esa

Ecosphere

SYNTHESIS & INTEGRATION

Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems

Michael A. Litzow[1,]† and Mary E. Hunsicker[2]

[1]*Farallon Institute for Advanced Ecosystem Research, Petaluma, California 94952 USA*
[2]*Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service, National Oceanic and Atmospheric Administration, Newport, Oregon 97365 USA*

Citation: Litzow, M. A., and M. E. Hunsicker. 2016. Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. Ecosphere 7(12):e01614. 10.1002/ecs2.1614

Abstract. Early warning signals (EWS) might dramatically improve our ability to manage nonlinear ecological change. However, the degree to which theoretical EWS predictions are supported in empirical systems remains unclear. The goal of this study is to make recommendations for identifying the types of ecological transitions that are expected to show EWS. We conducted a review and meta-analysis of published studies and comparative analysis of eight northeast Pacific Ocean time series to illustrate the importance of testing for nonlinearity in empirical EWS studies. We found that published studies demonstrating nonlinearity in ecosystem dynamics are more likely to support EWS predictions than studies with linear or undetermined dynamics. The northeast Pacific time series in our analysis were often too short for formal tests of nonlinearity, a common problem in empirical studies. To assess the evidence for nonlinear dynamics in these data, we tested for state-dependent driver–response relationships consistent with hysteresis, a central feature of nonlinear ecological models. This analysis supported the results of the literature meta-analysis. Four time series with driver–response relationships consistent with hysteresis generally supported theoretical EWS predictions, while four without evidence of hysteresis failed to support EWS predictions. Theoretical support for EWS is largely generated from nonlinear models, and we conclude that tests for either nonlinear dynamics or hysteresis are needed before employing EWS.

Key words: alternate states; early warning; hysteresis; leading indicator; nonlinearity; North Pacific; regime shift.

Received 7 June 2016; revised 18 October 2016; accepted 19 October 2016. Corresponding Editor: Julia A. Jones.

† **E-mail:** litzow@faralloninstitute.org

Introduction

Early warning signals (EWS) for abrupt change are one of the fastest-growing subfields in ecology. The general idea of EWS is that key system parameters should show characteristic statistical signals as the system approaches a transition, and that tracking these signals should provide information about the likelihood of abrupt change. Two classes of EWS are available: model-based and metric-based (Boettiger and Hastings 2012, Dakos et al. 2012*a*). Metric-based EWS, such as rising variance or rising autocorrelation in key parameters prior to a transition, require almost no assumptions about the correct model underlying system dynamics (Carpenter and Brock 2006, Scheffer et al. 2009), and this approach has dominated the empirical literature. The theoretical justification for EWS is most often based on critical slowing down as the system loses resilience and approaches a critical transition (Box 1; Scheffer et al. 2009). EWS arise from a long tradition of

Box 1 - *Definitions*

The terms describing nonlinear theory and early warning signals often have meanings that are highly specific and context-dependent. Here, we define some of these concepts as we use them in the paper.

Alternative stable states: Different configurations of a system that are able to exist at the same set of external conditions, corresponding to a stable equilibrium or basin of attraction in nonlinear response to external conditions.

Critical slowing down: Reduced speed of recovery from perturbation as a critical transition is approached, due to a decline in engineering resilience.

Critical transition: Abrupt shift in a system caused by nonlinear responses to external conditions.

Early warning signal/early warning indicator: Model- or metric-based statistic able to warn that the system is approaching a sudden change, most often associated with a critical transition.

Fold bifurcation/saddle-node bifurcation: A critical transition between alternative stable states, corresponding to the threshold in external conditions at which stable and unstable equilibria meet.

Hysteresis: Different critical transitions in response to increasing and decreasing external conditions; responses to external conditions that depend on system state and the direction of change in external conditions.

Linear: Systems with dynamics that can be expressed statistically by models in which the estimated parameters are combined by addition. Thus, a linear regression ($y = a + bx$) or a quadratic regression ($y = a + b_1x + b_2x^2$) both describe linear systems. In a linear system, the effect of any small perturbation decays in time.

Nonlinear: In a nonlinear system, a small perturbation may propel the system to another stable state, and dynamics are both state-dependent and sensitive to initial conditions. Statistically, the response variable cannot be summarized as a linear combination of estimated parameters.

Resilience: Ecological resilience is the ability of a system to remain in its current state when exposed to perturbation. Engineering resilience is the speed with which a system returns to equilibrium after perturbation.

using nonlinear models to explain complex ecological behavior. They represent the most concrete contribution of this tradition to problems of ecosystem management to date, in the form of explicit, testable predictions concerning the response of systems to external perturbation, as well as an approach for quantifying resilience, which has previously been more of a buzzword than a measurable quantity. The field of EWS presents great potential for addressing currently intractable problems in ecosystem management, such as the early recognition of impending transitions, the avoidance of ecological surprises, and the maintenance of systems in desired states.

But attempts to demonstrate the application of EWS theory in empirical systems have produced uneven results. Some studies have found EWS to fail completely in real systems (Bestelmeyer et al. 2011, Burthe et al. 2016), others have found some mix of success and failure (Lindegren et al. 2012, Litzow et al. 2013), and others have presented evidence supporting the predictions of theory (Dakos et al. 2008, Wouters et al. 2015). This mix of results is unsurprising when predictions derived from simple models are applied to the complex real world (Scheffer et al. 2015). However, progress in the use of EWS in empirical systems would benefit from attempts to elucidate the factors distinguishing successful and unsuccessful real-world applications.

A critical question in the application of EWS is whether changes in the system under consideration are the result of linear or nonlinear dynamics. The fold or saddle-node bifurcation model that underpins much of the theoretical EWS literature invokes strongly nonlinear dynamics, with multiple ecological responses possible at a single level of external forcing (Beisner et al. 2003). While there is theoretical evidence for EWS prior to ecological changes stemming from other dynamics (Kéfi et al. 2013), theoretical support for EWS is generally weaker in situations where biological parameters show a linear or threshold response to perturbation (Dakos et al. 2015). Empirical studies often assume fold bifurcation dynamics (Boettiger and Hastings 2012) or make heuristic arguments for the presence of "regime shifts" with similar dynamics (Hewitt and Thrush 2010, Litzow et al. 2013, Wouters et al. 2015), while other empirical studies present formal tests for nonlinearity (Carpenter et al. 2011, Wang et al. 2012). Given the central importance of nonlinearity in EWS theory, and the range of approaches taken by empirical studies, perspective on the best approach is needed.

Here, we elucidate the importance of tests for nonlinear dynamics in attempts to apply EWS theory to empirical systems, using a review of the literature, meta-analysis of published studies, and comparative analysis across multiple data sets. Our specific goals are to (1) review the state of empirical EWS research to date; (2) use meta-analysis to compare the results of published studies that do and do not demonstrate nonlinearity in study systems; and (3) analyze eight northeast Pacific Ocean time series to compare EWS results between systems that are and are not consistent with a model of hysteretic driver–response relationships, a central feature of nonlinear ecological systems.

Literature Review

Empirical EWS research, 2006–2015

Like most ideas in ecology, EWS have antecedents that stretch far back in the literature. Previous approaches for measuring ecological resilience and stability (Ives 1995, Ives et al. 2003) draw on foundational research on nonlinear ecological dynamics (May 1977, Wissel 1984) and the relationships between disturbance, nonlinearity, and statistical behaviors of ecosystems (van Nes and Scheffer 2003, Fraterrigo and Rusak 2008). However, the idea that characteristic statistical behaviors prior to critical transitions are a tool that can be used broadly in systems showing complex dynamics was advanced by Carpenter and Brock (2006), and the explosion in the EWS literature can reasonably be dated to that paper. Our review and meta-analysis of published studies covers papers appearing in print or online during 2006–2015. At least 94 EWS studies were published in ecology and climate science during that time. And while this literature has its origin in theory, nearly a third of the papers published during those years included empirical tests of proposed EWS behavior (Fig. 1A). Given this number of examples available in the literature, reasonable conclusions from cross-study comparisons may be possible. A detailed description of empirical studies is presented in Appendix S1: Table S1, and a complete list of published EWS studies appears in Appendix S1: Table S2.

Early warning signals have been described as "generic"—suitable for application across many system types, even if the underlying system dynamics are poorly understood (Carpenter and Brock 2006, Scheffer et al. 2009). This makes them extremely attractive for application to large, complex ecosystems, such as coastal and open ocean systems, where the mechanisms underlying complex reactions to external perturbation are almost always mysterious. However, this idea of generic application can be overstretched. It is relevant to the mechanisms of system behavior, but not to the class of model that describes system dynamics. If the system is best described by a model of linear relationships between the biological response and environmental drivers, there is little expectation for the presence of EWS. A central question, both in theoretical and in empirical EWS studies, is therefore the identity of the model describing system response to perturbation (Boettiger et al. 2013). Theorists are involved in a dialogue about which models may or may not produce EWS (e.g., Carpenter and Brock 2006, van Nes and Scheffer 2007, Hastings and Wysham 2010, Dakos et al. 2011, Seekell et al. 2011, Boettiger and Hastings 2012, Boerlijst et al. 2013, D'Odorico et al. 2013, Fung et al. 2013, Guttal et al. 2013, Kéfi et al. 2013, Dakos and Bascompte 2014, Clements et al. 2015, D'Souza et al. 2015, Lumi et al. 2015, Xu et al. 2015). Empiricists are faced with the double problem of sorting out the conflicting advice of the theoretical literature and then determining which theoretical model is appropriate for their system. This is a poorly tractable problem that predates the development of EWS. It brings together several threads of uncertainty in the literature, which may give rise to the mixed results in the application of EWS theory in empirical systems.

In the empirical studies to date, the ecosystem transitions for which EWS are being tested

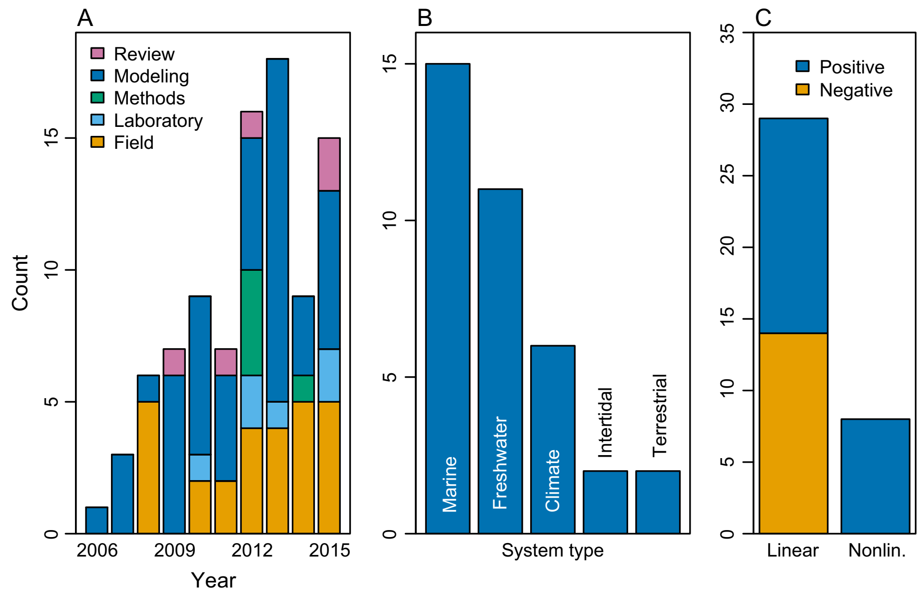

Fig. 1. Summary of early warning signals (EWS) studies published in print or online during 2006–2015. (A) Published papers classified by year and study type. (B) Empirical studies classified by system type (more than one system in some studies). (C) The results of quantitative EWS tests (positive or negative results) compared between study systems with linear or undefined dynamics vs. systems with nonlinear dynamics.

include trophic cascades, desertification, and shifts in species abundance, community composition, and climate patterns (Appendix S1: Table S1). Empirical examples come primarily from marine and freshwater ecology and climate studies (Fig. 1B). These empirical studies have faced two general challenges: the challenge of identifying and classifying ecosystem shifts that are suitable for EWS, and the statistical challenges that are presented by testing EWS with real data.

Sudden ecosystem transitions that might be preceded by EWS are often referred to as "regime shifts," but this term has never achieved a universal definition in the literature (Lees et al. 2006). The term first gained currency in marine ecology, where regime shifts may be linear biological reactions to abrupt change in an environmental driver (Rudnick and Davis 2003, Di Lorenzo and Ohman 2013). In the theoretical and general ecology literature, regime shifts have become synonymous with critical transitions among alternative states (Scheffer et al. 2001, Scheffer and Carpenter 2003, Guttal and Jayaprakash 2008, Pace et al. 2015). The difficulty in defining the term has meant that "regime shift" has also been used to mean any large, persistent change in a system, springing from a diverse set of underlying dynamics (Brock and Carpenter 2010). Given this vagueness in the literature, it is unsurprising that empirical studies have taken a variety of approaches to the problem of identifying ecological transitions that might produce EWS. Some studies rigorously test shifts for conformity with a model of nonlinear change between alternative states (e.g., Carpenter et al. 2011, Wang et al. 2012). More commonly, studies use simple temporal or spatial breakpoint analyses as evidence of a less rigorously defined shift of indeterminate nature (e.g., Beaugrand et al. 2008, Litzow et al. 2008, 2013, Carstensen and Weydmann 2012, Wouters et al. 2015, Burthe et al. 2016). Some of these studies explicitly acknowledge that tests for critical transitions are impossible given the data on hand, and recognize the use of breakpoint analysis as a necessary compromise (e.g., Carstensen and Weydmann 2012). Still other studies, without directly testing for nonlinearity, reference features of the study system that are heuristically consistent with critical transitions between alternative states (e.g., Litzow et al. 2008, 2013, Hewitt and Thrush 2010, Wouters et al. 2015). This reliance on heuristic arguments risks repeating earlier problems in the literature concerning the possible over-application of nonlinear theory to describe any abrupt ecological change (see *Hysteresis as a hallmark of nonlinearity*, below). Finally, some studies have tested for EWS against shifts detected in time series without any attention to the question of whether those shifts were the result of alternative states or other nonlinear dynamics. This type of application may involve confusion over the term "nonlinear" in the literature. For instance, Burthe et al. (2016) defined nonlinear change as turning points in predicted values from generalized additive models (GAMs) fit to time series, but this sort of time series variability may be parsimoniously explained by linear processes, such as red noise (Di Lorenzo and Ohman 2013). Similar confusion may exist in the literature on possible driver–response relationships (*Analysis of northeast Pacific time series*, below), with relationships that can be summarized through polynomial regression misclassified as nonlinear.

Empirical studies are also faced with the statistical problems posed by time series data that are short, noisy, temporally autocorrelated, open to confounding and unknown sources of variability, and inadequately replicated. Very helpful guides for statistical best practices in empirical studies have been published (e.g., Boettiger and Hastings 2012, Dakos et al. 2012*a*, Kéfi et al. 2014). However, in many cases, proposed statistical approaches are demonstrated with data produced by models, and may therefore not consider the real-world problems outlined above. For instance, metric-based indicators calculated from sliding windows within time series were first proposed in modeling studies where time series length is not a consideration, but this approach may offer inadequate statistical power to detect a trend in EWS within the confines of short real-world time series (Boettiger and Hastings 2012). Empirical studies to date have taken a variety of statistical approaches that may be classed into three groups: statistical hypothesis testing (e.g., Carstensen and Weydmann 2012, Litzow et al. 2013, Wouters et al. 2015, Burthe et al. 2016), a model selection approach (e.g., Carstensen and Weydmann 2012, Krkošek and Drake 2014), and shift detection routines applied simultaneously to EWS time series and time series measuring system state (e.g., Litzow et al. 2008, Lindegren et al. 2012). Generally speaking, these three approaches can be considered as a declining order of statistical rigor for demonstrating the operation of EWS. However, the magnitude of the problem of designing appropriate statistical tests for EWS within the constraints imposed by real-world data sets can be seen in the relatively large number of studies that have conducted qualitative assessments of EWS, usually through visual inspection of graphically presented results (e.g., Robinson and Uehlinger 2008, Bestelmeyer et al. 2011, Carpenter et al. 2011, Spanbauer et al. 2014; see Appendix S1: Table S1 for complete list). The relatively common use of qualitative assessments in EWS studies is an implicit recognition of the fact that rigorous tests are often impossible within the constraints of empirical data. Qualitative tests are certainly preferable to quantitative tests that are based on flawed assumptions, and they have been instrumental in making early progress in the empirical application of EWS. However, study designs that are able to produce quantitative assessments will be necessary to reduce subjectivity and truly assess the robustness and utility of EWS.

Hysteresis as a hallmark of nonlinearity

Since so much of EWS theory is based on nonlinear dynamics, and there is a greatly reduced expectation for EWS in linear systems, there is an obvious motivation for testing systems for nonlinear dynamics before they are used in EWS studies. However, this introduces another set of analytical difficulties. Formal tests for nonlinearity, such as the BDS test (after the initials of W. A. Brock, W. Dechert and J. Scheinkman; Brock et al. 1996, Dakos et al. 2012*a*) or S-maps (Hsieh et al. 2005), are data-intensive and often unsuitable for short ecological time series. Methods exist for joining short time series to produce synthetic data sets that are long enough for nonlinearity tests (Hsieh et al. 2008). However, these have not been evaluated in terms of their effects on EWS and shift detection. This difficulty in testing for nonlinearity mirrors long-standing difficulty in applying nonlinear models, such as the fold bifurcation, to real ecosystems. These models (and the theoretical justification for EWS that they produce) have been most successfully applied to simple ecosystems, such as shallow freshwater lakes (Scheffer et al. 1993). However, even in these simple systems, empirical evidence for nonlinear dynamics such as alternative states is surprisingly sparse at the ecosystem scale (Schröder et al. 2005, Pinto and O'Farrell 2014). In spite of this paucity of evidence in even simple systems, heuristic arguments have often been used to invoke the fold bifurcation model to explain shifts in large complex systems where empirical support is correspondingly more difficult to obtain. Often-cited putative examples of alternative states or critical transitions in large systems include ecosystem regime shifts in the North Pacific (Scheffer et al. 2001, Scheffer and Carpenter 2003), coral–algae transitions on tropical reefs (Hughes 1994, Mumby et al. 2006, Mumby and Steneck 2008, Norström et al. 2009), and fishing-induced community reorganization in the North Atlantic (Choi et al. 2004). However, these examples may be more parsimoniously explained, respectively, by autocorrelated random variability (i.e., red noise; Rudnick and Davis 2003, Di Lorenzo and Ohman 2013), linear, reversible tracking of perturbation (Dudgeon et al. 2010), and transient responses to perturbation (Frank et al. 2011). The possible over-application of nonlinear theory may have important consequences for ecological understanding (Dudgeon et al. 2010), and results in considerable confusion among empiricists and managers concerning the best approaches for studying and managing sudden ecological transitions (Bestelmeyer et al. 2011).

These issues in reconciling theory with empirical observations will not be cleared up soon. However, an intermediate step providing increased empirical rigor in the application of alternative state theory is possible: testing for signs of hysteresis in ecological driver–response relationships (Bestelmeyer et al. 2011, Hunsicker et al. 2016). This approach is consistent with theoretical constructs for the range of possible biological responses to environmental drivers, from linear tracking, to threshold responses, to hysteresis (Fig. 2; Scheffer et al. 2001, Andersen et al. 2009, Pace et al. 2015). Hysteresis is present when the biological response fails to reverse its initial path after a perturbation is reversed, due to positive feedbacks in the system or a difference in the set of variables controlling the biological response in different states (Scheffer et al. 2001, Beisner et al. 2003). While formal tests for hysteresis are extremely difficult in observational systems (Schröder et al. 2005, Dudgeon et al. 2010, Faassen et al. 2015), tests for state-dependent behavior in driver–response relationships are statistically straightforward and allow empiricists to test for dynamics that are consistent with hysteresis. Specifically, the expectation for linear tracking is that the response is a continuous (linear) function of the driver, across the time series. The expectation for hysteresis is that the response is quantitatively different in different system states, and the expectation for the threshold model is that the driver and response show a sigmoidal relationship (Fig. 2; Samhouri et al. 2010, Bestelmeyer et al. 2011). This test provides necessary but not sufficient evidence for the presence of hysteresis and

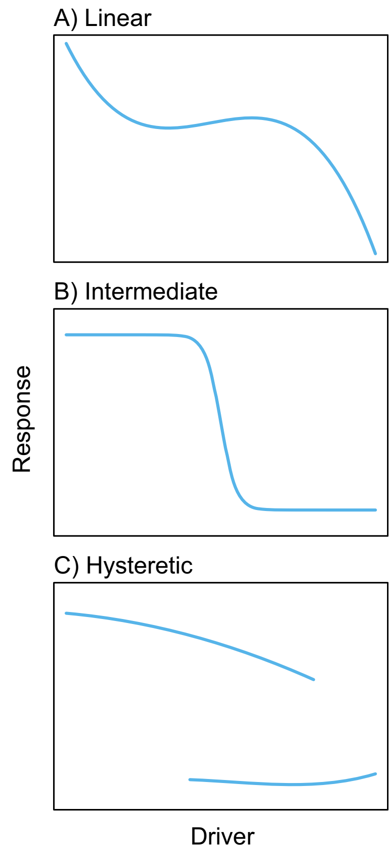

Fig. 2. Schematic of range of possible driver–response relationships, from (A) linear tracking of environmental conditions (in this case illustrated by a cubic function, which is a linear combination of model parameters); to (B) an intermediate response with a strong threshold; and (C) a state-dependent response consistent with hysteresis. Note that both (B) and (C) are nonlinear relationships.

alternative states (Beisner et al. 2003, Petraitis and Dudgeon 2016). This approach offers an attractive compromise between the need to demonstrate the suitability of a given system for EWS research, and the difficulties of testing for nonlinear dynamics in empirical systems.

Methods

Meta-analysis of published results

To compare the results of empirical EWS studies that do and do not demonstrate nonlinearity in study systems, we searched the Web Of Science database and references cited within the literature to identify examples of empirical EWS studies published in print or online from 2006 through 2015. We included only non-laboratory examples from the natural world (e.g., examples from economics were excluded) where the predictions of EWS theory were explicitly tested. Studies merely documenting trends in EWS (e.g., Boulton and Lenton 2015) or using EWS to test for nonlinear dynamics (e.g., Lenton et al. 2012) were not included. Our meta-analysis did include climate studies, as this literature includes a large proportion of the available empirical examples, and the utility of EWS in climate systems has direct bearing on ecology. Approximately one quarter of the empirical studies that we found presented only qualitative comparisons of EWS observations and predicted behavior (Appendix S1: Table S1). We judged that the results of these qualitative tests were not rigorous enough for the purposes of a meta-analysis of published analyses. Accordingly, we only included studies that presented some sort of quantitative test of EWS predictions—either a statistical hypothesis test, a model selection result, or the results of shift detection routines run simultaneously on time series of mean values and EWS.

We categorized the systems in these examples as either nonlinear or linear. In the former category, we included examples for which the results of a formal test for nonlinearity were either presented or cited. In the latter category, we included examples for which no formal test for nonlinearity was presented. None of the published papers presented tests for nonlinearity with negative results.

We then categorized each example as supporting or not supporting predictions for EWS. From the perspective of an empiricist looking for theoretical guidance on the expected behavior of EWS, the literature continues to make inconsistent predictions. For instance, while the most commonly invoked metric-based EWS are increasing variance, increasing autocorrelation, and increasing skewness prior to a transition, some models predict decreasing values of these metrics in some situations (Dakos et al. 2012*b*, Wang et al. 2012). In order to avoid the complication of comparing tests of inconsistent predictions, we limited the meta-analysis to examples where authors expected rising trends in the one or more of these three metrics prior to a transition. We then compared the proportion of positive and negative results between study systems for which nonlinear dynamics were or were not demonstrated. The sampling unit for this analysis was the test of an individual EWS metric within an individual study system. Individual studies often contained multiple EWS tests, sometimes in more than one study system.

Analysis of northeast Pacific time series

*Data.—*In order to further test for differences in the success of EWS applied to nonlinear vs. linear systems, and to provide an example of possible approaches for dealing with this distinction using real data sets, we conducted a comparative analysis of eight northeast Pacific time series. This analysis used four time series from Alaskan ecosystems and four time series from the northern California Current ecosystem. Data from Alaska include measures of community composition or distribution from trawl surveys in Pavlof Bay (Gulf of Alaska) and the Bering Sea, and time series of mean length for juveniles (age-1) for two gadid species (Pacific cod *Gadus macrocephalus* and walleye pollock *G. chalcogrammus*) from the Bering Sea. Data from the California Current are population abundance time series for three copepods (*Acartia longiremis*, *Pseudocalanus mimus*, and *Paracalanus parvus*) and a naturally produced population of coho salmon (*Oncorhynchus kisutch*). The eight time series respond to climate variability on monthly, annual, or decadal time scales. For detailed descriptions of each time series, see Appendix S1: Detailed Methods.

*Testing for nonlinearity.—*We considered and rejected several possible tests for nonlinearity in the northeast Pacific data. Formal tests for nonlinear time series behavior, such as the BDS test or

S-maps, generally require longer time series than the 31–34 years available in Alaska. Tests for multimodal distributions in system state have previously been used as informal tests for nonlinearity, though they may also indicate multimodality in environmental drivers (Scheffer and Carpenter 2003, Wang et al. 2012). However, when we applied Hartigans' dip test (Bestelmeyer et al. 2011) to test for multimodal distributions of the northeast Pacific time series, the results were difficult to interpret. In particular, analysis of distributions that appeared heuristically bimodal failed to reject the null hypothesis of unimodality (e.g., Pavlof Bay community time series; Appendix S1: Fig. S1). Bestelmeyer et al. (2011) noted the conservative nature of this test with small sample sizes, and we judged that more experience is needed with this test before it can be employed in restricted data settings such as ours.

We opted to test these time series for behavior consistent with hysteresis. Specifically, we used a model selection approach to test for state dependence in driver–response relationships (Bestelmeyer et al. 2011). Candidate models compared biological response to a climate driver between different climate states: either warm or cool states in the Pacific Decadal Oscillation (PDO) Index for the California Current time series, warm and cold states in the Pavlof Bay community composition time series, or periods of warming and cooling bottom temperature for Bering Sea time series. The inference here is that persistent shifts in the driver carry the response variable between different states, consistent with an extensive literature that ascribes different northeast Pacific community states to different states in environmental conditions (Mantua et al. 1997, Anderson and Piatt 1999, Benson and Trites 2002). "Linear" models invoked driver–response relationships that are unchanging across system states (linear regression or GAMs Fig. 2). Note that "linear" in this sense refers to dynamics that can be expressed by a statistical model containing linear combinations of model parameters, which includes polynomial regression or corollary dynamics in a GAM. "Hysteresis" models invoked responses that changed between system states, as reflected by the interactive effect between a state variable and the driver variable (Bestelmeyer et al. 2011). Candidate interactive models were constructed both with linear regression and with GAMs. We also tested three-parameter sigmoidal regressions to model strong threshold responses that were intermediate between the linear and hysteresis models (Samhouri et al. 2010, Bestelmeyer et al. 2011, Hunsicker et al. 2016). We limited the degree of smoothing in GAMs to three effective degrees of freedom in order to prevent over-fitting. Competing models were evaluated with Akaike's information criterion, adjusted for sample size (AIC_c).

Testing EWS.—All of the time series used in this analysis suffered from temporal autocorrelation that violated the assumption of independent samples. In order to conduct valid hypothesis tests for EWS in the presence of autocorrelation, we used a simple ad hoc randomization approach (Manly 2006) following four steps:

1. The desired statistical test was conducted on the data. The *P*-value from this test was ignored, but the test statistic (t or τ) was retained.
2. A distribution of the test statistic was generated from randomized time series with the same length and first-order autocorrelation as the original data. Unless noted otherwise, randomizations consisted of 10,000 permutations. Randomized time series were generated with function arima.sim in the "stats" package of computer language R (R Core Team 2016), which produces time series with a distribution of AR(1) values centered on the specified value. In an example, for 10,000 time series randomized with AR(1) = 0.6 specified, the mean AR(1) value over 100 time steps was 0.57, with SD of 0.08.
3. For a one-tailed test (i.e., for significantly rising or elevated values), the *P*-value of the observed results under the null hypothesis was estimated as the proportion of randomized tests with a test statistic equal to or greater than the test statistic from the observed data. Two-tailed *P*-values were estimated as the proportion of randomized absolute value test statistics equal to or greater than the absolute value of the test statistic from the actual data.

The EWS tested for in analysis differed among time series. For the Pavlof Bay community composition and gadid length time series, details of how the variables were calculated meant that rising spatial variability was the only EWS available. For the Bering Sea community distribution time series, rising spatial variability, rising spatial correlation, and rising temporal autocorrelation were used. For the four California Current time series, rising temporal variability and rising temporal autocorrelation were used.

For the Pavlof Bay community composition and the four northern California Current time series, our EWS test was formulated as a test of the hypothesis that shifts in mean values are preceded by rising EWS. For the three Bering Sea time series (community distribution and length for the two gadids), we tested the hypothesis that a persistent perturbation should be accompanied by rising EWS. For Pavlof Bay community composition, we used a one-tailed test on the linear trend in variability from 1972 until 1978, the year when a shift in variability was documented in the original study. For Bering Sea community distribution and the two juvenile gadid length time series, we used a one-tailed test for increased EWS during the cold anomaly compared with other years in the time series. For California Current time series, the hypothesis was tested with shifts in mean values detected using the shift detection approach presented by Bestelmeyer et al. (2011), using the "strucchange" package in R. This approach uses the CUSUM test (cumulative sums of ordinary least squares residuals) to identify break points in the data and Bayesian information criterion model selection to identify the best number of shifts for describing the time series. Shift detection analysis was conducted with raw time series values rather than detrended or differenced data. Copepod abundance EWS were calculated on 16-month sliding windows within data that had been detrended with Gaussian kernels, following the approach of Dakos et al. (2012*a*) and the package "earlywarnings" in R. The length of these sliding windows was based on the life span of the study species and the temporal scale of their response to PDO variability. The sliding window for EWS calculation in coho salmon abundance was set at 25% of time series length (16 yr). For all four time series, we tested for increasing EWS across 12 time steps prior to shifts in mean values, using a one-tailed randomization test.

The longer time series available for California Current copepods also allowed us to test for false-positive signals as an additional EWS test. This took the form of a test of the hypothesis that periods of rising trends in EWS are more likely to be followed by changes in the system mean than periods without rising EWS trends. To test this hypothesis, we first identified statistically significant increases in every EWS time series using a one-tailed randomization test. We then compared the slopes in mean values for the 12-month windows following these significant changes with other 12-month windows in the time series, using a two-tailed randomization test.

Comparison across time series.—We used the individual time series as the sampling unit in our comparison of EWS test results between time series with linear and hysteretic driver–response relationships. We first estimated the combined probability of observed EWS behavior from multiple EWS tests within a time series with Fisher's combined probability test (Sokal and Rohlf 1995, Dakos et al. 2008), which estimates the χ^2 statistic of the overall probability as

$$\chi^2_{2k} = -2\ln\sum_{i=1}^{k}\ln(P_i)$$

where k is the number of tests and P_i are the P-values from individual tests. This gave us a single P-value for each time series; a second round of the combined probability test was then used to evaluate the overall probability across the two groups of time series (linear and hysteretic).

Results

Meta-analysis of published studies

We retrieved explicit EWS tests from 25 study systems reported in 19 studies (Table 1). Nonlinearity was demonstrated in only six of the 25 systems. These systems produced eight positive EWS tests and no negative tests. There was no demonstration of nonlinearity for 19 study systems. These systems produced a much more mixed set of results (15 positive and 14 negative tests; test for independence in the linear and nonlinear distribution of positive and negative results, $\chi^2_1 = 4.33$, $P = 0.037$, Fig. 1C).

Comparative analysis of northeast Pacific data

Driver–response relationships.—In Alaska, the Pavlof Bay community composition (Fig. 3B), Bering Sea community distribution (Fig. 3E), and

Table 1. Summary of study type, analysis type, evidence for nonlinearities, and early warning signals test results retrieved from the 19 published studies (25 examples) used in the meta-analysis.

Study ID	Linear/nonlinear	System type	Analysis type	Replication	n	Test for false positives?	Var.	Autocor.	Skew.
1	Nonlinear	Fw	1	T	6	No	S	S	–
2	Nonlinear	Int	2	T	40	No	S	S	S
3	Nonlinear	Lake	2	S	1	No	S	–	–
4	Nonlinear	Fw	2	T	1	No	S	–	–
5	Nonlinear	Marine	2	T	1	No	S	–	–
6	Nonlinear	Fw	2	T	2	No	–	–	–
7	Linear	Marine	1	S	4	No	S	–	–
7	Linear	Marine	1	T	5	No	S	–	–
8	Linear	Fw	2	T	83	No	NS	NS	–
8	Linear	Fw	2	T	19	Yes	NS	NS	–
8	Linear	Marine	2	T	24	Yes	NS	NS	–
9	Linear	Climate	2	T	1	Yes	S	–	–
10	Linear	Climate	2	T	8	No	–	S	–
11	Linear	Climate	2	T	25	No	NS	NS	–
12	Linear	Climate	2	T	1	No	–	–	NS
13	Linear	Int	2	T	1	No	S	–	–
14	Linear	Mar/Fw	2	S	86	No	S	S	–
14	Linear	Mar/Fw	2	S	37	No	NS	NS	–
14	Linear	Mar/Fw	2	S	40	No	S	S	–
15	Linear	Marine	2	B	2	No	NS	NS	–
16	Linear	Marine	1	S	1	No	S	–	–
16	Linear	Marine	1	S	1	No	S	–	–
17	Linear	Marine	2	S	14	Yes	S	–	NS
18	Linear	Climate	1	T	7	No	–	S	–
19	Linear	Marine	2	T	2	No	S	S	–

Notes: References for the publications and additional study details are provided in Appendix S1: Table S1. Study ID is cross-referenced with Appendix S1: Tables S1 and S2. Fw, freshwater; Int, intertidal; Mar, marine; T, temporal; S, spatial; B, both; n, number of time series; Var., authors tested for rising variance; Autocor., authors tested for rising autocorrelation; Skew., authors tested for rising skewness; S, supporting evidence; NS, no supporting evidence. Analysis type: 1, shift timing; 2, significance test or model selection.

Pacific cod length time series (Fig. 3H) were best described by state-dependent driver–response relationships consistent with hysteresis. Walleye pollock length was best described by a linear model (Fig. 3K). In the California Current, *Pseudocalanus mimus* abundance was best described by a state-dependent driver–response relationship (Fig. 4E), while the other three time series showed linear relationships with the PDO (Fig. 4B, H, K). Complete model selection results are presented in Table 2.

EWS tests.—Predictions for EWS were supported in both community-level metrics from Alaska. Spatial variability increased prior to the shift in Pavlof Bay community composition ($P = 0.05$, Fig. 3C). Spatial variability, spatial autocorrelation, and temporal autocorrelation in Bering Sea community distribution increased during the persistent perturbation of the cold anomaly ($P < 0.00001$, Fig. 3F). However, negative results were obtained for Pacific cod and walleye pollock length; spatial variability for neither time series increased during the perturbation ($P \geq 0.2$; Fig. 3I, L).

In the California Current copepod data, the best shift detection models invoked either three of four shifts in each time series, with similar timing among time series, as is expected given their shared sensitivity to the PDO (Fig. 4A, D, G). Coho salmon abundance shifts were defined at 1976/1977, 1990/1991, and 2000/2001, corresponding to low-frequency changes in the sign of the PDO Index (Fig. 4J). However, these time series did not support EWS predictions; EWS trends (linear slopes on time) did not increase prior to these shifts in any of the four time series (Fig. 4C, F, I, L; $P \geq 0.18$). Randomization tests with Kendall's τ as the test statistic rather than t-values

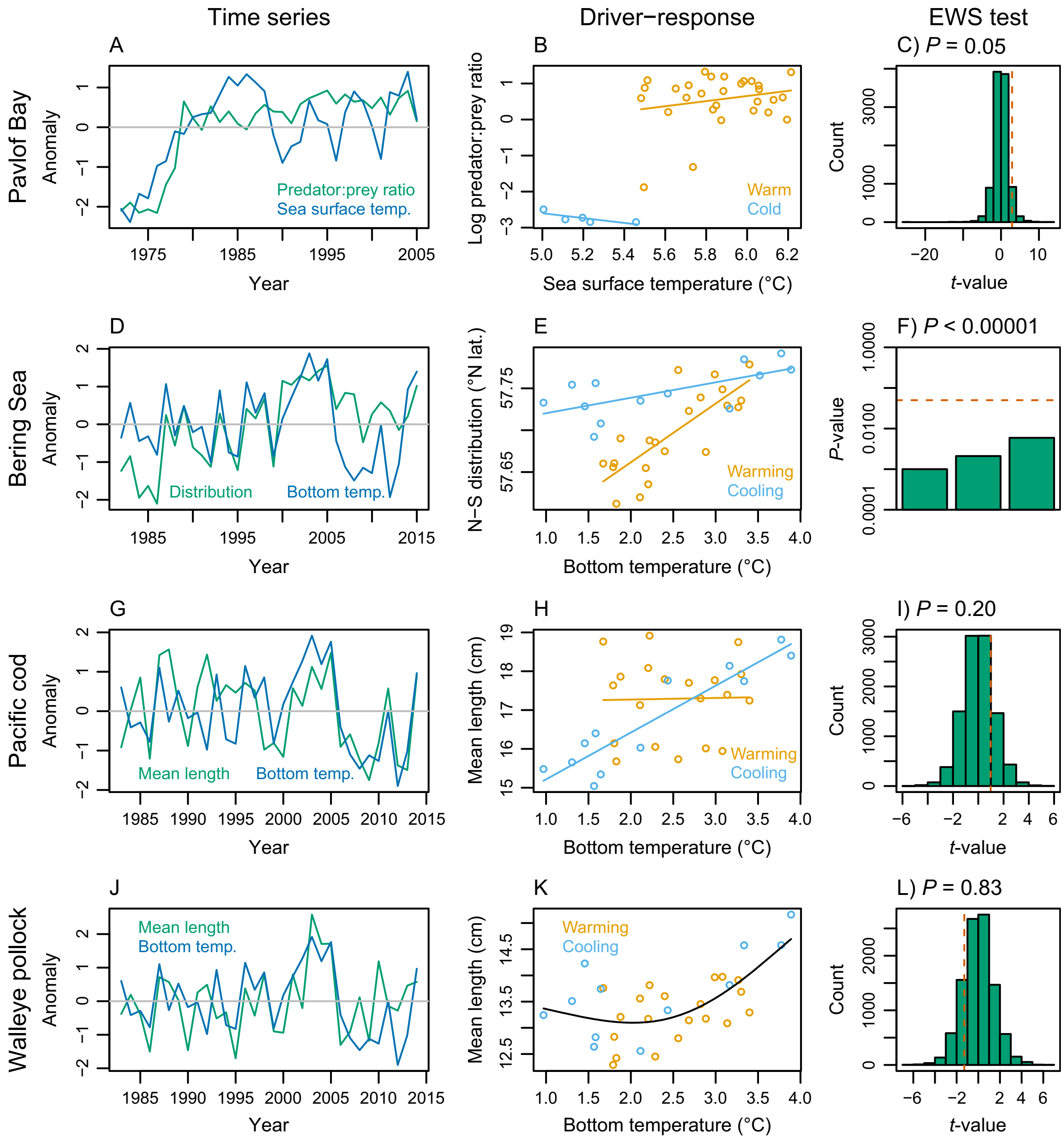

Fig. 3. Results of early warning signal (EWS) tests in Alaskan ecosystems: Pavlof Bay predator:prey ratios (A–C), Bering Sea community distribution (D–F), age-1 Pacific cod length (G–I), and age-1 walleye pollock length (J–L). Left-hand panels plot time series of standardized values of response values and their primary drivers. Middle panels plot best driver–response models and indicate different driver states (either warm/cold or warming/cooling). Right-hand panels plot results of randomization tests for increasing EWS prior to a historical shift (C), or during periods of increased perturbation (F, I, L). Bars in (C, I, L) indicate the distribution of *t*-values for linear models of variability trends from randomized data; vertical dashed lines indicate *t*-value observed in the system; bars in (F) plot the *P*-values from three separate EWS tests.

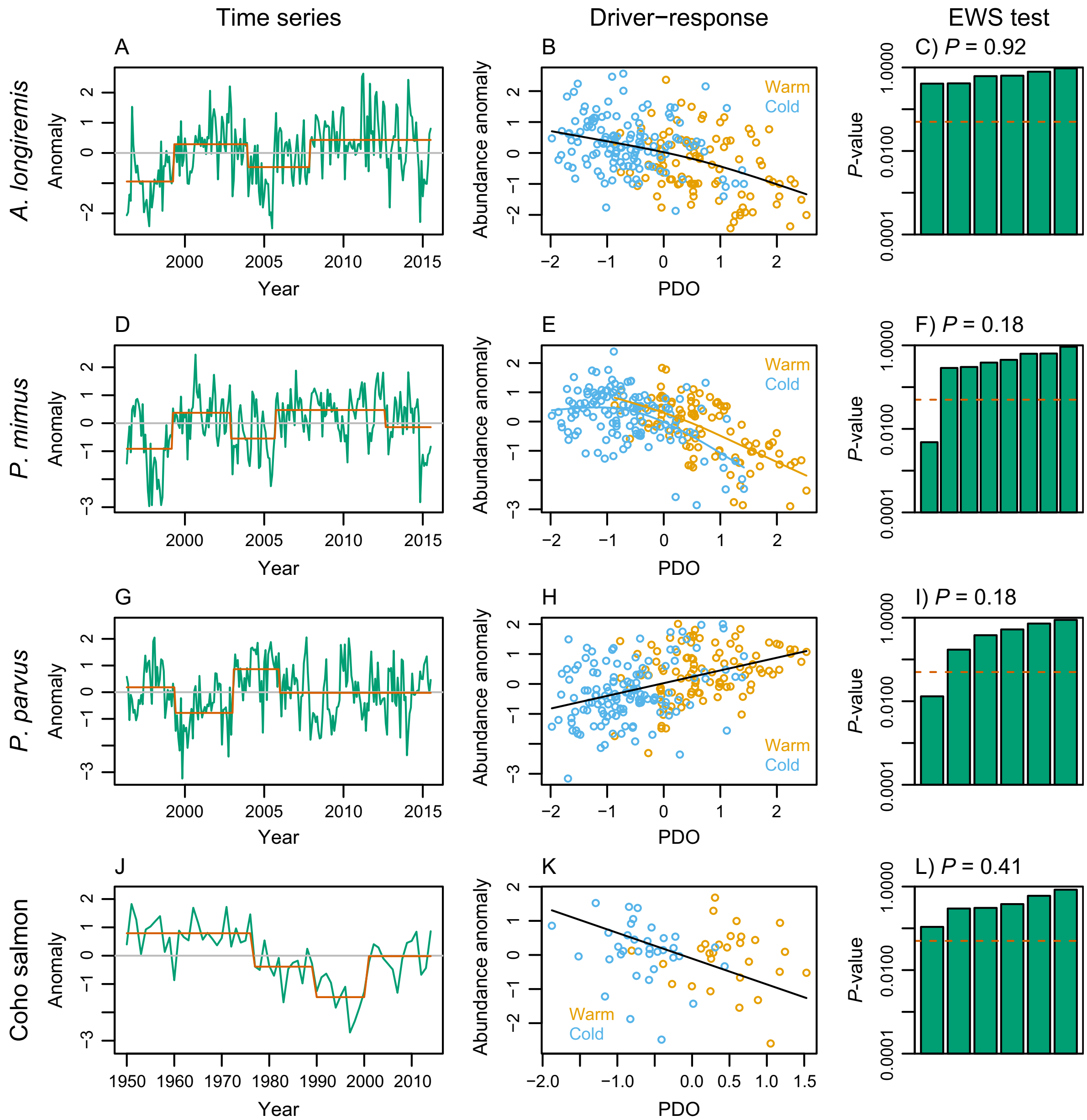

Fig. 4. Results of early warning signal (EWS) tests in the California Current ecosystem: abundance of *Acartia longiremis* (A–C), *Pseudocalanus mimus* (D–F), *Paracalanus parvus* (G–I), and coho salmon (J–L). Left-hand panels plot abundance time series (green lines) and states defined by shift detection routine (red lines). Middle panels plot best driver–response models and indicate warm/cold states in the PDO. Right-hand panels plot results of randomization tests for rising temporal variability and autocorrelation prior to the shifts indicated by red lines in (A, D, G). Horizontal dashed line indicates $P = 0.05$.

from linear regression produced similar results (data not shown). The tests for false positives in the copepod data reinforced the negative results of the primary EWS test. False positives were extremely common in the copepod abundance time series, and these data did not support predictions for EWS behavior. None of the three time series showed inconsistency with the null

Table 2. Results of model selection comparing linear, intermediate, and hysteretic driver–response models for eight northeast Pacific time series: Δ-AIC_c values.

Time series	n	Linear		Intermediate	Hysteretic	
		Linear regression	GAM	Sigmoidal‡	Linear interaction	GAM interaction
Acartia longiremis abundance	231	1.049	**0.000**	NA	0.563	0.619
Pseudocalanus mimus abundance	231	20.793	5.545	NA	12.489	**0.000**
Paracalanus parvus abundance	231	**0.000**	0.000†	NA	3.465	1.803
Coho salmon abundance	65	**0.000**	0.000†	NA	2.343	0.081
Pavlof Bay community composition	34	21.068	13.820	2.038	**0.000**	0.004
Bering Sea community distribution	34	24.029	13.835	22.459	**0.000**	13.478
Pacific cod juvenile length	31	3.841	3.841	NA	1.580	**0.000**
Walleye pollock juvenile length	31	6.979	**0.000**	2.822	4.781	1.683

Notes: Best model in each case is indicated in boldface. AIC_c, Akaike's information criterion, corrected for sample size; GAM, generalized additive model.
† Best GAM was identical to the linear regression.
‡ Sigmoidal model successfully fit only three time series.

hypothesis (combined probability over two EWS for each time series, $P > 0.1$; Appendix S1: Fig. S2).

Comparison across time series.—The combined probability of EWS tests for time series showing hysteretic driver–response relationships (Pavlof Bay community composition, Bering Sea community distribution, Pacific cod length, *P. mimus* abundance) led to rejection of the null hypothesis of no EWS behavior prior to shifts ($P < 0.00001$). On the other hand, combined probability for time series with linear driver–response relationships (walleye pollock length and *Acartia longiremis*, *Paracalanus parvus*, and coho salmon abundance) failed to reject the null hypothesis ($P = 0.67$).

Discussion

Early warning signals (EWS) may eventually prove to be immensely useful as management tools. Controlled tests based on manipulations in laboratories and simple ecosystems provide firm empirical support for EWS that is particularly important for demonstrating their potential for real-world applications (Drake and Griffen 2010, Carpenter et al. 2011, Dai et al. 2012, 2013, 2015, Seekell et al. 2012, Veraart et al. 2012, Pace et al. 2013, Cline et al. 2014). However, assessing EWS in larger, more complex ecosystems presents considerable difficulties in terms of study design and statistical approach. Recognizing these hurdles is essential for EWS to be properly assessed in real ecosystems. Our results suggest that testing for nonlinear dynamics or signs of hysteresis is a key step for improving field studies of EWS and hastening the testing and application of this promising idea.

The central role of nonlinearity

Our analysis of the literature suggests that tests in nonlinear systems are more likely to produce successful empirical examples of EWS than are systems where nonlinearity has not been formally demonstrated (Fig. 1C). Because the number of published examples is still quite small for the purposes of statistical analysis, we treated conclusions about separate EWS within studies, or conclusions from separate ecosystems within studies, as independent events in order to provide a reasonable sample size for the meta-analysis. This decision introduces the risk of pseudo-replication, so the result should be considered with caution. Still, the pattern is quite remarkable: Examples with a formal demonstration of nonlinearity produced exclusively positive EWS tests, while examples without such a test produced mixed outcomes.

A similar result was produced in our comparative analysis of northeast Pacific time series. The four time series with driver–response relationships consistent with hysteresis included two time series that supported EWS predictions (Pavlof Bay community composition and Bering Sea community distribution) and two that did not (*Pseudocalanus mimus* abundance and Pacific cod length). The overall distribution of P-values for these four time series rejects the null hypothesis of no EWS behavior ($P < 0.00001$). The four time series showing linear driver–response relationships (*A. longiremis, P. parvus*, and coho salmon abundance and pollock length), on the other hand, individually each failed to reject the null hypothesis, as did the distribution of *P*-values across these time series ($P = 0.67$).

We recognize that our model selection approach for determining whether driver–response relationships are or are not consistent with hysteresis is an imperfect test for assessing nonlinear dynamics. These results are vulnerable to the effects of small sample size, noise, confounding dynamics, and alternate interpretation that plague many attempts to test for complex dynamics with real-world observational data. Furthermore, the presence of hysteresis is not an ironclad guarantee that the fold bifurcation/alternative state model at the base of much EWS theory is actually present in these systems (Beisner et al. 2003). Nor can a simple statistical test for an interactive state effect in driver–response relationships distinguish hysteresis from other phenomena, such as transient responses and asymmetry in basins of attraction (Beisner et al. 2003, Frank et al. 2011). This approach should therefore be seen as a compromise between what we would wish and what is possible with real data. The approach is able to identify time series with strictly linear relationships to environmental drivers (e.g., *A. longiremis, P. parvus,* and coho salmon abundance; Fig. 4B, H, K), which are therefore poor candidates for the application of EWS. In cases where driver–response relationships consistent with hysteresis are detected, mechanistic understanding of the system can be used to evaluate the likelihood of nonlinear dynamics. For instance, interactions between Pacific cod and their prey in Pavlof Bay (Fig. 3A–C) are an example of oscillating trophic control, with the system progressing from bottom-up control at the start of the time series, through top-down control during the community reorganization, and back to bottom-up control (Litzow and Ciannelli 2007). These population controls that differ across the time series provide a possible explanation for state-dependent relationships between temperature and community state (Scheffer et al. 2001). The Pavlof Bay community composition time series is also notable because it is a reaction to a sudden perturbation in an external driver, the PDO. Normally such a sudden trigger is not expected to produce EWS, which are more likely in cases where the perturbation is slow relative to the other dynamics of the system (Dakos et al. 2015). However, in this case, the sudden perturbation appears to have produced a nonlinear biological response that was characterized by both state-dependent driver–response relationships and EWS. This contrasts with the four California Current time series, which are also driven by the PDO, and supports the prediction that EWS should not be present for reactions to sudden perturbation (Fig. 4). In the case of the Bering Sea community distribution time series (Fig. 3D–F), the mechanisms underlying state-dependent driver–response relationships are less clear. Community resilience increases with diversity (Frank et al. 2006), and warming since the 1980s has increased diversity of the Bering Sea demersal community (Mueter and Litzow 2008). This increase in diversity may therefore explain the apparent hysteresis in the community response to temperature as diversity-driven differences in resilience of the cold and warm community states. Beyond this general observation, understanding of the eastern Bering Sea ecosystem is not advanced enough to provide mechanistic understanding of possible hysteresis in the system. Similarly, state-dependent responses to temperature variability are known in gadids. For example, the sign of recruitment–temperature relationship in Atlantic cod (*Gadus morhua*) depends on the mean temperature (Drinkwater 2005). However, the mechanism creating apparent hysteresis in the Pacific cod length time series, with a switch from insensitivity to temperature during the warming period, to temperature sensitivity during the cooling period (Fig. 3G, H) remains mysterious. Walleye pollock length showed a linear relationship to temperature across the warming and cooling periods (Fig. 3K), and the reason for this different response between the two species is unknown. Finally, the generally linear nature of the relationships between the PDO and copepods and coho salmon in the northern California Current (Fig. 4B, H, K) is consistent with the expectation that copepod population dynamics are driven by linear relationships, characterized by reddening of environmental variability that is integrated over the mean life span of a population (Di Lorenzo and Ohman 2013).

The results of our meta-analysis and comparative analysis, demonstrating the importance of nonlinearity or hysteresis for the successful application of EWS, are unsurprising given the overwhelming importance of the fold bifurcation model in the theoretical literature. Indeed, EWS are so strongly associated with nonlinearity that they have been suggested as a means of distinguishing linear and nonlinear variability in time series (e.g., Ditlevsen and Johnsen 2010, Lenton et al. 2012). Given such strong theoretical and empirical association between EWS and nonlinearity, why did nearly 80% of the examples in our meta-analysis come from systems for which no tests for nonlinearity were conducted?

The scarcity of tests for nonlinearity in empirical EWS studies, and the central role of nonlinearity in the theory, is part of a broader pattern of differences between empirical and theoretical ecology. Nonlinear dynamics have been a cornerstone of mathematical ecology for decades (Lotka 1956, Lewontin 1968, Holling 1973, May 1977). But the majority of empirical research continues to be conducted with statistical tools assuming linear relationships (Deyle et al. 2013). We see three reasons for this dichotomy. The first, outlined in the Introduction, is the analytical difficulty in assessing nonlinearity with short time series from the real world. The second has to do with differences between mathematical models and reality (Scheffer et al. 2015). In real ecosystems, stochastic perturbations and weak interactions between a large number of interacting variables may prevent nonlinear dynamics from occurring (Bjørnstad 2015). Finally, both empirical (Schröder et al. 2005) and theoretical (Petraitis and Dudgeon 2016) studies indicate that the fold bifurcation and alternative state models are only one possible configuration of many that ecosystems may take on. Similarly, population variability and driver–response relationships may commonly reflect linear dynamics (Glaser et al. 2014, Hunsicker et al. 2016). In other words, the nonlinear dynamics that EWS are largely based on likely represent only a subset of ecological situations. These sources of uncertainty concerning the application of nonlinear models underscore the importance of the approach that we advance in our comparative analysis to test for signs of consistency with the models justifying EWS.

Acknowledgments

We thank two anonymous reviewers for constructive comments on an earlier version of this manuscript. For helpful discussions and providing data, we thank Jennifer Fisher, Bob Lauth, Peter Lawson, Dan Nichol, Bill Peterson, and the Ocean Tipping Points Project Team, including Rod Fujita, Carrie Kappel, Kendra Karr, and Courtney Scarborough. MAL was supported by a grant from the Pew Charitable Trusts.

Literature Cited

Andersen, T., J. Carstensen, E. Hernandez-Garcia, and C. M. Duarte. 2009. Ecological thresholds and regime shifts: approaches to identification. Trends in Ecology and Evolution 24:49–57.

Anderson, P. J., and J. F. Piatt. 1999. Community reorganization in the Gulf of Alaska following ocean climate regime shift. Marine Ecology Progress Series 189:117–123.

Beaugrand, G., M. Edwards, K. Brander, C. Luczak, and F. Ibanez. 2008. Causes and projections of abrupt climate-driven ecosystem shifts in the North Atlantic. Ecology Letters 11:1157–1168.

Beisner, B. E., D. T. Haydon, and K. Cuddington. 2003. Alternative stable states in ecology. Frontiers in Ecology and the Environment 1:376–382.

Benson, A. J., and A. W. Trites. 2002. Ecological effects of regime shifts in the Bering Sea and eastern North Pacific Ocean. Fish and Fisheries 3:95–113.

Bestelmeyer, B. T., et al. 2011. Analysis of abrupt transitions in ecological systems. Ecosphere 2:129.

Bjørnstad, O. N. 2015. Nonlinearity and chaos in ecological dynamics revisited. Proceedings of the National Academy of Sciences USA 112:6252–6253.

Boerlijst, M. C., T. Oudman, and A. M. de Roos. 2013. Catastrophic collapse can occur without early warning: examples of silent catastrophes in structured ecological models. PLoS ONE 8:e62033.

Boettiger, C., and A. Hastings. 2012. Quantifying limits to detection of early warning for critical transitions. Journal of the Royal Society Interface 9:2527–2539.

Boettiger, C., N. Ross, and A. Hastings. 2013. Early warning signals: the charted and uncharted territories. Theoretical Ecology 6:255–264.

Boulton, C. A., and T. M. Lenton. 2015. Slowing down of North Pacific climate variability and its implications for abrupt ecosystem change. Proceedings of the National Academy of Sciences USA 112:11496–11501.

Brock, W. A., and S. R. Carpenter. 2010. Interacting regime shifts in ecosystems: implication for early warnings. Ecological Monographs 80:353–367.

Brock, W. A., W. Dechert, J. Scheinkman, and B. LeBaron. 1996. A test for independence based upon the correlation dimension. Econometric Review 15:197–235.

Burthe, S. J., et al. 2016. Do early warning indicators consistently predict nonlinear change in long-term ecological data? Journal of Applied Ecology 53:666–676.

Carpenter, S. R., and W. A. Brock. 2006. Rising variance: a leading indicator of ecological transition. Ecology Letters 9:308–315.

Carpenter, S. R., et al. 2011. Early warnings of regime shifts: a whole-ecosystem experiment. Science 332:1079–1082.

Carstensen, J., and A. Weydmann. 2012. Tipping points in the Arctic: Eyeballing or statistical significance? Ambio 41:34–43.

Choi, J. S., K. T. Frank, W. C. Leggett, and K. Drinkwater. 2004. Transition to an alternate state in a continental shelf ecosystem. Canadian Journal of Fisheries and Aquatic Sciences 61:505–510.

Clements, C. F., J. M. Drake, J. I. Griffiths, and A. Ozgul. 2015. Factors influencing the detectability of early warning signals of population collapse. American Naturalist 186:50–58.

Cline, T. J., D. A. Seekell, S. R. Carpenter, M. L. Pace, J. R. Hodgson, J. F. Kitchell, and B. C. Weidel. 2014. Early warnings of regime shifts: evaluation of spatial indicators from a whole-ecosystem experiment. Ecosphere 5:102.

Dai, L., K. S. Korolev, and J. Gore. 2013. Slower recovery in space before collapse of connected populations. Nature 496:355–358.

Dai, L., K. S. Korolev, and J. Gore. 2015. Relation between stability and resilience determines the performance of early warning signals under different environmental drivers. Proceedings of the National Academy of Sciences USA 112:10056–10061.

Dai, L., D. Vorselen, K. S. Korolev, and J. Gore. 2012. Generic indicators for loss of resilience before a tipping point leading to population collapse. Science 336:1175–1177.

Dakos, V., and J. Bascompte. 2014. Critical slowing down as early warning for the onset of collapse in mutualistic communities. Proceedings of the National Academy of Sciences USA 111:17546–17551.

Dakos, V., S. R. Carpenter, E. H. van Nes, and M. Scheffer. 2015. Resilience indicators: prospects and limitations for early warnings of regime shifts. Philosophical Transactions of the Royal Society B: Biological Sciences 370:20130263.

Dakos, V., S. Kéfi, M. Rietkerk, E. H. van Nes, and M. Scheffer. 2011. Slowing down in spatially patterned ecosystems at the brink of collapse. American Naturalist 177:E153–E166.

Dakos, V., M. Scheffer, E. H. Van Nes, V. Brovkin, V. Petoukhov, and H. Held. 2008. Slowing down as an early warning signal for abrupt climate change. Proceedings of the National Academy of Sciences USA 105:14308–14312.

Dakos, V., et al. 2012*a*. Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. PLoS ONE 7:e41010.

Dakos, V., E. H. van Nes, P. D. D'Odorico, and M. Scheffer. 2012*b*. Robustness of variance and autocorrelation as indicators of critical slowing down. Ecology 93:264–271.

Deyle, E. R., M. Fogarty, C. H. Hsieh, L. Kaufman, A. D. MacCall, S. B. Munch, C. T. Perretti, H. Ye, and G. Sugihara. 2013. Predicting climate effects on Pacific sardine. Proceedings of the National Academy of Sciences USA 110:6430–6435.

Di Lorenzo, E., and M. D. Ohman. 2013. A double-integration hypothesis to explain ocean ecosystem response to climate forcing. Proceedings of the National Academy of Sciences USA 110: 2496–2499.

Ditlevsen, P. D., and S. J. Johnsen. 2010. Tipping points: early warning and wishful thinking. Geophysical Research Letters 37:L19703.

D'Odorico, P., L. Ridolfi, and F. Laio. 2013. Precursors of state transitions in stochastic systems with delay. Theoretical Ecology 6:265–270.

Drake, J. M., and B. D. Griffen. 2010. Early warning signals of extinction in deteriorating environments. Nature 467:456–459.

Drinkwater, K. F. 2005. The response of Atlantic cod (*Gadus morhua*) to future climate change. ICES Journal of Marine Science 62:1327–1337.

D'Souza, K., B. I. Epureanu, and M. Pascual. 2015. Forecasting bifurcations from large perturbation recoveries in feedback ecosystems. PLoS ONE 10: e0137779.

Dudgeon, S. R., R. B. Aronson, J. F. Bruno, and W. F. Precht. 2010. Phase shifts and stable states on coral reefs. Marine Ecology Progress Series 413:201–216.

Faassen, E. J., A. J. Veraart, E. H. V. Nes, V. Dakos, M. Lürling, and M. Scheffer. 2015. Hysteresis in an experimental phytoplankton population. Oikos 124:1617–1623.

Frank, K. T., B. Petrie, J. A. D. Fisher, and W. C. Leggett. 2011. Transient dynamics of an altered large marine ecosystem. Nature 477:86–89.

Frank, K. T., B. Petrie, N. L. Shackell, and J. S. Choi. 2006. Reconciling differences in trophic control in mid-latitude marine ecosystems. Ecology Letters 9:1096–1105.

Fraterrigo, J. M., and J. A. Rusak. 2008. Disturbance-driven changes in the variability of ecological patterns and processes. Ecology Letters 11:756–770.

Fung, T., R. M. Seymour, and C. R. Johnson. 2013. Warning signals of regime shifts as intrinsic properties of endogenous dynamics. American Naturalist 182:208–222.

Glaser, S. M., M. J. Fogarty, H. Liu, I. Altman, C. H. Hsieh, L. Kaufman, A. D. MacCall, A. A. Rosenberg, H. Ye, and G. Sugihara. 2014. Complex dynamics may limit prediction in marine fisheries. Fish and Fisheries 15:616–633.

Guttal, V., and C. Jayaprakash. 2008. Changing skewness: an early warning signal of regime shifts in ecosystems. Ecology Letters 11:450–460.

Guttal, V., C. Jayaprakash, and O. P. Tabbaa. 2013. Robustness of early warning signals of regime shifts in time-delayed ecological models. Theoretical Ecology 6:271–283.

Hastings, A., and D. B. Wysham. 2010. Regime shifts in ecological systems can occur with no warning. Ecology Letters 13:464–472.

Hewitt, J. E., and S. F. Thrush. 2010. Empirical evidence of an approaching alternate state produced by intrinsic community dynamics, climatic variability and management actions. Marine Ecology Progress Series 413:267–276.

Holling, C. S. 1973. Resilience and stability of ecological systems. Annual Review of Ecology and Systematics 4:1–23.

Hsieh, C. H., C. Anderson, and G. Sugihara. 2008. Extending nonlinear analysis to short ecological time series. American Naturalist 171:71–80.

Hsieh, C. H., S. M. Glaser, A. J. Lucas, and G. Sugihara. 2005. Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. Nature 435:336–340.

Hughes, T. P. 1994. Catastrophes, phase shifts, and large-scale degradation of a Caribbean coral reef. Science 265:1547–1551.

Hunsicker, M. E., C. V. Kappel, K. A. Selkoe, B. S. Halpern, C. Scarborough, L. Mease, and A. Amrhein. 2016. Characterizing driver-response relationships in marine pelagic ecosystems for improved ocean management. Ecological Applications 26:651–663.

Ives, A. R. 1995. Measuring resilience in stochastic systems. Ecological Monographs 65:217–233.

Ives, A. R., B. Dennis, K. L. Cottingham, and S. R. Carpenter. 2003. Estimating community stability and ecological interactions from time series data. Ecological Monographs 73:301–330.

Kéfi, S., V. Dakos, M. Scheffer, E. H. Van Nes, and M. Rietkerk. 2013. Early warning signals also precede non-catastrophic transitions. Oikos 122:641–648.

Kéfi, S., V. Guttal, W. A. Brock, S. R. Carpenter, A. M. Ellison, V. N. Livina, D. A. Seekell, M. Scheffer, E. H. van Nes, and V. Dakos. 2014. Early warning signals of ecological transitions: methods for spatial patterns. PLoS ONE 9:e92097.

Krkošek, M., and J. M. Drake. 2014. On signals of phase transitions in salmon population dynamics. Proceedings of the Royal Society B: Biological Sciences 281:20133221.

Lees, K., S. Pitois, C. Scott, C. Frid, and S. Mackinson. 2006. Characterizing regime shifts in the marine environment. Fish and Fisheries 7:104–127.

Lenton, T. M., V. N. Livina, V. Dakos, and M. Scheffer. 2012. Climate bifurcation during the last deglaciation? Climate of the Past 8:1127–1139.

Lewontin, R. C. 1968. The meaning of stability. Brookhaven Symposia in Biology 22:13–23.

Lindegren, M., V. Dakos, J. P. Groger, A. Gardmark, G. Kornilovs, S. A. Otto, and C. Möllmann. 2012. Early detection of ecosystem regime shifts: a multiple method evaluation for management application. PLoS ONE 7:e38410.

Litzow, M. A., and L. Ciannelli. 2007. Oscillating trophic control induces community reorganization in a marine ecosystem. Ecology Letters 10: 1124–1134.

Litzow, M. A., F. J. Mueter, and J. D. Urban. 2013. Rising catch variability preceded historical fisheries collapses in Alaska. Ecological Applications 23:1475–1487.

Litzow, M. A., J. D. Urban, and B. J. Laurel. 2008. Increased spatial variance accompanies reorganization of two continental shelf ecosystems. Ecological Applications 18:1331–1337.

Lotka, A. J. 1956. Elements of mathematical biology. Dover Publications, New York, New York, USA.

Lumi, N., K. Laas, and R. Mankin. 2015. Rising relative fluctuation as a warning indicator of discontinuous transitions in symbiotic metapopulations. Physica A: Statistical Mechanics and Its Applications 437:109–118.

Manly, B. F. J. 2006. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, New York, New York, USA.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. Bulletin of the American Meteorological Society 78:1069–1079.

May, R. M. 1977. Thresholds and breakpoints in ecosystems with a multiplicity of stable states. Nature 269:471–477.

Mueter, F. J., and M. A. Litzow. 2008. Sea ice retreat alters the biogeography of the Bering Sea continental shelf. Ecological Applications 18:309–320.

Mumby, P. J., and R. S. Steneck. 2008. Coral reef management and conservation in light of rapidly evolving ecological paradigms. Trends in Ecology and Evolution 23:555–563.

Mumby, P. J., et al. 2006. Fishing, trophic cascades, and the process of grazing on coral reefs. Science 311:98–101.

Norström, A. V., M. Norström, J. Lokrantz, and C. Folke. 2009. Alternative states on coral reefs: beyond coral–macroalgal phase shifts. Marine Ecology Progress Series 376:295–306.

Pace, M. L., S. R. Carpenter, and J. J. Cole. 2015. With and without warning: managing ecosystems in a changing world. Frontiers in Ecology and the Environment 13:460–467.

Pace, M. L., S. R. Carpenter, R. A. Johnson, and J. T. Kurtzweil. 2013. Zooplankton provide early warnings of a regime shift in a whole lake manipulation. Limnology and Oceanography 58:525–532.

Petraitis, P. S., and S. R. Dudgeon. 2016. Cusps and butterflies: multiple stable states in marine systems as catastrophes. Marine and Freshwater Research 67:37–46.

Pinto, P. D., and I. O'Farrell. 2014. Regime shifts between free-floating plants and phytoplankton: a review. Hydrobiologia 740:13–24.

R Core Team. 2016. R: a language and environment for statistical computing. Version 3.2.4. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org

Robinson, C. T., and U. Uehlinger. 2008. Experimental floods cause ecosystem regime shift. Ecological Applications 18:511–526.

Rudnick, D. L., and R. E. Davis. 2003. Red noise and regime shifts. Deep-Sea Research Part I: Oceanographic Research Papers 50:691–699.

Samhouri, J. F., P. S. Levin, and C. H. Ainsworth. 2010. Identifying thresholds for ecosystem-based management. PLoS ONE 5:e8907.

Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara. 2009. Early-warning signals for critical transitions. Nature 461:53–59.

Scheffer, M., and S. Carpenter. 2003. Catastrophic regime shifts in ecosystems: linking theory to observations. Trends in Ecology and Evolution 18:648–656.

Scheffer, M., S. R. Carpenter, V. Dakos, and E. H. van Nes. 2015. Generic indicators of ecological resilience: inferring the chance of a critical transition. Annual Review of Ecology, Evolution and Systematics 46:145–167.

Scheffer, M., S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature 413:591–596.

Scheffer, M., S. H. Hosper, M. L. Meijer, B. Moss, and E. Jeppesen. 1993. Alternative equilibria in shallow lakes. Trends in Ecology and Evolution 8:275–279.

Schröder, A., L. Persson, and A. M. De Roos. 2005. Direct experimental evidence for alternative stable states: a review. Oikos 110:3–19.

Seekell, D. A., S. R. Carpenter, T. J. Cline, and M. L. Pace. 2012. Conditional heteroskedasticity forecasts regime shift in a whole-ecosystem experiment. Ecosystems 15:741–747.

Seekell, D. A., S. R. Carpenter, and M. L. Pace. 2011. Conditional heteroscedasticity as a leading indicator of ecological regime shifts. American Naturalist 178:442–451.

Sokal, R. R., and F. J. Rohlf. 1995. Biometry: the principles and practice of statistics in biological research. Third edition. Freeman, New York, New York, USA.

Spanbauer, T. L., C. R. Allen, D. G. Angeler, T. Eason, S. C. Fritz, A. S. Garmestani, K. L. Nash, and J. R. Stone. 2014. Prolonged instability prior to a regime shift. PLoS ONE 9:e108936.

van Nes, E. H., and M. Scheffer. 2003. Alternative attractors may boost uncertainty and sensitivity in ecological models. Ecological Modelling 159: 117–124.

van Nes, E. H., and M. Scheffer. 2007. Slow recovery from perturbations as a generic indicator of a nearby catastrophic shift. American Naturalist 169:738–747.

Veraart, A. J., E. J. Faassen, V. Dakos, E. H. van Nes, M. Lürling, and M. Scheffer. 2012. Recovery rates reflect distance to a tipping point in a living system. Nature 481:357–359.

Wang, R., J. A. Dearing, P. G. Langdon, E. Zhang, X. Yang, V. Dakos, and M. Scheffer. 2012. Flickering gives early warning signals of a critical transition to a eutrophic lake state. Nature 492:419–422.

Wissel, C. 1984. A universal law of the characteristic return time near thresholds. Oecologia 65:101–107.

Wouters, N., V. Dakos, M. Edwards, M. P. Serafim, P. J. Valayer, and H. N. Cabral. 2015. Evidencing a regime shift in the North Sea using early-warning signals as indicators of critical transitions. Estuarine, Coastal and Shelf Science 152:65–72.

Xu, C., E. H. Van Nes, M. Holmgren, S. Kéfi, and M. Scheffer. 2015. Local facilitation may cause tipping points on a landscape level preceded by early-warning indicators. American Naturalist 186: E81–E90.

Supporting Information

Additional Supporting Information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/ecs2.1614/full