**The Regime Shift Detector: a model to identify changes in dynamic rules governing populations**

C.A. Bahlai1,2 And E.F. Zipkin2

1 Department of Biological Sciences, Kent State University. Kent, Ohio, United States of America. Email for correspondence: [cbahlai@kent.edu](mailto:cbahlai@kent.edu)

2 Department of Integrative Biology; Program in Ecology, Evolutionary Biology and Behavior, Michigan State University, East Lansing, Michigan, United States of America.

**Abstract**

1: Environmental factors may interact with internal rules of population regulation, sometimes perturbing systems to alternate dynamic states. Yet, pinpointing when sustained state changes occur in naturally fluctuating populations is difficult. An analytical approach that can identify the timing and magnitude of such changes, or “regime shifts”, would facilitate understanding of abrupt ecological transitions with potential to inform management of species of economic or conservation concern.

2: We develop a generalizable model, the “Regime Shift Detector” for detecting changes in the parameters of a simple density dependent model governing the temporal fluctuations of a population. The model is a generalizable tool comprised of a suite of functions for examining population time series data for the presence, location, and magnitude of shifts. The model uses an iterative approach to fitting subsets of time series data, and then ranks the fit of a particular break point combination using model selection, and a separate tool calculates a relative weight of each break, aiding in interpretation of more subtle patterns. We examined the performance of this model with simulated data and two real-world case studies each with >20-years of data: one species with an invasion dynamic, the multicolored Asian ladybeetle (*Harmonia axyridis*), and one that has been declining, the eastern monarch butterfly (*Danaus plexippus*).

3: We evaluated the regime shift detector model from two perspectives: in the context of the set of equivalently performing break point combinations, and in the context of the top ranked break point combination alone. We found that under low environmental/sampling error, the break point sets selected by the regime shift detector contained the simulation parameters of with 70%-100 accuracy. We found that the regime shift detector model was most accurate for more complex models when considered from the perspective of top-ranked break point combinations alone: top ranked break point combinations were most likely to identify additional potential breaks in the time series data. The weighting tool generally separated breaks intentionally placed in simulated data to those due to sampling error, although the magnitude of the difference varied with other simulation parameters.

4: In our case study examining the invasion process of Multicolored Asian Ladybeetle (*Harmonia axyridis*), the regime shift detector identified shifts in population cycling associated with known variation in prey availability. However, the results for the eastern monarch population were more ambiguous, within greater uncertainty about the number and location of breaks in the time series data.

5: When interpreted in the context of known species biology, the regime shift detector has the potential to aid management decisions and identify critical drivers of change in species’ dynamics. In an era of rapid global change, such tools can increase understanding of the conditions under which population dynamics can shift to other states.

**Introduction**

Population dynamics are governed by internal, biotic rules and also abiotic factors, often leading to both stochastic and deterministic forces governing abundance patterns (Bjørnstad and Grenfell 2001). External perturbations to population processes can lead to regime shifts, where the internal rules that govern fluctuations in population abundance transition to another state (Hare and Mantua 2000, Carpenter et al. 2008). Understanding how and when external environmental factors interact with internal density dependent regulation remains a fundamental challenge in population ecology (Sutherland et al. 2013, Eason et al. 2016).

Density dependent tools for modelling population time series were developed and championed during the 1950s and 60s. Examples include the Ricker and Beverton-Holt models, which were initially developed for fisheries management (Ricker 1954, Beverton and Holt 1957) and describe the expected population size in year *t* as a function of the population size in year *t-1*. The accuracy of these simple density dependent models is generally highest for populations fluctuating around their carrying capacity (Sabo et al. 2004), with compensatory density dependence, and a short development period to minimize the effect of stage-structured lags (Bjørnstad and Grenfell 2001). Although this deterministic approach to population modelling has largely fallen out of favor for more complex strategies involving nonlinear stochastic elements (May 1976, Bjørnstad and Grenfell 2001, Barraquand et al. 2017, Boettiger 2018), these models remain useful, in large part due to their simplicity and ecologically meaningful interpretations (Gadrich and Katriel 2016). Differences in parameter values between populations, or a change within a single population, can suggest differences in environmental constraints governing a population, providing a quantitative measure of environmental change impacts (Forchhammer and Asferg 2000, Berryman and Lima 2006, Zipkin et al. 2009, Bahlai, vander Werf, et al. 2015).

Although theoretically well-described, identifying abrupt transitions in ecological systems is challenging using real-world data due to a lack of systematic approaches (Bestelmeyer et al. 2011). In many cases, transition points are applied to time series data *ad hoc*, based on data visualization or specific hypotheses surrounding factors affecting population fluctuations (Hare and Mantua 2000, Toms and Lesperance 2003, Weimerskirch et al. 2003, Berryman and Lima 2006, Knapp et al. 2012), creating the potential for biases in selecting break points. Break point analysis tools eliminate this bias by locating change points with a variety of optimization strategies, including linear and moving average methods (Braun and Muller 1998, Zeileis et al. 2001, Killick and Eckley 2014, Priyadarshana and Sofronov 2015). However, these methods do not work on data with internal, density dependent structure inherent to population time series data. Density-dependent population growth has the potential to mask transition points. For example, transient dynamics occurring immediately after a temporary disturbance can result in a change in population size, but not necessarily in the rules governing population fluctuations. Wavelet analysis has been applied to population time series to address changes in cycling patterns (Jenouvrier et al. 2005) but this method also does not account for density-dependence processes as an explicit mechanism governing changes in abundance (Cazelles et al. 2008). A robust, unbiased tool for identifying regime shifts would simultaneously allow us to identify when shifts in population cycling processes occur and further quantify the specific changes to the underlying dynamic driving population size changes.

In this paper, we develop a generalizable tool for detecting shifts in dynamic regimes within density-dependent populations using time series data. We illustrate our approach using the Ricker model because of its simplicity and high performance under a variety of realistic environmental scenarios. Yet, the techniques described in our paper can be readily adapted for other model structures, including more complicated processes such as seasonal periodicity or lag effects. We describe the basic structure of our model and how it can be used to evaluate the presence, location, and magnitude of shifts in the population parameters governing population dynamics. We demonstrate the utility our model through a series of simulations and apply the model to real-world case studies of two populations of conservation and economic concern. First, we examine the invasion process of the multicolored Asian ladybeetle (*Harmonia axyridis*), a cosmopolitan invasive, in the two decades following its arrival in Midwestern US agricultural ecosystems. Then, we examine the declining eastern monarch butterfly (*Danaus plexippus*) population using census data collected on its overwintering grounds in Mexico over a similar two decade period.

**The Regime Shift Detector model**

We assume that the population of interest follows a Ricker model structure such that the population size in time t+1, , is dependent on the population size in time t, , the carrying capacity of the system, , and the per-capita annual growth rate, (Turchin 2003):

.

We further assume that observed annual population abundance is partially stochastic and may be influenced by either environmental variation and/or sampling error. As such, we include an error term to represent process noise, which follows a normal distribution centered around zero with a constant variance of . The parameters K, r, and are estimated from the population timer series data (N1, N2, … N*t*). The Ricker model is a useful starting point for breakpoint analyses because 1) it does not rely on any external information (other than the previous year’s population size; 2) only three parameters need to be estimated, and those parameters have ecologically meaningful interpretations; and 3) it is an extremely flexible distribution, taking a variety of forms, from linear to compensatory to over-compensatory, and thus has a wide range of applications across a wide variety of taxa (Ricker 1954, Brook and Bradshaw 2006).

To build the Regime Shift Detector, we use an iterative, model-selection based process to determine if, and when, shifts in parameter values occurred within a given time series. To achieve this, we first fit the Ricker model to the entire time series of available data. Then the population time series is subdivided into all possible combinations of 2, 3, …, n subsets of sequential data points (hereafter, ‘break point combination’) and the Ricker model is fitted to each of the subsets produced for each break point combination. We constrain break point combinations to only include subsets with at least four sequential data points to avoid over-fitting.

After fitting each subset for a given break point combination, we evaluated the model’s Akaike Information Criteria for small sample sizes (AICc) by calculating AICc values for each segment and sum them together (Hall et al. 2013). AICc values were used to rank fits for each break point combination, and fits for break point combinations with lower AICc values were considered to have better performance. When AICc values differed by two units or less, models were considered equivalent (Burnham and Anderson 2002). To evaluate the strength of evidence associated with a given break in a time series identified by the regime shift detector, we developed an extension of Burnham and Anderson’s (2002) Relative Variable Importance method. To accomplish this, we computed the Akaike weight *w*i for each break point combination, and, for each potential break point, summed the Akaike weights across all break point combinations in which that break point appeared, allowing us to compute a relative ‘break weight.’

The ‘regime shift detector’ is implemented as a series of R functions to enable a user to quickly generate a report on the fit of the model and the location of break points with their own data. Users provide the model with a data frame of population observations at standard time intervals. The model (and all subsequent simulations and case studies) were scripted and run in R Version 3.3.3 (R Development Core Team 2017) . All data manipulations, analyses and figure scripts, including the complete development history, are publicly available in a Github repository at <https://github.com/cbahlai/monarch_regime>. We summarize the role of each function in the supplemental files (see S1), but readers can download the script file directly for implementation details.

*Simulation study*

We conducted a series of simulations to test the accuracy of the regime shift detector under a variety of scenarios. For a given time series, the regime shift detector produces both a top break point combination, and a set of equally performing break point combinations for cases in which model fits produce equivalent AICc values (i.e. within 2 units of the best-performing fit). We selected AICc as our information criterion to use for model selection within the regime shift detector because it provided a balance of simplicity and sensitivity, however, we also completed a parallel analysis with an identical procedure, using AIC as the information criterion for decision-making, which is documented in appendix S2. We tested the performance of the regime shift detector model under a variety of plausible parameter spaces. This was achieved by altering the Ricker parameters to simulate changes in dynamic rules. First, we selected a starting a base scenario with a starting population *Nt* of 3000, and a starting values of *r* and *K* as 2 and 2000, and a duration of 20 years, to start the scenario in a known, stable dynamic. We then modified *r* and *K* individually by a set of intervals from their starting values (0%, 10%, 25%, 50%, 75%) while holding all other parameters constant. We examined the effect of the size of *r* on model performance by creating scenerios with different starting values of *r* (0.5, 1, 1.5, 2). We also evaluated how the magnitude of stochasticity in the system (as measured by the error term ) influenced model performance. For generalizability of our simulation results, we simulated error as a percentage of the mean population size, rather than as absolute value (described in the model above). For each (true) annual population size in the simulated dataset, a random value was selected from a normal curve of mean 0 and standard deviation of (where = 1%, 2%, 5%, 10%, 15%)*.* To scale the standard deviation to the population size used in the simulations, and the expected value of N(t+1) was multiplied by this value to simulate error σ. Each of these scenarios was run on 250 simulated datasets for breakpoint combinations with 0, 1, 2 and 3, respectively. We also tested the impact of time series length by modifying length of the simulated time series at five year intervals (over a range from 15 – 30 years) while holding other parameters constant.

We examined the regime shift detector’s performance for all test scenarios outlined above from two perspectives. First, we evaluated the ability of the model to detect scenario initialization conditions within the set of equivalent break point combinations (Fig 1). Then, we examined the performance of the singular top-ranked break point combination in greater depth (Fig. 2). We also examined the performance of the break-point weighting tool from the perspective of its average and ‘worst-case’ weightings of correct and erroneous break points (Figs. 3, 4).

In general, initiation conditions were detected within the equivalently performing break point combination sets with >70% accuracy under nearly all parameterization scenarios (Fig. 1). However, the model’s absolute top-ranked break point combination reflected parameterization conditions much more inconsistently. The regime shift detector correctly identified initiation break point combinations within its equivalent break point set in virtually 100% of scenarios initiated with two break points, >90% of scenarios initiated with one break point and >80% of zero break simulations, and this performance remained roughly consistent with varied experimental noise(Fig. 1 A). Top ranked break point combinations were most likely to detect more complex initiation conditions (i.e. more break points in the simulated data), but the accuracy was generally quite low, particularly for zero and one break point scenarios. A very similar trend of model performance was observed over varying starting values of *r*, with roughly consistent performance of the regime shift detector model amongst the equivalent perfoming break point combinations, regardless of *r*, and minor variation in performance amongst the top ranked break point combination, with accuracy reduced at extreme values of *r* ( Fig. 1 B). The regime shift detector model performance was reduced with smaller shifts in K (<10%; Fig 1. C) and larger changes to *r* (>25%; Fig. 1 D. This result is somewhat counter-intuitive, as we would generally expect larger shift in parameters to be more easily detected. However, because the Ricker model produces chaotic dynamics with high values of r (r > 2.3) and has the potential for large deviations from equilibrium with high K values, a large shift in parameters could potentially result in a situation where multiple break point fits would perform equally well. Finally, the likelihood that the regime shift detector accurately identified scenario conditions decreased as scenario length increased, likely because of the factorial increase in potential break point combinations with additional data in the time series (Fig. 1 E).

When we examined regime shift detector’s top ranked break point combinations in more depth, we found that the majority of the errors in detecting the initiation conditions could be attributed to finding additional, erroneous breaks, particularly in scenarios initiated with one or two breaks (Fig. 2). We found the top ranked break point’s combination varied in performance similarly to that of the model set, but in more pronounced ways to extreme values of parameterization: model performance decreased with increasing experimental noise (Fig 2 A), decreased slightly at low and high starting values of *r* (Fig. 2 B), increased with increasing change in *K* (Fig. 2 C), decreased with increasing change in *r* (Fig 2 D). However, unlike the equivalently performing break point sets, top models containing the correct parameterization, plus additional break points remained roughly stable with increasing time series length (Fig. 2 E).

We subjected the simulation data to a weighting analysis under the same conditions, providing a relative weight for each prospective break point identified by the regime shift detector model. We found that in the vast majority of parameterization cases, the average weight of a ‘true’ break (i.e. one that was intentionally simulated in the data) typically exceeded a value of 0.8 (Fig. 3), whereas the weight of erroneous breaks averaged less than 0.2 in weight. The notable exception occurs when ‘true’ breaks correspond to very small shifts in K (Fig. 3 C). We also examined the break weights from a ‘worst-case’ perspective: i.e. under idealized conditions, find the minimum weights we observe for ‘true’ break points, and simultaneously, the maximum weights we’d expect to observe for erroneous break points (Fig. 4). In general, even in these conditions, the lowest weight observed for a ‘true’ break was generally higher than the highest weight observed for an erroneous break, but the division between the two became less clear in higher complexity scenarios with more break points (Fig. 4). Increasing experimental noise (Fig. 4A), smaller shifts in K (Fig. 4 C), larger shifts in r (Fig. 4 D), and longer time series (Fig. 4 E) were all associated with less sensitivity in discerning between the true and erroneous break point signal.

**Applications**

We tested the performance of the regime shift detector model with two cases using population time series data from field observation studies. Both case studies involve approximately two decades of observations of economically or culturally important insect species: one case examines an invasion process and the other examines a population decline, both occurring over the same time period in recent history.

Multicolored Asian ladybeetles in southwestern Michigan

The 1994 invasion of multicolored Asian ladybeetles to southwestern Michigan, United States was documented in monitoring data collected on agriculturally-important Coccinellidae (ladybeetles) in landscapes dominated by field crops. Population density of ladybeetles is monitored in 10 plant communities weekly over the growing season using yellow sticky card glue traps starting in 1989 at the Kellogg Biological Station at Michigan State University. We used data on the captures of adults at the site from 1994-2017, culled at day of year 222 (August 10) to minimize the effect of year-to-year variation in the sampling period. We then calculated the average number of adults captured per trap, across all traps deployed within a sampling year, and used this value as the dependent variable in our analysis. Detailed sampling methodology is available in previous work (Bahlai et al. 2013, Bahlai, Colunga-Garcia, et al. 2015, Bahlai, vander Werf, et al. 2015).

Two break points, one occurring after 2000 and one occurring after 2005, were observed in the top break point combination model (Fig. 5 A, AICc=-18.02). However, the regime shift detector indicated that two additional break point combinations, a single break after 2000 (AICc=-17.46), and a no break series (AICc=-17.64), had equivalent performance. Break weight analysis suggested a weight of 0.56 for the 2000 break, and a weight of 0.29 for the break after 2005. Although the weights of both of these break points fall into a range intermediate to what we observed between ‘true’ and ‘erroneous’ breaks observed from our simulations, we expect that natural systems would behave in a less idealized manner, thus we conclude that there is reasonably strong evidence of a shift in dynamic rule after 2000, and a moderate-weak evidence for a shift after 2005. The shift from ‘phase A’ to ‘phase B’ in 2000 is characterized by substantial increases in the fitted values for K and r, with approximate increases of 75% and 40% over their initial estimates, respectively (Table 1). The shift from 'phase B' to 'phase C' in 2005 is characterized by a return to parameter estimates that were nearly identical to those observed for ‘phase A’ (Table 1, Fig. 5 B).

These observations can be explained in the context of the known ecology of this ladybeetle. Dynamics of the ladybeetle invasion appear to be closely coupled with prey availability (Bahlai and Sears 2009, Heimpel et al. 2010, Rhainds et al. 2010, Bahlai, Colunga-Garcia, et al. 2015), which, in turn, is driven by documented pest management practices (neonicotinoid insecticide use; Bahlai, vander Werf, et al. 2015) leading to relatively simple pulsed changes in dynamics. The first shift in the dynamics of the Asian ladybeetle*,* after 2000, corresponds to the well documented arrival and establishment of a preferred prey item from its native range, soybean aphid, to North America (Ragsdale et al. 2004, Wu et al. 2004). The invasion of this aphid dramatically increased resources available to *H.axyridis* in habitats where the beetles were already well-established (Bahlai et al. 2013), supporting both a higher carrying capacity and a greater intrinsic growth rate. The second shift, after 2005, was weaker is more difficult to directly explain, as the prey item still persists in the landscapes where *H. axyridis* occurs. However, landscape-scale use of neonicotinoid insecticides decreased prey numbers, particularly during the spring when aphids are colonizing new hosts, which could be limiting early season reproduction of ladybeetles(Bahlai, vander Werf, et al. 2015)*.* Indeed, in this case, we would expect a weaker shift in dynamics as the prey item is incompletely controlled, and control tactics were not uniformly adopted across the prey’s range all at one time.

Monarch butterflies in Mexican overwintering grounds

The eastern population of the North American monarch butterfly (*Danaus plexippus*) is migratory, with the majority of individuals overwintering in large aggregations in Oyamel fir forests within the transvolcanic mountains in the central region of Mexico (Urquhart and Urquhart 1978). Monarchs are highly dispersed over their breeding season, occupying landscapes throughout the agricultural belt in central and eastern United States and southern Canada (Flockhart et al. 2017). As such, estimates of the overwintering population size can provide a convenient and inclusive annual metric of the size of the eastern migratory population (Prysby and Oberhauser 2004). This population of monarchs has been in dramatic decline in recent decades, although the degree and cause of this decline is hotly debated (Sarkar 2017). We used data on the total area occupied by monarchs from 1995 to 2017 (based on early winter surveys conducted in December), compiled by the World Wildlife Fund Mexico (available at MonarchWatch; Lovett 2017).

The regime shift detector estimated that the best break point combination fit for the monarch overwintering density data was a single break after 2003 (Fig. 6; AICc=120.18). However, the regime shift detector indicated that two additional break point combinations, a single break after 2006 (AICc=121.87), and a two break combination of 2003 and 2008 (AICc=-121.86), had equivalent performance. The weight analysis suggested weights of 0.49, 0.14 and 0.26, for 2003, 2006 and 2008 respectively. The weight analysis also indicated that there was a weight of 0.12 associated with a break at 2007. As above, with our ladybeetle case study, the strength of evidence was strongest for the first break in 2003, and weaker for the secondary break. The shift corresponds with a >50% reduction in K at 2003, and, if the secondary break is taken at 2008, a further reduction of K nearing 50% again at that point (Table 2; Fig. 6 B). However, model selection results were somewhat ambiguous: this secondary break did not appear in the top selected model, and the set of equivalently performing models contained contradictory structures, that is, they did not all select overlapping break point combinations. Yet, this apparent contradiction is not unexpected in the context of the biology of this species.

Abiotic drivers of monarch population dynamics are complex and can interact at local, regional, and continental scales (Saunders et al. 2017). Loss of milkweed hostplants due to changing agricultural practices on Midwestern breeding grounds (Hartzler 2010, Pleasants and Oberhauser 2013) is hypothesized to be a major cause of the decline. However, other studies have implicated climate (Zipkin et al. 2012), extreme weather events (Brower et al. 2004), changing habitat availability on wintering grounds (Vidal and Rendón-Salinas 2014), and mortality during the fall migration (Imanine et al. 2016?) as possible factors influencing monarch population dynamics. With many super-imposed drivers, monarch dynamics are likely to be driven by both press and pulsed processes, making the detection of discrete break points associated with regime shifts more difficult.

Thus, there is biological support for essentially all the break points selected by our model- or, indeed, changes occurring in a combination of pulsed and smooth processes. For example, changing herbicide use practices in central North America have largely eliminated milkweed hostplants from agricultural field crops, with fairly consistent, low levels of milkweed on the landscape starting from about 2003-2005 (Zaya et al. 2017). Although glyphosate tolerant soybeans and maize were introduced to the US market in 1996 and 1998 respectively (Duke and Powles 2009), actual glyphosate use lagged behind, with dramatic increases in use of the pesticide in 1998- 2003 in soybean, and 2007-2008 in maize (Baker 2017).

**Discussion**

The regime shift detector provides a novel and objective tool for evaluating changes in parameters, such as carrying capacity and growth rates, that govern natural populations. The regime shift detector allows for comparisons of population processes over time and between ecosystems. As illustrated with our case studies, model results not only identify and quantify parameters changes, but can provide insight into the potential drivers in the systems under study and help focus the time frames that should be examined more closely. The regime shift detector, as describes here, uses the Ricker model as its central structure. However, the central dynamic model could be changed or modified to incorporated more complicated population processes as well as covariates (for example, if a population had a known response to temperature or another environmental variable).

The model in the regime shift detector is a compromise between sensitivity and simplicity. We found that, in general, the model results were more easily interpreted when using a more conservative information criterion to select results in simulations, and thus, we recommend users of the regime shift detector model use AICc when examining new data. However, there may be cases where it is desirable to gain a more liberal estimate of changes in patterns: in this case, the more sensitive AIC can be used to rank break point combinations. We found final results of the regime shift detector model were similar, regardless of the information criterion used, except that using AIC was more likely to indicate more candidate breaks, and weight both true and erroneous breaks with higher values (Appendix S2). Thus, this more sensitive approach may be most useful in the context of hypothesis generation, rather than as an explicit hypothesis test.

Additionally, we recommend users carefully consider the limitations of the tool in the context of the raw data presented: if phases of change are close together (e.g., less than 3-4 years or time periods), the resultant variation may reduce the model’s sensitivity for detecting shifts in the data. For instance, we use the discrete form of the Ricker model, with a yearly time step, but end users can modify model structure and temporal resolution appropriate to their systems and available data. As is, the results of the model should be interpreted with caution in populations undergoing rapid change in their environments or internal dynamics.

Our results clearly indicate the examination of any single statistic or test is insufficient to determine which break points, and in which combination, represent ‘truth’ for a given data set. However, we found that taken in combination, our set of tools provides a means to identify candidate breaks and evaluate their relative strength of evidence. For instance, we found that, amongst top-ranked break point combinations, the regime shift detector’s most common error was to over-estimate, i.e. to ‘find’ breakpoints where they did not occur (Fig. 2). Amongst top selected break point combinations, the regime shift detector model found extra breaks >50% of the time in scenarios parameterized to have one or two breaks, and zero-break parameterizations were very rarely ranked highest. However, the exact parameterization of the scenario was listed in the equivalent break point combination set >75% of the time, across almost all parameterization scenarios. Additionally, the equivalent set approach was superior in detecting scenarios which had been initiated with no break points at all: for these scenarios, a break point combination with no peaks listed among the equivalently performing set more than 80% of the time (Fig. 1)

In both of these evaluations, we found that the regime shift detector model was able to detect scenario conditions in practically all cases, but presented a problem with respect to over-sensitivity: essentially, more ‘suspected’ break points were identified by the model than were intentionally placed in the simulated data, creating ambiguities in interpretation. Thus, both the ‘model set’ and the ‘top model’ approach could be used to identify a list of potential break points and break point combinations, and an additional tool could be used to evaluate the strength of evidence for each. In our simulations, we found that, on average, ‘true’ breaks based on scenario conditions and erroneous breaks could be easily discerned by their computed weights, while erroneous breaks typically averaged less than 0.2 in weight, whereas in most scenarios, true breaks had weights of >0.8 (Fig. 3). Evaluating these weights from a ‘worst case’ perspective, where the minimum weight ‘true’ break was compared to the maximum weight ‘erroneous’ break observed for a given scenario, however, yielded more ambiguity: in some simulations and for certain parameterizations, the weight values of true and erroneous breaks overlapped. (Fig. 4). However, even in this case, the majority of the ‘true’ breaks had higher weights than the erroneous ones under most parameterization conditions.

Applying these insights to our case studies upon evaluating our case studies, we found interpretation of the ladybeetle example was relatively straightforward (Fig. 5). Our top break point combination and the equivalently-performing set did not contain contradictory information: each candidate set was simply a subset of breakpoints from the most complex set, and only two break points were found. Both of these break points were associated with moderate or greater weights. The monarch butterfly case study was more ambiguous (Fig. 6). Although the top two break point weights were nearly identical to those observed in the ladybeetle study, the top-ranked break point combination only had one break, and the set of equivalently performing break point combinations had varied structures. We repeated these analyses using AIC to examine whether a more liberal approach would clarify results (S2). We found that the AIC analysis reenforced and simplified the results observed in the ladybeetle case, but introduced more ambiguity (by suggesting even more weakly supported candidate break points).

We used simulations to understand how changing various inputs affected the likelihood of the regime shift detector in identifying the conditions under which the data were produced. Performance of the regime shift detector model declined with increasing levels of stochasticity, or error, in the data, particularly in regard to the performance of the top-selected break point combination (Fig. 2 A), a behavior that is expected of any statistical tool. Yet, performance was relatively stable amongst model sets across the range of noise tested in our experiment (Fig. 1 A), suggesting that if a user has reason to believe their data has a higher degree of noise, relying on the model set to identify candidate breaks in the dats, and suggest overall break point structure, may be more reliable than examining the top break point combination exclusively. Thus, whenever possible, the sampling error of the data should be estimated to help evaluate the model’s results in the context of variation within the data due to sampling error.

Other input conditions also impacted the performance of the regime shift detector model. The effect of dynamic shift size was dependent on which parameter was changed and by how much (Figs. 1, 2 C, D). Larger shifts in *K* were more detectable both by the top break point combination and within the equivalent break point combinations set, but this effect was more pronounced amongst the top break point combination analysis (Fig. 2 C). However, increasing change in *r* made it less likely for the parameterization conditions to be detected, but this effect varied in strength between parameterization complexity and whether the break point set or just the top break point combination was considered (Figs. 1, 2 D). Although larger shifts in regression parameters would, intuitively, lead to a higher likelihood of detection, these larger shifts in *r* would also be more likely to induce large variations in transient dynamics in the years immediately following the shift, potentially making the timing of shifts more difficult to pinpoint. Similarly, longer time series yielded results that were more error prone (Fig. 1-4 E). This likely because, firstly, there were simply more possible break-point combinations for the model to select from, and secondly, because the penalty for increasing parameterization (i.e. AICc) would decrease as sample sizes grew, leading to increasing likelihood of identifying extra breaks.

When interpreted in the context of known species biology, the regime shift detector model has the potential to aide management decisions and identify, and rank critical drivers of change in internal dynamics of a population. Our modeling framework should be useful in this era of rapid global change affecting species dynamics, as it can help evaluate population changes relative to internal regulators.

**References**

**Bahlai, C. A., M. Colunga-Garcia, S. H. Gage, and D. A. Landis**. **2013**. Long term functional dynamics of an aphidophagous coccinellid community are unchanged in response to repeated invasion. PLoS One. 8: e83407.

**Bahlai, C. A., and M. K. Sears**. **2009**. Population dynamics of Harmonia axyridis and Aphis glycines in Niagara Peninsula soybean fields and vineyards. Journal of the Entomological Society of Ontario. 140: 27–39.

**Bahlai, C. A., W. vander Werf, M. O’Neal, L. Hemerik, and D. A. Landis**. **2015**. Shifts in dynamic regime of an invasive lady beetle are linked to the invasion and insecticidal management of its prey. Ecological Applications.

**Bahlai, C., M. Colunga-Garcia, S. Gage, and D. Landis**. **2015**. The role of exotic ladybeetles in the decline of native ladybeetle populations: evidence from long-term monitoring. Biol Invasions. 17: 1005–1024.

**Baker, N. T.** **2017**. Estimated annual agricultural pesticide use by crop group for states of the conterminous United States, 1992-2014. National Water Quality Assessment Program.

**Barraquand, F., S. Louca, K. C. Abbott, C. A. Cobbold, F. Cordoleani, D. L. DeAngelis, B. D. Elderd, J. W. Fox, P. Greenwood, and F. M. Hilker**. **2017**. Moving forward in circles: challenges and opportunities in modelling population cycles. Ecology letters. 20: 1074–1092.

**Berryman, A., and M. Lima**. **2006**. Deciphering the effects of climate on animal populations: diagnostic analysis provides new interpretation of Soay sheep dynamics. The American Naturalist. 168: 784–795.

**Bestelmeyer, B. T., A. M. Ellison, W. R. Fraser, K. B. Gorman, S. J. Holbrook, C. M. Laney, M. D. Ohman, D. P. C. Peters, F. C. Pillsbury, A. Rassweiler, R. J. Schmitt, and S. Sharma**. **2011**. Analysis of abrupt transitions in ecological systems. Ecosphere. 2: art129.

**Beverton, R. J., and S. J. Holt**. **1957**. On the dynamics of exploited fish populations. Springer Science & Business Media.

**Bjørnstad, O. N., and B. T. Grenfell**. **2001**. Noisy Clockwork: Time Series Analysis of Population Fluctuations in Animals. Science. 293: 638.

**Boettiger, C.** **2018**. From noise to knowledge: how randomness generates novel phenomena and reveals information. Ecology Letters. 21: 1255–1267.

**Braun, J. V., and H.-G. Muller**. **1998**. Statistical Methods for DNA Sequence Segmentation. Statistical Science. 13: 142–162.

**Brook, B. W., and C. J. Bradshaw**. **2006**. Strength of evidence for density dependence in abundance time series of 1198 species. Ecology. 87: 1445–1451.

**Brower, L. P., D. R. Kust, E. Rendon-Salinas, E. G. Serrano, K. R. Kust, J. Miller, C. Fernandez del Rey, and K. Pape**. **2004**. Catastrophic winter storm mortality of monarch butterflies in Mexico during January 2002. The Monarch butterfly: biology and conservation. 151–166.

**Burnham, K. P., and D. R. Anderson**. **2002**. Model selection and multimodal inference: a practical information-theoretic approach, 2nd ed. Springer Science + Business Media, LLC, New York.

**Carpenter, S. R., W. A. Brock, J. J. Cole, J. F. Kitchell, and M. L. Pace**. **2008**. Leading indicators of trophic cascades. Ecology Letters. 11: 128–138.

**Cazelles, B., M. Chavez, D. Berteaux, F. Ménard, J. O. Vik, S. Jenouvrier, and N. C. Stenseth**. **2008**. Wavelet analysis of ecological time series. Oecologia. 156: 287–304.

**Duke, S. O., and S. B. Powles**. **2009**. Glyphosate-resistant crops and weeds: now and in the future. AgBioForum. 12: 346–357.

**Eason, T., A. S. Garmestani, C. A. Stow, C. Rojo, M. Alvarez-Cobelas, and H. Cabezas**. **2016**. Managing for resilience: an information theory-based approach to assessing ecosystems. J Appl Ecol. 53: 656–665.

**Flockhart, D. T. T., L. P. Brower, M. I. Ramirez, K. A. Hobson, L. I. Wassenaar, S. Altizer, and D. R. Norris**. **2017**. Regional climate on the breeding grounds predicts variation in the natal origin of monarch butterflies overwintering in Mexico over 38 years. Glob Change Biol. 23: 2565–2576.

**Forchhammer, M. C., and T. Asferg**. **2000**. Invading parasites cause a structural shift in red fox dynamics. Proceedings of the Royal Society of London B: Biological Sciences. 267: 779–786.

**Gadrich, T., and G. Katriel**. **2016**. A Mechanistic Stochastic Ricker Model: Analytical and Numerical Investigations. Int. J. Bifurcation Chaos. 26: 1650067.

**Hall, A. R., D. R. Osborn, and N. Sakkas**. **2013**. Inference on Structural Breaks using Information Criteria. The Manchester School. 81: 54–81.

**Hare, S. R., and N. J. Mantua**. **2000**. Empirical evidence for North Pacific regime shifts in 1977 and 1989. Progress in Oceanography. 47: 103–145.

**Hartzler, R. G.** **2010**. Reduction in common milkweed (Asclepias syriaca) occurrence in Iowa cropland from 1999 to 2009. Crop Protection. 29: 1542–1544.

**Heimpel, G., L. Frelich, D. Landis, K. Hopper, K. Hoelmer, Z. Sezen, M. Asplen, and K. Wu**. **2010**. European buckthorn and Asian soybean aphid as components of an extensive invasional meltdown in North America. Biological Invasions. 12: 2913–2931.

**Jenouvrier, S., H. Weimerskirch, C. Barbraud, Y.-H. Park, and B. Cazelles**. **2005**. Evidence of a shift in the cyclicity of Antarctic seabird dynamics linked to climate. Proceedings of the Royal Society B: Biological Sciences. 272: 887–895.

**Killick, R., and I. Eckley**. **2014**. changepoint: An R package for changepoint analysis. Journal of Statistical Software. 58: 1–19.

**Knapp, A. K., M. D. Smith, S. E. Hobbie, S. L. Collins, T. J. Fahey, G. J. A. Hansen, D. A. Landis, K. J. La Pierre, J. M. Melillo, T. R. Seastedt, G. R. Shaver, and J. R. Webster**. **2012**. Past, present, and future roles of long-term experiments in the LTER Network. Bioscience. 62: 377–389.

**Lovett, J.** **2017**. Monarch Population Status.

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

**Pleasants, J. M., and K. S. Oberhauser**. **2013**. Milkweed loss in agricultural fields because of herbicide use: effect on the monarch butterfly population. Insect Conservation and Diversity. 6: 135–144.

**Priyadarshana, W., and G. Sofronov**. **2015**. Multiple break-points detection in array CGH data via the cross-entropy method. IEEE/ACM Transactions on Computational Biology and Bioinformatics (TCBB). 12: 487–498.

**Prysby, M. D., and K. S. Oberhauser**. **2004**. Temporal and geographic variation in monarch densities: citizen scientists document monarch population patterns. The monarch butterfly: Biology and conservation. 9–20.

**R Development Core Team**. **2017**. R: A Language and Environment for Statistical Computing 3.3.3. R Foundation for Statistical Computing.

**Ragsdale, D. W., D. J. Voegtlin, and R. J. O’Neil**. **2004**. Soybean aphid biology in North America. Annals of the Entomological Society of America. 97: 204–208.

**Rhainds, M., H. J. S. Yoo, P. Kindlmann, D. Voegtlin, D. Castillo, C. Rutledge, C. Sadof, S. Yaninek, and R. J. O’Neil**. **2010**. Two-year oscillation cycle in abundance of soybean aphid in Indiana. Agricultural and Forest Entomology. 12: 251–257.

**Ricker, W. E.** **1954**. Stock and Recruitment. J. Fish. Res. Bd. Can. 11: 559–623.

**Sarkar, S.** **2017**. What Is Threatening Monarchs? BioScience. 67: 1080–1080.

**Saunders, S. P., L. Ries, K. S. Oberhauser, W. E. Thogmartin, and E. F. Zipkin**. **2017**. Local and cross-seasonal associations of climate and land use with abundance of monarch butterflies Danaus plexippus. Ecography. n/a-n/a.

**Sutherland, W. J., R. P. Freckleton, H. C. J. Godfray, S. R. Beissinger, T. Benton, D. D. Cameron, Y. Carmel, D. A. Coomes, T. Coulson, M. C. Emmerson, R. S. Hails, G. C. Hays, D. J. Hodgson, M. J. Hutchings, D. Johnson, J. P. G. Jones, M. J. Keeling, H. Kokko, W. E. Kunin, X. Lambin, O. T. Lewis, Y. Malhi, N. Mieszkowska, E. J. Milner-Gulland, K. Norris, A. B. Phillimore, D. W. Purves, J. M. Reid, D. C. Reuman, K. Thompson, J. M. J. Travis, L. A. Turnbull, D. A. Wardle, and T. Wiegand**. **2013**. Identification of 100 fundamental ecological questions. J Ecol. 101: 58–67.

**Toms, J. D., and M. L. Lesperance**. **2003**. Piecewise regression: a tool for identifying ecological thresholds. Ecology. 84: 2034–2041.

**Turchin, P.** **2003**. Complex population dynamics: a theoretical/empirical synthesis. Princeton University Press.

**Urquhart, F. A., and N. R. Urquhart**. **1978**. Autumnal migration routes of the eastern population of the monarch butterfly (Danaus p. plexippus L.; Danaidae; Lepidoptera) in North America to the overwintering site in the Neovolcanic Plateau of Mexico. Can. J. Zool. 56: 1759–1764.

**Vidal, O., and E. Rendón-Salinas**. **2014**. Dynamics and trends of overwintering colonies of the monarch butterfly in Mexico. Biological Conservation. 180: 165–175.

**Weimerskirch, H., P. Inchausti, C. Guinet, and C. Barbraud**. **2003**. Trends in bird and seal populations as indicators of a system shift in the Southern Ocean. Antarctic Science. 15: 249–256.

**Wu, Z., D. Schenk-Hamlin, W. Zhan, D. W. Ragsdale, and G. E. Heimpel**. **2004**. The soybean aphid in China: a historical review. Annals of the Entomological Society of America. 97: 209–218.

**Zaya, D. N., I. S. Pearse, and G. Spyreas**. **2017**. Long-Term Trends in Midwestern Milkweed Abundances and Their Relevance to Monarch Butterfly Declines. BioScience. 67: 343–356.

**Zeileis, A., F. Leisch, K. Hornik, and C. Kleiber**. **2001**. strucchange. An R package for testing for structural change in linear regression models.

**Zipkin, E. F., C. E. Kraft, E. G. Cooch, and P. J. Sullivan**. **2009**. When can efforts to control nuisance and invasive species backfire? Ecological Applications. 19: 1585–1595.

**Zipkin, E. F., L. Ries, R. Reeves, J. Regetz, and K. S. Oberhauser**. **2012**. Tracking climate impacts on the migratory monarch butterfly. Glob Change Biol. 18: 3039–3049.

**Table 1:** Ricker model fits for each phase between break points resulting from fitting population data of *Harmonia axyridis* from Kellogg Biological Station, 1994-2015. Regression parameters r represent the per capita yearly intrinsic rate of increase and K the carrying capacity, based on population numbers expressed as average number of adult *H. axyridis* captured per trap, per year. The ‘Phase’ column gives a shorthand for referring to the data subsetting structure indicated by the regime shift detector model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data**  **structure** | **Phase** | **Years in subset** | **r ( ± SE)** | **K ( ± SE)** |
| Two breaks | A | 1994-2000 | 1.3 ± 0.3 | 0.31 ± 0.02 |
|  | B | 2001-2005 | 2.3 ± 0.3 | 0.43 ± 0.03 |
|  | C | 2006-2013 | 1.6 ± 0.3 | 0.27 ± 0.03 |
|  |  |  |  |  |

**Table 2.** Ricker model fits for each phase between break points resulting from fitting population data documenting the area occupied by overwintering Monarch butterflies in their winter habitat in the Mexico, 1995-2016. Regression parameters r represent the per capita yearly intrinsic rate of increase and K the carrying capacity, in units of hectares occupied. The ‘Phase’ column gives a shorthand for referring to the data subsetting structure indicated by the regime shift detector model.

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Data structure** | **Phase** | **Years in subset** | **r ( ± SE)** | **K ( ± SE)** |
| Two breaks | A | 1995-2003 | 1.0 ± 0.5 | 10.1 ± 1.9 |
|  | B | 2004-2008 | 1.6 ± 0.2 | 5.6 ± 0.3 |
|  | C | 2009-2016 | 1.2 ± 0.4 | 2.8 ± 0.5 |
|  |  |  |  |  |

**Figure captions**

Figure 1: **Performance the regime shift detector model under varying conditions.** Proportion of results where initial conditions were detected by the top break point combination (circles) or within the equivalent model set (triangles) under A) varied noise (in the form of normally distributed error) B) varied starting values of the r constant, C) varied % changes in the K constant in the Ricker model D) varied % changes in r, the intrinsic rate of increase in the Ricker model and E) simulated time series length. Sets of 0, 1, 2 and 3 break points were randomly generated from within the set of possible values each scenario was iterated 250 times.

Figure 2: **Performance the top ranked break point combination selected by the regime shift detector model under varying conditions.** Proportion of results where initial conditions were detected by the top break point combination (circles) or contained within the parameter set of the top break point combination (i.e. scenario conditions detected, plus additional break points found; triangles) under A) varied noise (in the form of normally distributed error) B) varied starting values of the r constant, C) varied % changes in the K constant in the Ricker model D) varied % changes in r, the intrinsic rate of increase in the Ricker model and E) simulated time series length. Sets of 0, 1, 2 and 3 break points were randomly generated from within the set of possible values each scenario was iterated 250 times.

Figure 3: **Average break weight of break points found under varying parameterization conditions.** Break weight is computed based on a modification of “Relative Variable Importance” formula, where each break point is multiplied by the Akaike weight of each model in which it appears. Average weights of break points identified by the regime shift detector model reflecting true parameterization conditions (circles) or erroneous breaks suggested by the model (triangles) under A) varied noise (in the form of normally distributed error) B) varied starting values of the r constant, C) varied % changes in the K constant in the Ricker model D) varied % changes in r, the intrinsic rate of increase in the Ricker model and E) simulated time series length. Sets of 0, 1, 2 and 3 break points were randomly generated from within the set of possible values each scenario was iterated 250 times.

Figure 4: **‘Worst-case’ break weight of break points found under varying parameterization conditions.** Break weight is computed based on a modification of “Relative Variable Importance” formula, where each break point is multiplied by the Akaike weight of each model in which it appears. Minimum weights of break points identified by the regime shift detector model reflecting true parameterization conditions (circles) and maximum weights of erroneous breaks suggested by the model (triangles) under A) varied noise (in the form of normally distributed error) B) varied starting values of the r constant, C) varied % changes in the K constant in the Ricker model D) varied % changes in r, the intrinsic rate of increase in the Ricker model and E) simulated time series length. Sets of 0, 1, 2 and 3 break points were randomly generated from within the set of possible values each scenario was iterated 250 times.

Figure 5: **Regime shift detector breaks and Ricker model fits for an invasive ladybeetle**. Population data documenting the invasion of *Harmonia axyridis*, a ladybeetle native to eastern Asia, to plots at the Kellogg Biological Station in southwestern Michigan, USA, 1994-2015 A) Time series documenting average number of adults captured, per trap, per year. Vertical blue lines indicate timings of shifts in dynamics, as indicated by the regime shift detector model. B) Ricker fits of phases of population dynamics as indicated by the regime shift detector model.

Figure 6: **Regime shift detector breaks and Ricker model fits for a species of conservation concern.** Population data documents the area occupied by overwintering Monarch butterflies in their winter habitat in the Mexico, 1995-2017. A) Time series documenting raw data of estimated area occupied by overwintering monarchs by year. Vertical blue line indicates timing of a shift in dynamics, as indicated by the regime shift detector model. B) Ricker fits of phases of population dynamics as indicated by the regime shift detector model.