting three models: ARX(0), ARX(1) and ARX(2). The best model was chosen based on minimum BIC. This procedure was repeated 5000 times at various time series lengths. The type I error for lag effects was taken as the proportion of simulations in which ARX(1) or ARX(2) models were chosen to be the best model,

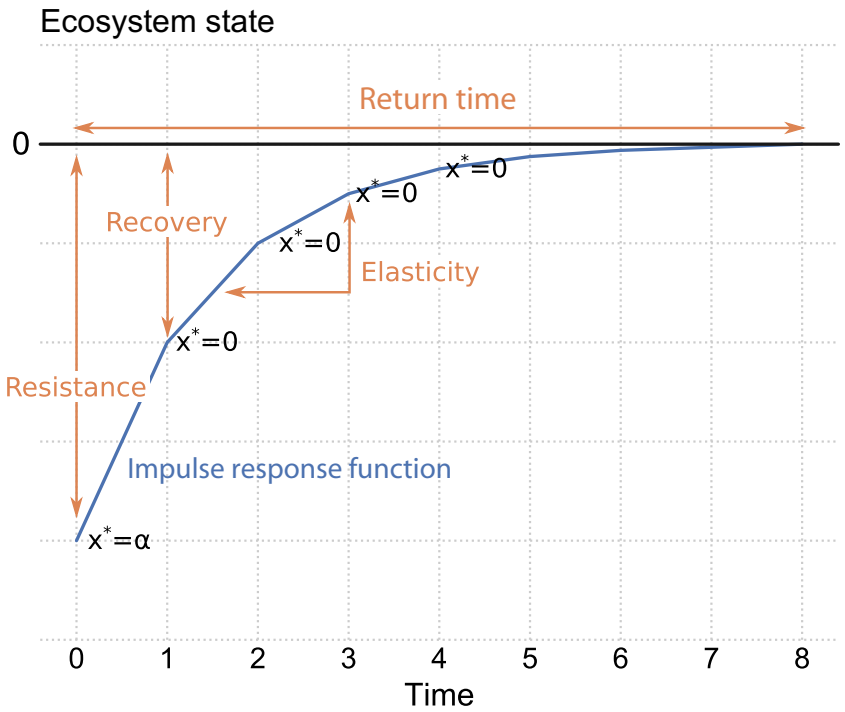

Figure 3. Graphical depiction of an IRF. In this IRF, the ecosystem is subject to an initial shock in the exogenous predictor with magnitude α. The system then returns to its average conditions for the rest of the time series. Resistance is measured by the decline in ecosystem state during the shock, with more negative values implying less resistance. Recovery is the extent to which an ecosystem remains altered post-disturbance, with more negative values implying less recovery. Elasticity is the rate at which the system recovers (i.e. slope, $\Delta y/\Delta x$), and return time is the amount of time it takes for a system to return to nominal levels. Note that this example implies a harmful disturbance. However, the sign of all values and the curve could flip for positive disturbances, such as a pulse of nitrogen enrichment on plant production.

thereby misidentifying lag effects. The type I error rate was high (approximately 0.25) for short time series, but quickly dropped to negligible levels for long time series (Fig. 4A). Even the high type I error rate for short time series was approximately half the type I error rate for the 95% CI approach (Fig. 2C), indicating that using information criteria to select ARX models provides a marked improved in accuracy.

To judge the accuracy of IRFs in estimating resistance and recovery, I conducted a second simulation experiment. For each of 10 000 simulations, I set $\alpha = -2$ to estimate resistance and recovery from a 2σ drought. I then drew a value of resistance from a Uniform(−2,0) distribution, and calculated the corresponding precipitation coefficient as $\beta = \text{Resistance}/\alpha$. I also drew the autoregressive parameter φ_1 from a Uniform(0,1) distribution and calculated $\text{Recovery} = \text{Resistance} \times \varphi_1$. I standardized the shortgrass steppe precipitation record, and simulated standardized ANPP data from an ARX(1) model based on the simulated β and φ_1 parameters. I back-converted standardized ANPP to the original scale assuming a mean of 100 g m^{-2} and a standard deviation of 15 g m^{-2}. I added noise to the data using an error rate of 15 g m^{-2}. I then re-standardized the noisy ANPP data and fit an ARX(1) model to the simulated data. From the ARX(1) model, I calculated the observed resistance and recovery based on IRFs for a 2σ drought ($\alpha = -2$). To further evaluate how ARX models improve estimates of resistance and recovery, for each simulation I also calculated resistance to the 2012 drought using the ratio-based method on the raw, unstandardized (noise-present) data.

ARX models and IRFs substantially improved the accuracy and precision of ecosystem responses to disturbance. Resistance to a 2σ drought was estimated accurately, the median points fall along the 1:1 line (Fig. 4B). Estimates of recovery were less accurate, and tended to be overestimated although the mismatch was slight (Fig. 4C). The degree of mismatch depends on the relative variation in the response and exogenous predictor; if the exogenous predictor is highly variable relative to the response, estimates of φ_1 and therefore recovery will be inaccurate (Supporting information). Both resistance and recovery were estimated imprecisely (Fig. 4B–C), but this is to be expected when estimating resistance to and recovery from an unknown disturbance. Estimates based on a known disturbance, such as the drought of 2012, were highly accurate and precise, especially compared to the common ratio-based method. As above, the ratio-based method based on raw data (including observation errors) were accurate but imprecise: a true resistance of 0.7 yielded estimates from below 0.4 to above 1.1 (Fig. 4D). In contrast, estimates of resistance to the 2012 drought based on noise-free values predicted by an ARX(1) model were highly precise; a true resistance of 0.7 yielded estimates of between 0.6 and 0.8 (Fig. 4D), a 2.5-fold increase in precision.

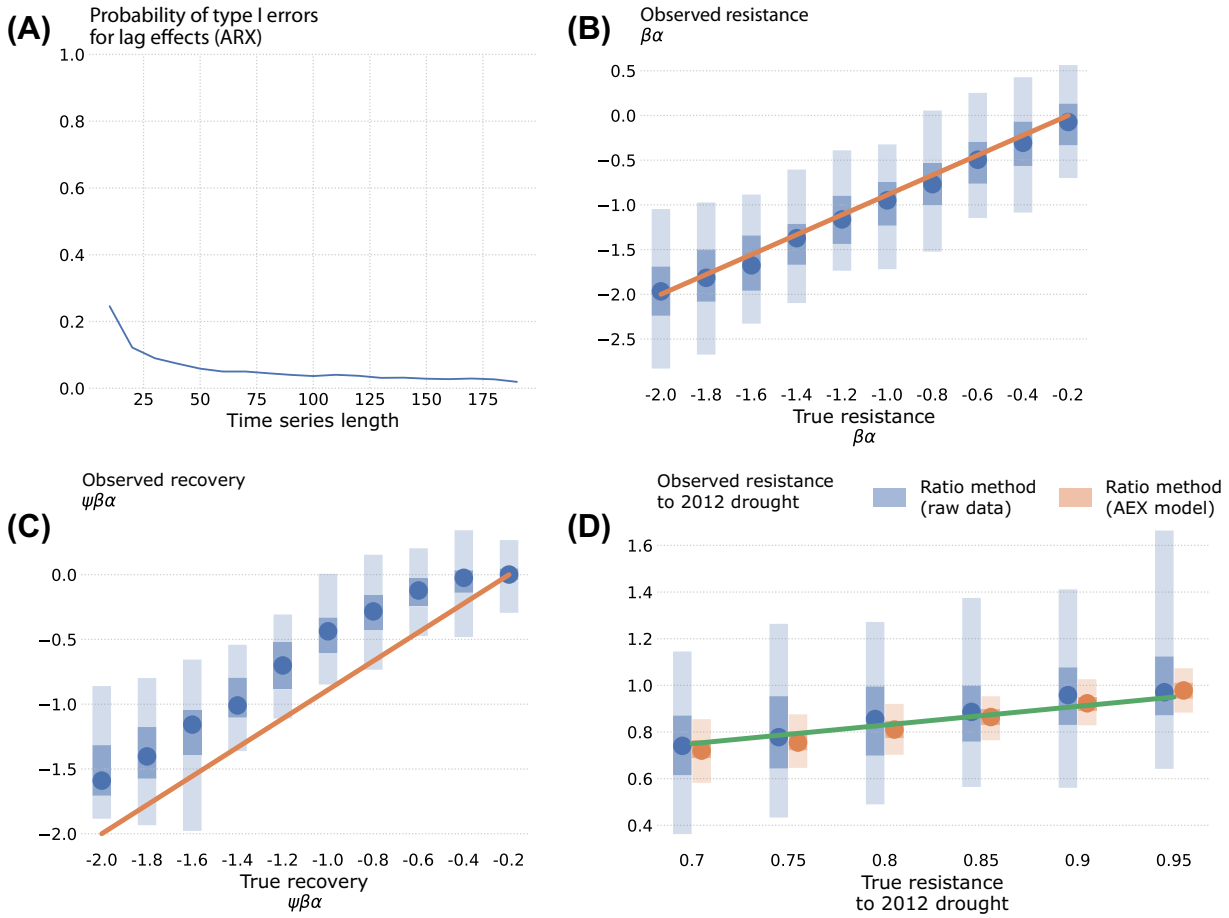

Figure 4. (A) Type I error rate for identifying lag effects using information criteria to choose the appropriate autoregressive order for ARX(*p*) models. The error rate was calculated as the proportion of simulations in which information theory incorrectly identified autoregressive structure to data simulated with no autoregressive structure. (B) Resistance to a 2σ drought using simulated ANPP data. A known resistance value was used to generate simulated ANPP data, which were then subject to IRF analysis to estimate resistance as βα. The orange line shows the 1:1 relationship, points show the medians, dark areas show the 50% quantile, and light areas show the 95% quantile. (C) Recovery from a 2σ drought using simulated ANPP data. A known recovery value was used to generate simulated ANPP data, which were then subject to IRF analysis to estimate recovery as φβα. The orange line shows the 1:1 relationship, points show the medians, dark areas show the 50% quantile, and light areas show the 95% quantile. (D) Comparison of resistance estimates based on the traditional ratio-based method applied to raw simulated data, or the noise-free fitted values based on an ARX(1) fit to simulated data. The green line shows the 1:1 relationship, points show the medians, dark areas show the 50% quantile, and light areas show the 95% quantile.

Case studies

The following section demonstrates the utility of ARX models and IRFs to different ecological questions. It is important to note that these case studies are intended as demonstrations only and do not represent thorough analyses of the actual phenomena in question. Thus, some data pre-treatment (i.e. gap-filling) presented here must be carefully considered in any actual model-fitting exercise. Further, in some examples, I make simplifying assumptions (i.e. lions prey only on wildebeest) in order to demonstrate the utility of IRFs. As with any statistical method, treatment of the data put into the model and careful consideration of the ecological process in question are critically important. Further, some example data are better suited to other methods, such as population models, but are used here as they illustrate the concept.

Crab larvae and sea surface temperature of the North Sea

Populations might, in some cases, exhibit lag effects by means of 'reproductive inertia'. For example, consider a population that has a large recruitment year due to favorable environmental conditions, such as warmer temperatures. If conditions return to nominal, the next year should still see elevated recruitment due to the higher-than-normal abundances of adults from the previous year. Such 'inertia' would manifest as lag effects, and stronger lag effects yield longer residual effects. However, the stochastic nature of both environments and populations makes understanding how long the effects of a single disturbance persist difficult. IRFs can illuminate just how long such effects persist by removing temporal stochasticity in the exogenous predictor.

As an illustration, I extracted data from a 47-year time series (1958–2005) of sea surface temperatures (SST) and decapod larval abundances in the North Sea compiled by Continuous Plankton Recorder Survey (Kirby et al. 2008) (Fig. 5A). Prior to analyses, I standardized and detrended both SST and larval abundances. Then, I fit the following three models:

$$\mathrm{ARX}(0) \quad \mathrm{DLA}_t = \beta \mathrm{STT}_t$$

$$\mathrm{ARX}(1) \quad \mathrm{DLA}_t = \beta \mathrm{STT}_t + \varphi_1 \mathrm{DLA}_{t-1}$$

$$\mathrm{ARX}(2) \quad \mathrm{DLA}_t = \beta \mathrm{STT}_t + \varphi_1 \mathrm{DLA}_{t-1} + \varphi_2 \mathrm{DLA}_{t-2}$$

where DLA is decapod larval abundance and SST is sea-surface temperatures. I chose the best model using BIC.

Decapod larval abundances were best described by an ARX(1) model (Table 3). Decapod larval abundances increased strongly with increasing SST ($\beta = 0.42 \pm 0.13$, $p = 0.001$). In addition, there was a strong, positive lag effect ($\varphi_1 = 0.74 \pm 0.01$, $p < 0.001$). This ARX(1) model allowed me to assess how decapod larval abundance responds to a disturbance of anomalously warm SST. In particular, I calculated the IRF for decapod larval abundance to a 2σ increase in SST in order to estimate the ecosystem responses to disturbance. Such effects are difficult to determine using raw time series due to stochastic fluctuations in SST that might mask autoregressive patterns. By using IRFs, I was able to simulate every year at average temperatures following a warm year.

Decapod larval abundances were not resistant to disturbance; a 2σ increase in SST yielded a large, 0.8σ increase in decapod larval abundances (Fig. 5B). However, strong autocorrelation ($\varphi_1 = 0.74$) indicated lag effects of decapod larval abundances. The lag effect resulted in low recovery, low elasticity and long return times. In the year following the warm disturbance, larval abundances recovered to 0.6σ above what would be expected based on temperature alone (Fig. 5B). The low elasticity suggested a low rate of recovery and produced a long return time, decapod abundances required almost 10 years to stabilize at pre-disturbance levels (Fig. 5B). Such autocorrelation likely results from population inertia, where an individual pulse of recruitment returns higher than average population sizes that diminish through time.

Wildebeest and lions of the Serengeti

Predator populations often exhibit lag effects in both prey and predator abundance. That is, predator abundances are sometimes determined by prey abundance in the prior year(s), which affects breeding success, fecundity and juvenile survival. Common examples include coyotes and lynx preying on snowshoe hares (O'Donoghue et al. 1997) and wolves and moose on Isle Royale National Park in the northern United States (McLaren and Peterson 1994). In other cases, predator populations respond instantaneously (i.e. in the same year) to changes in prey abundance (Samhouri et al. 2017). Lions and wildebeest of the Serengeti are an example of a predator–prey system that changes synchronously (Samhouri et al. 2017). As a result, lion population growth rates are sensitive to changes in wildebeest abundance in the same year. For example, wildebeest populations crashed in the mid-20th century due to an outbreak of rinderpest (Sinclair 1973), and lions populations subsequently declined precipitously due to a lack of wildebeest and other prey animals, such as buffalo. Herbivore populations increased quickly after the disease disappeared, followed by a marked increase in lion populations (Sinclair 1973). However, lions have low fecundity rates and require several years to reach sexual maturity. Populations might therefore be slow to increase despite higher food availability simply due to low initial population sizes (i.e. lag effects). In such cases, IRFs can be used to answer the question, 'How long does it take for a predator population to recover after a severe reduction in prey abundance?'

To demonstrate, I extracted a 33-year time series on wildebeest and lion populations from (Samhouri et al. 2017)

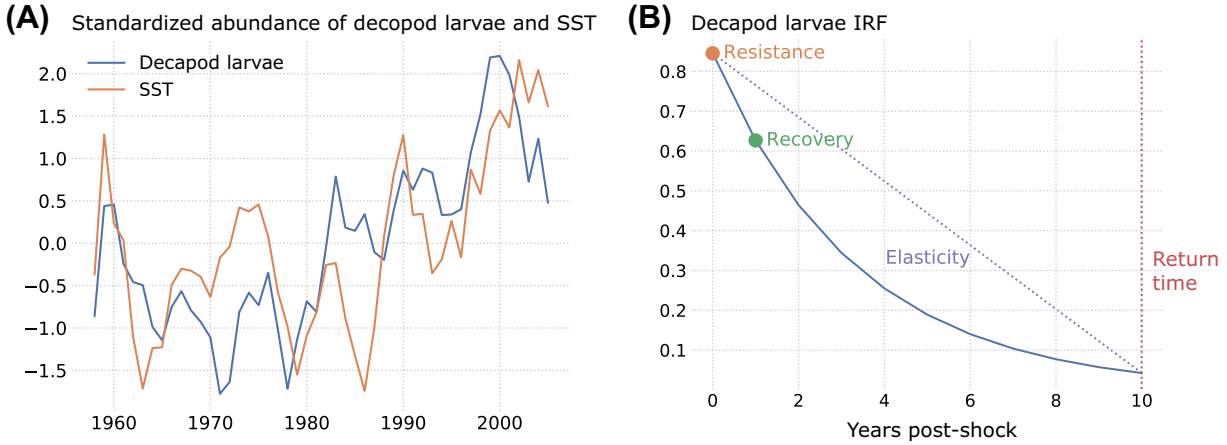

Figure 5. (A) Time series of standardized decapod larvae abundance and sea surface temperatures of the North Sea. (B) Impulse response function for decapod larvae following a 2σ increase in sea surface temperature in the first time point. Sea surface temperatures returned to average for the remaining time points.

Table 3. Bayesian information criteria (BIC) for three ARX(*p*) models fit to decapod larval abundances and SST time series. Bold denotes the best model, chosen by ΔBIC < 2.

Model	BIC	ΔBIC
ARX(0)	132.5	32.8
ARX(1)	**99.9**	**0.0**
ARX(2)	101.4	1.7

(Fig. 6A). Since ARX models require contiguous time series, I interpolated missing years using the average of the two nearest years. Prior to analysis, both time series were detrended by extracting residuals from linear regressions against year, and residuals were then standardized to N(0,1). I fit three models:

$$\mathrm{ARX}(0) \quad L_t = \beta W_t$$

$$\mathrm{ARX}(1) \quad L_t = \beta W_t + \varphi_1 L_{t-1}$$

$$\mathrm{ARX}(2) \quad L_t = \beta W_t + \varphi_1 L_{t-1} + \varphi_2 L_{t-2}$$

where L is lion abundance and W is wildebeest abundance. Note that these models are identical to autoregressive models of population abundance that include interspecific effects (Hansen et al. 1999), such that ARX models can be used to determine how populations respond to changes in prey, competitor or predator abundances. I chose the best model using BIC.

The best model describing lion abundance was an ARX(2) model that included current-year wildebeest abundance ($p = 0.001$), previous year lion abundance ($p < 0.001$) and two-year previous lion abundance ($p = 0.001$) (Table 4). With the model identified, I quantified how an extreme disturbance in wildebeest populations affects lion populations. Specifically, I calculated the lion IRF to an instantaneous 2σ reduction in wildebeest populations, followed by a stable increase to mean wildebeest population size for the remainder of the IRF. The IRF was calculated recursively (Supporting information). Since ARX(2) models do not have easily identified solutions for recovery time or elasticity due to oscillations (Supporting information), I chose recovery time as the point when the IRF stabilized around the equilibrium and elasticity as the rate at which lion populations first rebounded to average levels (Fig. 6B).

Lion populations were highly sensitive (i.e. not resistant) to reductions in wildebeest populations, declining by 1.1σ during the 2σ reduction in wildebeest (Fig. 6B). In the following year, strong lag effects prevented recovery and lion populations remained suppressed below 1σ (Fig. 6B). However, lion populations increased rapidly in subsequent years (elasticity = 0.4σ year^{-1}). Populations oscillated for several years, a consequence of the second autoregressive parameter and a common pattern in discrete-time predator–prey cycles. Recovery time was 14 years, at which point lion populations stabilized at average size (Fig. 6B). Thus, this example demonstrates that lion populations respond rapidly to a decline in wildebeest numbers, can remain suppressed for up to three years following a disease outbreak, and will fluctuate for nearly a decade before stabilizing.

Resistance and recovery of grasslands from extreme drought

Ecologists commonly use long-term time series to assess how grasslands vary in drought resistance and recovery across global or continental climate gradients (Sala et al. 2012, Knapp et al. 2015a). However, previous efforts have been hindered by the difficulties in standardizing drought effects across sites and accurately quantifying drought resistance free from temporal stochasticity (Table 1). Here, I demonstrate how IRFs can resolve this complexity by calculating ecosystem resistance to and recovery from drought for 14 globally-distributed herbaceous sites previously identified to possess significant lag effects.

Following Sala et al. 2012, I identified 14 datasets composed of both annual precipitation and ANPP in herbaceous

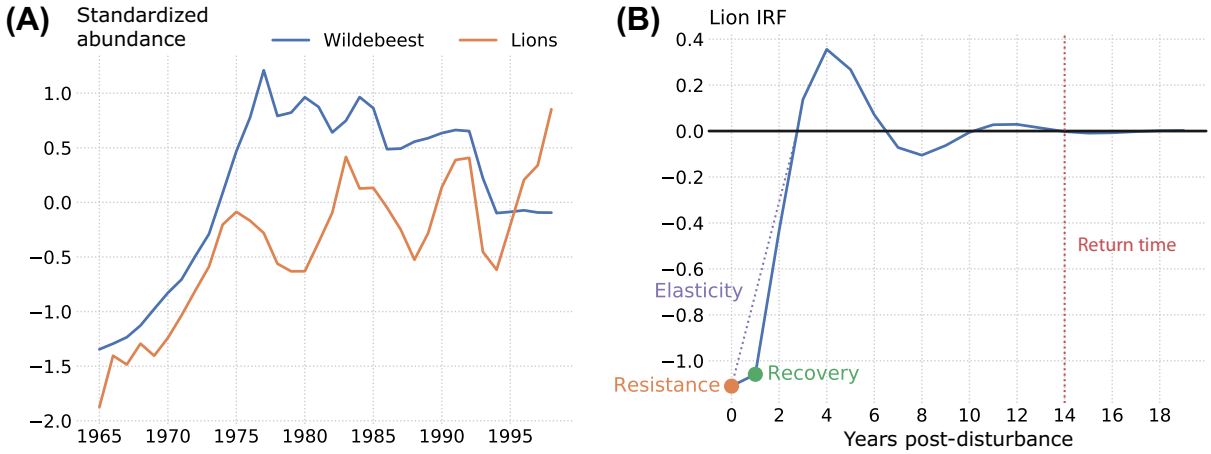

Figure 6. (A) Time series of standardized wildebeest and lion population sizes. (B) Impulse response function for lions following a 2σ decline in wildebeest abundances in the first time point. Wildebeest abundances returned to average for the remaining time points.

Table 4. Bayesian information criteria (BIC) for three ARX(*p*) models fit to lion and wildebeest time series. Bold denotes the best model, chosen by ΔBIC < 2.

Model	BIC	ΔBIC
ARX(0)	88.9	15.7
ARX(1)	78.4	5.2
ARX(2)	**73.2**	**0.0**

communities. Gap years in either ANPP or primary production were filled using a radial basis function. Radial basis functions (i.e. Gaussian processes) are commonly used to impute missing data in time series; however there are many imputation methods, including most common attribute, mean value, K-nearest neighbors, K-means clustering, expectation maximization, singular value decomposition or multivariate imputation by chained equations (Luengo et al. 2010, van Buuren and Groothuis-Oudshoorn 2011). I chose the simplest method because the emphasis here is on ARX examples, but ecologists should carefully consider the various methods of missing data imputation for real datasets. I kept only datasets with ten or more years. Prior to analyses, I standardized and detrended ANPP and precipitation for each dataset. I then fit the following models to each dataset:

$$\mathrm{ARX}(0) \quad \mathrm{ANPP}_t = \beta \mathrm{PPT}_t$$

$$\mathrm{ARX}(1) \quad \mathrm{ANPP}_t = \beta \mathrm{PPT}_t + \varphi_1 \mathrm{ANPP}_{t-1}$$

$$\mathrm{ARX}(2) \quad \mathrm{ANPP}_t = \beta \mathrm{PPT}_t + \varphi_1 \mathrm{ANPP}_{t-1} + \varphi_2 \mathrm{ANPP}_{t-2}$$

where ANPP is standardized, detrended ANPP and PPT is standardized, detrended precipitation. Following model fitting, I chose the best model for each dataset using BIC. After identifying the appropriate ARX model for each site, I calculated resistance and recovery following a 2σ decline in precipitation. In this way, ecosystem resistance and recovery were all calculated for the same magnitude of rainfall reduction relative to ambient conditions at each site. After calculating resistance and recovery of each site, I regressed each metric against mean annual precipitation derived from WorldClim.

Using the IRF method, primary production at the majority of sites (71%) was best described by an ARX(0) model, indicative of no lag effects (Table 5). Of the four sites exhibiting lag effects, three were best fit by an ARX(1) model and only one site was best fit by an ARX(2) model (Table 5). Resistance to a 2σ decline in precipitation varied among sites from a minimum of 2σ decline in ANPP at XLN to a 0.5σ increase in ANPP at NRB (Fig. 7A). Indeed, a significant positive relationship between drought resistance and mean annual precipitation ($p = 0.008$) indicated that drier herbaceous sites were generally less resistant to drought than mesic systems. Yet the relationship was not strong ($R^2 = 0.41$); even dry sites varied significantly in drought resistance. For example, JRN possesses roughly the same mean annual precipitation as XLN, yet JRN was 68% more resistant to a 2σ reduction in rainfall than XLN (Fig. 7A). Such variability in drought resistance among sites of similar precipitation has been previously reported (Huxman et al. 2004a) and could derive from differences in species composition, rainfall patterns (e.g. monsoonal, Mediterranean, etc.) or management history among sites. Relatively few grasslands demonstrated lag effects, such that most sites exhibited perfect recovery in the year following drought (Fig. 7B). There was no relationship between mean annual precipitation and the strength of recovery (Fig. 7B).

Table 5. ΔBIC values of ARX(*p*) models for 14 grassland sites used in (Sala et al. 2012). Bold denotes the best model, chosen by ΔBIC < 2. In the case of multiple competing models, I chose the simplest model following the principle of parsimony.

Site	ARX(0)	ARX(1)	ARX(2)
Badkhyz, Turkmenistan (BDK)	**0.0**	2.8	6.3
Cheyenne, Wyoming (CHY)	**0.0**	2.4	4.3
Dzhanybek, Kazakhstan (DZH)	**0.0**	3.0	6.5
Jornada, New Mexico (JRN)	**0.0**	3.2	4.8
Konza Prairie, Kansas (KNZ)	**0.0**	3.3	6.7
Kursk, Russia (KRS)	**0.0**	1.3	3.5
Manyberries, Alberta (MBR)	4.1	**0.0**	2.1
Nairobi, Kenya (NRB)	**0.0**	0.8	2.6
Niwot Ridge, Colorado (NWT)	14.0	**0.0**	0.4
Rio Mayo, Argentina (RMY)	**0.1**	2.1	0.0
Sevilleta, New Mexico (SEV)	**0.0**	0.9	1.3
Fort Collins, Colorado (SGS)	**0.0**	0.0	2.7
Tumugi, China (TMG)	1.6	**0.0**	2.0
Xilingol, China (XLN)	1.90	1.80	**0.0**

Conclusions

Given the expected increase in both the severity and intensity of extreme disturbance events, it is imperative that we accurately quantify how ecosystems respond to disturbance. Estimating ecosystem vulnerability to disturbance using long-term time series data is a promising approach, but ecologists have not yet coupled time series data with the appropriate statistical time-series tools. Most current methods possess flaws that potentially bias estimates of ecosystem susceptibility to disturbance and potentially misidentify legacy effects. To resolve these issues, I advocate for using IRFs derived from autoregressive time series models as a single quantitative framework that can accurately estimate ecosystem resistance, recovery, elasticity and return time from disturbance events. Impulse response functions have numerous advantages over prior techniques, including the separation of observation and process errors, standardizing disturbance severity among different locations, and rigorously testing for legacy effects.

One advantage of the method proposed here is the relative ease with which ARX models can be fit and IRFs calculated in common statistical programming languages. The following recommendations would prove beneficial for ecologists wishing to implement the method outlined here:

1. *Properly pre-treat data* – Data must be processed properly prior to analysis with autoregressive models. First, data must be examined for gaps, as simple ARX models

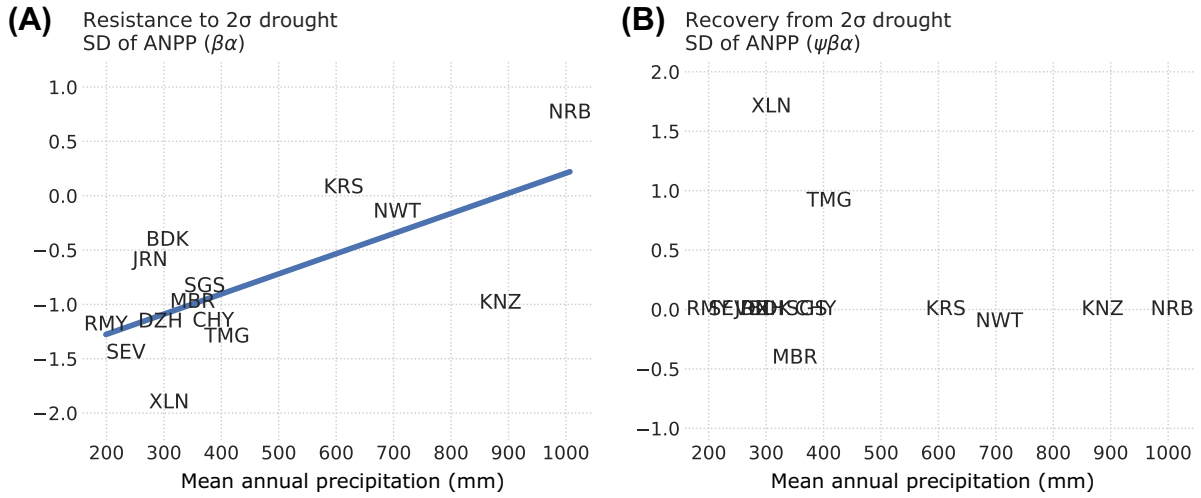

Figure 7. (A) Relationship between resistance to a 2σ drought, calculated from impulse response functions and mean annual precipitation for 14 herbaceous systems. Line shows the best fit by least squares. (B) Relationship between recovery from a 2σ drought, calculated from impulse response functions and mean annual precipitation for 14 herbaceous systems.

proposed here do not function with non-contiguous data. Small gaps can be filled with a data imputation function (e.g. radial basis functions, used here). Second, data must be detrended. ARX models assume stationarity, wherein the mean and variance do not change through time. Detrending data by using the residuals from a linear regression against time can stabilize the mean through time, but variances must still be checked visually. Third, data should be standardized in order to facilitate comparison among sites by subtracting the mean and dividing by the standard deviation. For example, if precipitation is not standardized, then $\alpha = -2$ for the IRF is only a 2 mm decline in rainfall, rather than an extreme 2σ event. In some cases, it might be preferable not to standardize. In the crab larvae example, not standardizing SST and using $\alpha = 2$ tests the system's response to a 2°C increase in temperature, relevant for some climate scenarios. However, in most cases, standardizing the data will allow researchers to compare IRFs among disparate systems, locations or stress responses. The decision whether or not to standardize data should be made carefully depending on the question asked.

2. *Use a 2σ increase or decrease in the exogenous predictor* – If all ecologists use a 2σ change in the exogenous predictor, then results are perfectly comparable among studies. I chose 2σ because it represents an extreme event. For example, a 1σ decline in rainfall is the 16% quantile, whereas a 2σ decline in rainfall represents a drought falling in the 2% quantile (assuming a normal distribution), thereby representing an extreme stress event. However, ecologists are also free to use whatever α they feel appropriate for the question at hand, so long as it is reported and justified.
3. *Report the autoregressive order and parameter values* – Reporting the parameters enables future researchers to easily extract the IRF and calculate ecosystem disturbance responses under different x^*. For example, ecologists could standardize all IRFs to a 2σ stress if variation exists in the literature, or could assess ecosystem recovery using values different from a 50% return in ecosystem function. Alternatively, future researchers could use IRFs to assess how systems respond to multiple disturbance events of either identical or varying magnitude.
4. *Use designated AR model fitting functions* – The ARX functions specified here could all be fit using least squares. Doing so, however, requires trimming the first two data points from all model fits because we cannot use information theory or likelihood ratio tests to compare models fit to different datasets (e.g. n points for ARX(0), n − 1 points for ARX(1), n − 2 points for ARX(2), etc.). For small datasets, the loss of two data points can substantially alter the results. For example, using OLS to fit an ARX(0) model to the RMY data without the first two data points ($n = 8$) results in no relationship between primary production and precipitation ($p = 0.65$) because the first two data points are the driest and wettest years. Using the full dataset ($n = 10$) yields a stronger primary production–precipitation relationship ($p = 0.12$). Common statistical languages have ARIMAX functions (R: TSA library, Python: statsmodels module) wherein the user can specify the AR order, incorporate an exogenous predictor, and utilize the full dataset.

Using time series of adequate length is perhaps the most important consideration for ARX models and subsequent IRFs. As described above, ARX models are relatively straightforward linear models, and as such the same recommendations of sample size for linear models apply to ARX models. It is generally recommended that a dataset contain 10 observations per parameter being estimated. Thus, ARX(2) models should ideally be fit to time series with at least 30 contiguous points. However, I recognize that most time series are considerably shorter. It is possible to use short time series in ARX(2)

models; I advise no fewer than 10 points. This provides data points per parameter (plus one). Furthermore, longer time series have a greater probability of including an extreme event and thus will model IRFs more accurately.

A second important consideration is the relationship between the exogenous predictor and the response. The exogenous predictor should be linearly related to the ecosystem response, otherwise ARX models and IRFs will yield nonsensical results. This is best illustrated with an example. An outbreak of crown-of-thorns starfish occurred on the coral reefs around Moorea in the mid-2000s, devastating coral cover. Corals slowly recovered after starfish returned to normal densities. I attempted to model this with an ARX model and IRF, but the effect of starfish on coral cover was non-significant or even positive! The model correctly identified the slow recovery time of corals, but gave the impression that starfish had minimal or even positive effects on coral cover. The reason for this behavior was the dichotomous nature of starfish data: starfish were either hyperabundant or nearly absent, and the relationship between starfish abundance and coral cover was therefore nonlinear and exhibited a threshold response. ARX models will probably struggle with hurricanes, forest fires or other stochastic, 'all-or-nothing' disturbances. The ARX model works best for assessing disturbance from an exogenous predictor that exhibits a range of values, such as precipitation, temperature, salinity, etc.

In conclusion, IRFs provide ecologists with a quick and simple means for quantifying ecosystem responses to disturbance, while enabling ecologists to capitalize on the increased availability of long-term, observational time series data. Ecologists can use this method to quantify the components of ecosystem disturbance response in a standardized way across many sites. Site-specific information on species composition, long-term climate, rainfall patterns or any other important variable can then be used to identify the abiotic and biotic factors that dictate ecosystem stress response. For example, the brief analyses presented here suggest that dry grasslands are often more sensitive to drought than wet grasslands, but also that our understanding of differential ecosystem sensitivity to drought remains incomplete. ARX models can even be used when there are multiple exogenous drivers, such as temperature and precipitation, by including an additional predictor in the model. As a result, IRFs should greatly improve our ability to predict how ecosystems will respond to the increased severity and frequency of extreme disturbance events in the future.

Data availability statement

There are no new data in this manuscript. All data used here are freely available online. The Supplementary information contain all code for reproducing analyses and figures.

Acknowledgements – I would like to thank the newly minted Drs. R. Griffin-Nolan and A. Hoffman for their helpful comments on drafts of this manuscript. This work was improved through the helpful comments of P. R. Peres-Neto.

Funding – This work was funded by an NSF DEB award (1941390) to NPL.

Author contributions

Nathan Lemoine: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (lead).

References

Anderegg, W. R. L. et al. 2015. Pervasive drought legacies in forest ecosystems and their implications for carbon cycle models. – Science 349: 528–532.

Bisgaard, S. and Kulahci, M. 2011. Time series analysis and forecasting by example. – Wiley.

Box, G. E. P. et al. 2015. Time series analysis: forecasting and control. – Wiley.

Chang, R. W. et al. 2017. Determining arterial wave transit time from a single aortic pressure pulse in rats: vascular impulse response analysis. – Sci. Rep. 7: 1–9.

Ciais, P. et al. 2005. Europe-wide reduction in primary productivity caused by the heat and drought in 2003. – Nature 437: 529–533.

Creal, D. D. and Wu, J. C. 2017. Monetary policy uncertainty and economic fluctuations. – Int. Econ. Rev. 58: 1317–1354.

Gambetti, L. and Musso, A. 2017. Loan supply shocks and the business cycle. – J. Appl. Econ. 32: 764–782.

Gazol, A. et al. 2017. Impacts of droughts on the growth resilience of Northern Hemisphere forests. – Global Ecol. Biogeogr. 26: 166–176.

Gazol, A. et al. 2018. Forest resilience to drought varies across biomes. – Global Change Biol. 24: 2143–2158.

Griffin-Nolan, R. J. et al. 2018. Legacy effects of a regional drought on aboveground net primary production in six central US grasslands. – Plant Ecol. 219: 505–515.

Hansen, T. F. et al. 1999. Interspecific and intraspecific competition as causes of direct and delayed density dependence in a fluctuating vole population. – Proc. Natl Acad. Sci. USA 96: 986–991.

Heisler-White, J. L. et al. 2008. Increasing precipitation event size increases aboveground net primary productivity in a semi-arid grassland. – Oecologia 158: 129–140.

Hodgson, D. J. et al. 2015. What do you mean, 'resilient'? – Trends Ecol. Evol. 30: 503–506.

Huxman, T. E. et al. 2004a. Convergence across biomes to a common rain-use efficiency. – Nature 429: 651–654.

Huxman, T. E. et al. 2004b. Precipitation pulses and carbon fluxes in semiarid and arid ecosystems. – Oecologia 141: 254–268.

Joos, F. et al. 2013. Carbon dioxide and climate impulse response functions for the computation of greenhouse gas metrics: a multi-model analysis. – Atmos. Chem. Phys. 13: 2793–2825.

Kannenberg, S. A. et al. 2020. Ghosts of the past : how drought legacy effects shape forest functioning and carbon cycling. – Ecol. Lett. 23: 891–901.

Kirby, R. R. et al. 2008. Climate-induced effects on the meroplankton and the benthic–pelagic ecology of the North Sea. – Limnol. Oceanogr. 53: 1805–1815.

Knapp, A. K. et al. 2015a. Characterizing differences in precipitation regimes of extreme wet and dry years: implications for climate change experiments. – Global Change Biol. 21: 2624–2633.

Knapp, A. K. et al. 2015b. Differential sensitivity to regional-scale drought in six central US grasslands. – Oecologia 177: 949–957.
Le Nohaïc, M. et al. 2017. Marine heatwave causes unprecedented regional mass bleaching of thermally resistant corals in northwestern Australia. – Sci. Rep. 7: 14999.
Lirman, D. et al. 2011. Severe 2010 cold-water event caused unprecedented mortality to corals of the Florida reef tract and reversed previous survivorship patterns. – PLoS One 6: e23047.
Lloret, F. et al. 2011. Components of tree resilience: effects of successive low-growth episodes in old ponderosa pine forests. – Oikos 120: 1909–1920.
Lodge, D. J. and McDowell, W. H. 1991. Summary of ecosystem-level effects of Caribbean hurricanes. – Biotropica 23: 373–378.
Luengo, J. et al. 2010. A study on the use of imputation methods for experimentation with radial basis function network classifiers handling missing attribute values: the good synergy between RBFNs and EventCovering method. – Neural Netw. 23: 406–418.
Lutkepohl, H. and Kratzig, M. 2004. Applied time series econometrics. – Cambridge Univ. Press.
McLaren, B. and Peterson, R. O. 1994. Wolves, moose and tree rings on Isle Royale. – Science 266: 1555–1558.
Millar, J. R. et al. 2017. A modified impulse-response representation of the global near-surface air temperature and atmospheric concentration response to carbon dioxide emissions. – Atmos. Chem. Phys. 17: 7213–7228.
Nimmo, D. G. et al. 2015. Vive la résistance: reviving resistance for 21st century conservation. – Trends Ecol. Evol. 30: 516–523.
O'Donoghue, M. et al. 1997. Numerical responses of coyotes and lynx to the snowshoe hare cycle. – Oikos 80: 150–162.
Oesterheld, M. et al. 2001. Inter-annual variation in primary production of a semi-arid grassland related to previous-year production. – J. Veg. Sci. 12: 137–142.
Perkins, S. E. et al. 2012. Increasing frequency, intensity and duration of observed global heatwaves and warm spells. – Geophys. Res. Lett. 39: L20714.
Sala, O. E. et al. 2012. Legacies of precipitation fluctuations on primary production: theory and data synthesis. – Phil. Trans. R. Soc. B 367: 3135–3144.
Samhouri, J. F. et al. 2017. Rapid and direct recoveries of predators and prey through synchronized ecosystem management. – Nat. Ecol. Evol. 1: 1–6.
Schultz, W. et al. 2015. Phasic dopamine signals: from subjective reward value to formal economic utility. – Curr. Opin. Behav. Sci. 5: 147–154.
Senbet, D. 2016. Measuring the channels of monetary policy transmission: a factor-augmented vector autoregressive (FAVAR) approach. – J. Cent. Bank. Theory Pract. 5: 5–40.
Sinclair, A. R. E. 1973. Population increases of buffalo and wildebeest in the Serengeti. – Afr. J. Ecol. 11: 93–107.
Smale, D. A. et al. 2019. Marine heatwaves threaten global biodiversity and the provision of ecosystem services. – Nat. Clim. Change 9: 306–312.
Smith, M. D. 2011. The ecological role of climate extremes: current understanding and future prospects. – J. Ecol. 99: 651–655.
Stuart-Haëntjens, E. et al. 2018. Mean annual precipitation predicts primary production resistance and resilience to extreme drought. – Sci. Total Environ. 636: 360–366.
Thompson, M. V and Randerson, J. T. 1999. Impulse response functions of terrestrial carbon cycle models: methods and application. – Global Change Biol. 5: 371–394.
van Buuren, S. and Groothuis-Oudshoorn, K. 2011. MICE: multivariate imputation by chained equations. – J. Stat. Softw. 45: 1–65.
Wilcox, K. R. et al. 2017. Asymmetric responses of primary productivity to precipitation extremes: a synthesis of grassland precipitation manipulation experiments. – Global Change Biol. 23: 4376–4385.
Zeng, S. et al. 2017. The response of the Beijing carbon emissions allowance price (BJC) to macroeconomic and energy price indices. – Energy Policy 106: 111–121.