bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Repeated regime shifts in a desert rodent community

Authors and Affiliations

Erica M. Christensen, Department of Wildlife Ecology and Conservation, University of Florida, ericachristensen@ufl.edu

David J. Harris, Department of Wildlife Ecology and Conservation, University of Florida, dave@harris-research.me

S. K. Morgan Ernest, Department of Wildlife Ecology and Conservation, University of Florida, skmorgane@ufl.edu

Short running title: Repeated regime shifts in a rodent community

Keywords: community dynamics, Latent Dirichlet Allocation, regime shift, desert rodents, temporal dynamics, extreme climatic events

Type of article: Letters

Number of words in abstract (150); number of words in main text (excluding abstract, acknowledgements, references, and table and figure legends: 4272)

Number of references (61)

Number of figures (5)

Number of tables (0)

Number of text boxes (0)

Corresponding author:

Erica M. Christensen

110 Newins-Ziegler Hall

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

PO Box 110430

Gainesville, FL 32611

Phone: 352-294-2083

Fax: 352-392-6984

Email: ericachristensen@ufl.edu

Statement of authorship: EMC conceived the original study; DH contributed the LDA methods and developed the change-point model; EMC and SKME interpreted the results, and all authors contributed to writing the manuscript.

Data accessibility statement: We confirm that, should this manuscript be accepted, the data supporting the results will be archived in an appropriate public repository such as Dryad or Figshare and the data DOI will be included at the end of the article.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Abstract

Many communities are experiencing temporal changes in the abundances and/or identities of the species that comprise them. Understanding whether communities are changing gradually or via rapid regime shifts is important because different mechanisms are implicated in different patterns of change. We used a machine learning approach to examine the long-term dynamics of a desert rodent community undergoing compositional change. Our approach detected three regime shifts, where changes in the relative abundances of dominant and rare species occurred, and a period of increased variance in the structure of the community. These events coincided with extreme climate events (two high rain events and two droughts) that caused major declines in rodent abundances, suggesting a possible link between climate and regime shifts at this site. Continued study of the dynamics of community change will provide important information on the processes structuring communities and provide guidance for forecasting how community changes in the future.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Introduction

In many places on the planet, the ecological template of nature is changing: temperature is increasing, nutrient distributions are changing, and habitats are being altered and fragmented (Walther 2010). Because every species has environmental conditions and resources that they prefer or require, these changes will alter the communities of species that live there. However, exactly how the composition of a community will temporally track those changes is less clear. There are theoretical reasons to expect both slow and rapid changes in community composition. Slow changes in communities emerge from either stochastic turnover events (Magurran & Henderson 2003; Orrock & Watling 2010) or because a species' ability to thrive gradually improves or degrades as the environment changes, leading to niche-based turnover events (Beaugrand 2004; Tingley et al. 2009). Rapid changes in communities, or regime shifts, can emerge from intrinsic or extrinsic drivers (Williams et al. 2011). With intrinsic regime shifts, gradual changes in the environment eventually push the ecosystem past a threshold, triggering rapid shifts to a new stable state (Scheffer & Carpenter 2003). Extrinsic regime shifts occur either from niche-based tracking as the environment rapidly shifts from one state to another (e.g., shift in climate from a drier to a wetter state; Rowe 2007; Wirth et al. 2008), or via extreme events, which cause cascading changes in species populations that alter how the community recovers post-disturbance (Smith 2011). Understanding the prevalence of rapid and gradual shifts in community structure is important for understanding the drivers of community reorganization and whether we should expect changes to occur quickly or over the span of years or decades.

While a growing number of studies document the occurrence of regime shifts (Österblom et al. 2007; Isbell et al. 2013; Jochimsen et al. 2013), meta-analyses compiling data from many long-term studies indicate that most communities are changing gradually over time (La Sorte & Boecklen 2005; Dornelas et al. 2014; Magurran et al. 2015). However, determining the long-term dynamics of turnover in species composition is complicated because communities are

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

comprised of many species, making temporal turnover an inherently multivariate problem. Distance metrics are commonly used to reduce this complexity to a single axis, and a linear regression is used to test whether samples are becoming more different as the time between those samples increases (e.g. Fig. 1) (Thibault et al. 2004; Magurran and Henderson 2010; Magurran et al. 2015; McGill et al. 2015). Distance-based metrics and linear quantification of change are useful for determining whether long-term compositional change has occurred, but they are less useful for determining what types of dynamics underlie that long-term signal (Bagchi et al. 2017). As a result, while we know that many communities are changing, our current approaches provide little insight into the nature of this change.

To better document how communities are actually changing, what is required is a more sensitive approach that allows rapid changes in community composition – if they occur -- to be better visualized and quantified. Recently, a machine learning approach was introduced to ecology that is capable of quantifying the dynamics of species composition change across gradients (see Valle et al. 2014). Latent Dirichlet Allocation (LDA) was originally developed as an alternative to cluster analysis for summarizing documents based on the words they contain (Blei et al. 2003). Rather than assigning each document to a single cluster, LDA identifies the proportion of words in each document that belong to a given “topic.” For example, an LDA analysis of documents produced by NCEAS working groups assigned each document a vector of topic proportions describing how many words in a given document came from each of sixteen different topics including “Ecosystem Services,” “Population Dynamics,” “Evolution,” and “Forest Ecology” (Rebich-Hespanha et al. 2015). This method has recently been applied to ecological data as well, where it has been used to summarize species assemblages based on the proportion of analogous “topics,” such as “Coniferous Forest” and “Oak-dominated Forest,” found in each community across spatial gradients (Valle et al. 2014). While Valle et al. (2014) demonstrated how this technique can be used with spatial data, it has not been widely adopted in ecology and has not been used to examine the temporal dynamics of a community

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

undergoing long-term compositional change.

Here, we apply LDA to examine community change over time in a desert rodent community. Censused monthly from 1977 to 2015, this rodent community has undergone significant temporal turnover in response to shifts in habitat from an arid desertified grassland to a shrub-dominated system (Brown et al. 1997; Thibault et al. 2004; Ernest et al. 2008). Both rapid and gradual reorganization have been invoked to explain changes in rodent composition. Using linear approaches applied to annual or seasonal data (Fig. 1), temporal turnover was interpreted as indicative of gradual long-term temporal turnover in response to habitat shifts from open arid grassland to shrubland (Thibault et al. 2004; Ernest et al. 2008). Other studies, focused on shorter-term dynamics associated with climate events, have proposed that temporal turnover occurred rapidly during finite periods (Valone et al. 1995; Thibault and Brown 2008). Thus this is an ideal site for examining how using an approach that does not assume linear change through time can provide insight into the mechanisms driving long-term temporal turnover.

Methods

We use a two-step process to model community dynamics from species count data collected at many time steps. In the first step, we use Latent Dirichlet Allocation (LDA) to reduce species composition to a small number of “topics” (which we will hereafter refer to as community-types), and describe change over time at the site in terms of these community-types. In the second step, we apply a change-point model to the resulting timeseries of community-types to identify specific points in time when the character of the community dynamics changed. To demonstrate the effectiveness as well as some limitations of our approach, we first test these methods on a set of simulated data before applying them to the observed rodent data. We construct the simulation using the simple scenario of two community-types, consisting of two non-overlapping sets of species with evenly apportioned abundance across species (Fig. 2A).

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We used these two communities to simulate three types of dynamics: a fast transition from one community-type to the other with stability before and after, a slow transition occurring over most of the time series, and constant proportions of the two community-types over time (Fig. 2 B-D). By multiplying the species composition matrix by the matrix of community dynamics over time, we obtain three simulated datasets of species counts at each time step.

Quantifying community change using Latent Dirichlet Allocation

We use LDA to reduce complex species composition data to a small number of readily-interpretable community-types (Fig. 3A) to describe changes in their relative contributions to the data (Fig. 3B). LDA is easiest to understand as a generative model, moving from the top of Figure 3 to the bottom. LDA postulates that species assemblages are composed of random samples from a relatively small number of community-types, which are defined in terms of relative species composition (Fig. 3A). In this context, community-types (referred to as "topics" in most of the LDA literature), could represent guilds, clades, or other groups of species whose abundances tend to rise and fall together. Assemblages (Fig. 3C) are generated by randomly drawing species from each of these community-types, with the relative frequency of each community-type determined by the proportions in Fig. 3B.

Fitting an LDA model thus involves simultaneously determining two sets of numbers: one defining the community-types (Fig. 3A), and one describing the observed species assemblages in terms of those types (Fig. 3B). This requires working backwards from observed patterns to infer the parameters of a possible underlying process – i.e., moving in Figure 3 from panel D up to panels A and B [see Blei et al. (2003) for the mathematical details]. Due to the complex relationship between the way each community-type is defined (Fig. 3A) and its influence on individual assemblages (Fig. 3B), exact inference is not possible in this model. Here, we used Blei et al.'s variational approximation, which simplifies this relationship. As discussed in Wainwright & Jordan (2008), variational methods are optimal in the sense that they minimize the

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

approximation error (i.e. they minimize the KL divergence from the exact model). This approximation allowed us to use Blei et al.'s (2003) iterative "variational expectation-maximization" procedure for jointly optimizing both sets of parameters (specifically for maximizing a lower bound on the log-likelihood), as implemented in the `topicmodels` package (Hornik & Grün 2011) for R 3.3.2 (R Core Team 2016). This gave us community-type compositions as well as estimates of the influence of each community-type on every assemblage observed in our data set (in our case, "assemblage" refers to each monthly census of the rodent community).

Like many clustering and ordination methods, LDA requires the number of community-types to be specified (i.e. it does not determine the number of community-types supported by the data). Unfortunately, perhaps the largest downside of LDA is that evaluating the predictive performance of a fitted model quantitatively (e.g. for choosing the optimal number of community-types) is notoriously difficult (Wallach et al. 2009). The difficulty is compounded by the large number of local maxima on the likelihood surface that prevent researchers from finding the global maximum or from effectively sampling the model's posterior distribution. These issues prevented us from using traditional model comparison to choose the number of community-types to include in the model. Instead, we used an approximate AIC procedure (based on a variational approximation to the full likelihood function) to inform the appropriate number of topics to include, regardless of the model's initial conditions (Fig. S1, in Supplement). To confirm that our results did not depend heavily on the random seed used to initialize the model, we calculated the Hellinger distance between the best-fitting models and the 99 next-best fitting ones (see Fig. S2, in Supplement).

We applied the LDA model to the three simulated datasets (Fig. 2 E-G). Our AIC procedure correctly predicted that two community-types was the best fit in all three cases. The LDA model was able to accurately reproduce the community compositions and the dynamics of the community-types used to create the simulated data for two of the three datasets (Fig. 2 E,F),

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

but for the dataset which represents constant proportions of the communities over time, the model was not able to distinguish the two communities (Fig. 2G).

Quantifying when change occurs using a change-point model

While LDA simplifies multivariate species composition to better visualize community dynamics, its outputs are continuous: they do not tell us if (or when) a discrete change in community structure has occurred. We fit a change-point model (Western & Kleykamp 2004) to identify abrupt transitions in the time series of community-type prevalence generated by the LDA model. Change-point models break up a time series into intervals, with a different set of parameters to describe the time series during each interval. Between each pair of change-points, we modeled each community-type's prevalence as a sinusoid with a period of one year, to control for seasonal fluctuations. Each interval was modeled using a separate multinomial generalized linear model fit with the nnet package (Venables & Ripley 2002). This gave us a quasi-likelihood for each interval (McCullagh & Nelder 1989). Since our quasi-likelihood did not inflate the variance, these quasi-likelihoods can be interpreted on the same scale as a conventional likelihood for purposes of model comparison (Anderson et al. 1994). The product of interval-level quasi-likelihoods yields the quasi-likelihood for the full data set; this value will be largest when the change-points break the time series into relatively stable intervals that can be explained well by the generalized linear model.

The number of possible change-point locations was too large to evaluate all the possibilities exhaustively, so we used Markov chain Monte Carlo (MCMC) to collect a representative sample of change-point locations that are consistent with the data. Initial experiments with Metropolis sampling showed poor mixing, so we implemented a parallel tempering sampler (also called Metropolis-coupled MCMC and replica-exchange MCMC) to facilitate movement of the Markov chain between modes via exchanges with auxiliary Markov chains that rapidly explore the space of possible change-points (Earl & Deem 2005). We fit

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

change-point models with up to five change-points for the rodent data, and evaluated model performance by comparing the average log-likelihood to the number of model parameters (see Fig. S6; Gelman et al. 2014).

We applied our change-point model to the LDA results for the simulations to determine where it identified changes in the community (Fig. 2 H-J). The results of the change-point model differed by the type of change being estimated: the fast change resulted in a narrow peak in the distribution of change point estimates; the slow change resulted in a wider distribution with a peak centered at the midpoint of the slow change; and the uniform dynamics resulted in a uniform distribution. In this way, the change-point model can provide information not only on the timing of community change, but also the rate of change.

Application to real data: desert rodents in Arizona

We applied the above methods to a 38-year time series (1977-2015) of a desert rodent community. Located on 20 hectares of Chihuahuan Desert near the town of Portal, Arizona, this site has undergone considerable habitat change: a 3-fold increase in woody vegetation between 1977 and the mid-1990s transitioned this site from an open desert grassland with widely scattered woody shrubs to a desert shrubland (Fig. 1; Brown et al. 1997). Small mammal data is collected at this site on 24 permanent 50m by 50m plots, censused at monthly intervals with no major changes in methodology since 1977. The 24 plots are divided into three experimental treatments, though for this project we only use data from the 8 unmanipulated control plots which allow unrestricted access to all species from the regional pool (see Brown 1998 for details). As input to the LDA model, we pooled capture data from the 8 control plots, obtaining a table of counts for each species for each month of the time series. This amounted to 436 time steps and 21 species. There has been considerable turnover in species composition through time, with only 2 species captured consistently at almost every census since 1977.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Results

Over the nearly 40 years of the study, the dynamics of the rodent community are best described with four different community-types, with the majority of random seeds (75%) supporting four community-types for the best-fitting model (See Fig. S1, Supplement). Although some seeds supported three or five community-types, the qualitative results we discuss are generally robust to this choice (see Fig. S3, Fig. S4, Fig. S5 in Supplement for comparison). We found that community-types had consistent species compositions across the different seed runs (the average Hellinger distance from the model presented here was 0.085, on a scale from 0 to 1; see Fig. S2, Supplement), suggesting that the inferred community-types were also robust to the choice of random seed. The four community-types differ in the composition of both common and rare species (Fig. 4A). The most abundant members of community-types 1 and 2 are kangaroo rats from the genus *Dipodomys*: community-type 1 is co-dominated by *D. spectabilis* and *D. merriami*, while community-type 2 is dominated by *D. merriami* alone. In contrast, the most abundant members of community-types 3 and 4 are pocket mice from the genus *Chaetodipus*; community-type 3 is dominated by *C. baileyi* and community-type 4 by *C. penicillatus*. In Fig. 4A, the 21 species of small mammals are arranged on the x-axes in order of decreasing body size and grassland-affiliated species are denoted with bold outlines on their bars, to demonstrate that the four component communities differ not only in the identity of species making up the community, but also the distribution of body sizes and habitat preferences contained in the community.

Through time, the different community-types vary in their prevalence and dynamics (Fig. 4B). When the study began, the desert rodent community mainly consisted of community-type 1 (Fig. 4, light blue). In the mid-1980s, the rodent community transitioned to community-type 2 (Fig. 4, dark blue) and then transitioned again in the late 1990s to become a mix of community-types 2, 3, and 4 (Fig. 4, dark blue, gold, and grey, respectively). Finally, around 2010, the community entered its current state which is seasonal oscillations between community-types 2

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

and 4 (Fig. 4, dark blue and grey). These dynamics and community-types are consistent with previous studies that documented the decline of *D. spectabilis* (the co-dominant species of community-type 1) in the mid-1980s (Valone et al. 1995), the colonization and rise to dominance of *C. baileyi* (the dominant species in community-type 3) in the late-1990s (Ernest & Brown 2001), shifts in the body size structure of the community from large species to smaller species (White et al. 2004), and a general decline in grassland-affiliated species and an increase in shrubland-affiliated species (Ernest et al. 2008).

Visually, the LDA results suggest that major shifts in community dynamics occurred multiple times over the study. Using our change-point approach, we found that four transition periods were best supported by the data (for comparison of models using 2, 3, 4, and 5 change-points see Fig. S6 in supplement). Histograms showing the locations of these four change-points are shown in Fig. 4C, with the distribution of each point shown in a different shade of gray. Using these distributions, we located the 95% confidence interval for when each of these transitions occurred: December 1983-July 1984, October 1988-January 1996, September 1998-December 1999, and June 2009-September 2010. Fig. 4D shows the change-point model's estimate of how the prevalence of the four community-types differs before and after each transition event, demonstrating that three of these events (1984, 1998-1999, 2009-2010) are driven by a shift in which community-type is most prevalent, marking a major shift in community structure. These three major shifts in community structure are separated from each other by 10-15 years. Between the 1998-1999 change-point and the 2009-2010 change-point, the model did not detect any shifts in the average prevalence of the community-types (i.e. intercept) or in the phase/amplitude of the annual cycles. The 1988-1996 transition differs from the other three major reorganization events. It has a broader change-point distribution, and there is no change in which community-type is dominant. However, this change-point indicates an increase in the variance (amplitude) of the dominant community-type through this period. Minor changes in community structure also occur as community-type 1 disappears and community-types 3 and 4

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

increase their prevalence. Together these dynamics suggest the 1988-1996 transition is not a rapid shift in community structure like the other events, but is instead either an increase in the variance of community structure and/or a signal of a very slow shift in species composition that was abruptly terminated with the major reorganization event of 1998-1999.

Discussion

Our results suggest that long-term composition change occurred primarily through three distinct reorganization events in 1984, 1999, and 2010. Our approach does not allow us to definitively determine how long each of these events lasted, but based on similarities to the simulation results, the three major reorganization events appear to have been relatively rapid. Our pattern of discrete and relatively rapid transitions is consistent with either intrinsic or extrinsic regime shifts. In an intrinsic regime shift, gradual changes in the environment eventually override intrinsic stabilizing mechanisms causing sudden reorganization events (Scheffer et al. 2001). Extrinsic regime shifts are generated either by rapid shifts in the environment which trigger bursts of niche-based reorganization as the community tracks the environment (Wirth et al. 2008), or by extreme events which can trigger alternative community assembly pathways (Chase 2003; Fukami et al. 2010). The shift in rodent composition from grassland- to shrubland-affiliated species (Ernest et al. 2008; Fig. 4A) suggests that the community is tracking an environmental change: the long-term shift in shrub cover. Thus, for the extrinsic niche-based scenario to occur, our site would need to have experienced multiple rapid bursts of shrub growth and establishment within the 38 years of this study. Unfortunately, without a relatively high frequency time series on shrub cover, we cannot distinguish between multiple intrinsic regime shifts in response to continual change in shrub cover and niche-based regime shifts in response to bursts in increasing shrub cover. However, given the slow dynamics of shrub growth relative to the observed dynamics of the rodent community (Goslee et al. 2003; Okin et al. 2009), we suspect that the latter scenario is unlikely.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

We can use the history of the site to assess whether extreme events could be related to our transition events. Extreme climatic events have been documented to trigger reorganization of communities, but it is difficult to predict when an extreme event will cause reorganization versus when the community will simply return to its previous state (e.g. Arthur et al. 2006). Over the nearly 40 years of this study, four unusual climate events (Fig. 5, gray bars) have been documented that happen to occur in close proximity to each of our transition events (Fig. 5, horizontal red lines): an intense tropical storm in October 1983 (Valone et al. 1995), a large flood event in August 1999 (Thibault & Brown 2008), and regional droughts in 1993-1994 and in 2009-2010. Each of these correspond to known impacts on the abundance of rodents: either the abundance across all species (Thibault et al. 2004; Allington et al. 2013; Fig. 5), or of important species such as the dominant keystone *D. spectabilis* (Valone et al. 1995; Brown et al. 1997). While we cannot prove that these events are related to our transition events, the coincidence is striking.

Extreme events are hypothesized to trigger reorganization events when they push the community past a threshold that allows a new species configuration to establish (Smith et al. 2009). If extreme climate events are causing our reorganization events, our results suggest that part of this threshold could involve either major declines in the dominant species or by reducing the community overall to low total abundances. Many studies have empirically documented how removing or reducing dominant species restructures communities (e.g. Paine 1966; Smith & Knapp 2003; Sasaki & Lauenroth 2011). Alternatively, reducing a community to low total abundance can trigger reorganization events via various processes. At low community abundances, the role of stochasticity in determining competitive outcomes increases, allowing alternative community compositions to establish (Orrock & Fletcher 2005; Chase et al. 2009). During recovery from a low abundance event, the order of species recolonization can impact the trajectory of community assembly as early colonizing species preclude later species from colonizing (i.e. priority effects; Fukami et al. 2010). These established species then preempt

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

resources and space, impeding colonization and population growth of new species better suited to the environment (i.e. incumbency effects; Thibault & Brown 2008). Thus by decimating the dominant species or the entire community, an extreme climatic event may create a threshold event that allows regime shifts to occur.

While the low abundance of the dominant species or entire community may create the potential for reorganization events, it seems unlikely that stochastic processes alone explain the shift in our community from grassland- to shrubland-affiliated species. This niche-based change in the community suggests that the increase in woody vegetation is also playing a role in these reorganization events. We hypothesize that as the environment changed, incumbency effects prevented the rodent community composition from tracking those changes, leading to a larger and larger difference between the current state of the environment and the niche characteristics in the community. When this happens, extreme events may act as a reset button, by allowing the establishment and/or population growth of species better suited to the current conditions. However, as the environment continues to change, the difference between environment and community state grows again, until the next extreme climatic event resets the community. The increased variability in the dominant community-type during the 1988-1996 transition event could indicate erosion of stabilizing mechanisms as the current state of the system was no longer well-suited to the environment (i.e., leading variance before a regime shift; Carpenter & Brock 2006). If extreme events indeed trigger reorganization events by allowing communities to re-establish equilibrium with their environment, then this may also explain why it is difficult to predict when an extreme event will trigger community reorganization. It depends both on the magnitude of the impact on either dominant species or total community abundances, and whether the community is out of sync with current environmental conditions. If the niche characteristics of the species present were already in sync with the environment, no reorganization may occur even if the impact on abundances was large.

Whether or not our reorganization events are being caused by extreme climatic events

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

or simply coincident with them, our results show that long-term community change did not occur continuously in a gradual fashion. Although changes in rodent community composition appear to track changes in the environment at the multi-decadal scale, the fact that community change occurred in discrete jumps reveals that the system also contains mechanisms that resist change at small time scales (e.g. incumbency effects, priority effects). Ignoring these resistance mechanisms, or assuming that community change must be a gradual process, will lead to inaccurate predictions of how a system will respond to future perturbations. It is therefore critical to consider time scale when using past dynamics to infer mechanisms of change or to predict future dynamics. These insights would not have been possible without the long-term, high frequency monitoring of this study system, emphasizing growing concerns (Lindenmayer et al. 2012; Hofmann et al. 2013; Hughes et al. 2017) that maintaining long-term studies will be critical for detecting, understanding, and predicting future changes in nature.

Acknowledgements

This research was supported in part by the Gordon and Betty Moore Foundation's Data-Driven Discovery Initiative through Grant GBMF4563 to E.P. White. We would also like to thank D. Valle and E. Damschen for informative discussions on LDA and stochastic community assembly, respectively. The Portal Project has been funded nearly continuously by the National Science Foundation, most recently by DEB-1622425, which has also supported E. Christensen.

References

Allington, G.R.H., Koons, D.N., Ernest, S.K.M., Schutzenhofer, M.R., & Valone, T.J. (2013). Niche opportunities and invasion dynamics in a desert annual community. *Ecol. Lett*., 16, 158–166.

Anderson, D.R., Burnham, K.P., & White, G.C. (1994). AIC Model Selection in Overdispersed

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Capture-Recapture Data. *Ecology,* 75, 1780-1793.

Arthur, R., Done, T.J., Marsh, H., & Harriott, V. (2006). Local processes strongly influence post-bleaching benthic recovery in the Lakshadweep Islands. *Coral Reefs*, 25, 427–440.

Bagchi, S., Singh, N.J., Briske, D.D., Bestelmeyer, B.T., McClaran, M.P., and Murthy, K. (2017). Quantifying long-term plant community dynamics with movement models: implications for ecological resilience. *Ecol. Appl.,* 27, 1514–1528.

Beaugrand, G. (2004). The North Sea regime shift: Evidence, causes, mechanisms and consequences. *Prog. Oceanogr*., 60, 245–262.

Biggs, R., Carpenter, S.R., & Brock, W.A. (2009). Turning back from the brink: Detecting an impending regime shift in time to avert it. *Proc. Natl. Acad. Sci*., 106, 826–831.

Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet Allocation. *J Mach Learn Res*, 3, 993–1022.

Brown, J. H. (1998). The desert granivory experiments at Portal. In: *Experimental Ecology,* Editors Resetarits, W.J. Jr. and Bernardo, J. Oxford University Press, New York, New York, USA, pp. 71-95.

Brown, J.H., Valone, T.J., & Curtin, C.G. (1997). Reorganization of an arid ecosystem in response to recent climate change. *Proc. Natl. Acad. Sci*., 94, 9729–9733.

Carpenter, S.R., & Brock, W.A. (2006). Rising variance: a leading indicator of ecological

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

transition. *Ecol. Lett.*, 9, 311–318.

Chase, J.M. (2003). Community assembly: when should history matter? *Oecologia*, 136, 489–498.

Chase, J.M., Biro, E.G., Ryberg, W.A., & Smith, K.G. (2009). Predators temper the relative importance of stochastic processes in the assembly of prey metacommunities. *Ecol. Lett*., 12, 1210–1218.

Dornelas, M., Gotelli, N.J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C., et al. (2014). Assemblage Time Series Reveal Biodiversity Change but Not Systematic Loss. *Science*, 344, 296–299.

Earl, D.J., & Deem, M.W. (2005). Parallel tempering: Theory, applications, and new perspectives. *Phys. Chem. Chem. Phys.,* 7, 3910-3916.

Ernest, S.K.M., & Brown, J.H. (2001). Delayed Compensation for Missing Keystone Species by Colonization. *Science*, 292, 101–104.

Ernest, S.K.M., Brown, J.H., Thibault, K.M., White, E.P., & Goheen, J.R. (2008). Zero Sum, the Niche, and Metacommunities: Long-Term Dynamics of Community Assembly. *Am. Nat*., 172, E257–E269.

Ernest, S.K., Valone, T.J., & Brown, J.H. (2009). Long-term monitoring and experimental manipulation of a Chihuahuan Desert ecosystem near Portal, Arizona, USA. *Ecology*, 90, 1708. doi: 10.1890/08-1222.1

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Ernest, S.K.M., Yenni, G.M., Allington, G., Christensen, E.M., Geluso, K., Goheen, J.R., et al. (2016). Long-term monitoring and experimental manipulation of a Chihuahuan desert ecosystem near Portal, Arizona (1977–2013). *Ecology*, 97, 1082. doi: 10.1890/15-2115.1

Fukami, T., Dickie, I.A., Paula Wilkie, J., Paulus, B.C., Park, D., Roberts, A., et al. (2010). Assembly history dictates ecosystem functioning: evidence from wood decomposer communities. *Ecol. Lett*., 13, 675–684.

Gelman, A., Hwang, J., & Vehtari, A. (2014). Understanding predictive information criteria for Bayesian models*. Stat. Comput*. 24, 997–1016.

Goslee, S.C., Havstad, K.M., Peters, D.P.C., Rango, A., & Schlesinger, W.H. (2003). High-resolution images reveal rate and pattern of shrub encroachment over six decades in New Mexico, U.S.A*. J. Arid Environ*., 54, 755–767.

Hornik, K., & Grün, B. (2011). topicmodels: An R Package for Fitting Topic Models. *J. Stat. Softw.* 40, 1–30. doi: 10.18637/jss.v040.i13 (URL: http://doi.org/10.18637/jss.v040.i13).

Hofmann, G.E., Blanchette, C.A., Rivest, E.B., & Kapsenberg, L. (2013). Taking the Pulse of Marine Ecosystems: The Importance of Coupling Long-Term Physical and Biological Observations in the Context of Global Change Biology. *Oceanography*, 26, 140–148.

Hughes, B.B., Beas-Luna, R., Barner, A.K., Brewitt, K., Brumbaugh, D.R., Cerny-Chipman, E.B., et al. (2017). Long-Term Studies Contribute Disproportionately to Ecology and Policy. *BioScience,* 67, 271–281.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Isbell, F., Tilman, D., Polasky, S., Binder, S., & Hawthorne, P. (2013). Low biodiversity state persists two decades after cessation of nutrient enrichment. *Ecol. Lett.,* 16, 454–460.

Jiguet, F., Brotons, L., & Devictor, V. (2011). Community responses to extreme climatic conditions. *Curr. Zool.,* 57, 406–413.

Jochimsen, M.C., Kümmerlin, R., & Straile, D. (2013). Compensatory dynamics and the stability of phytoplankton biomass during four decades of eutrophication and oligotrophication. *Ecol. Lett*., 16, 81–89.

La Sorte, F.A., & Boecklen, W.J. (2005). Temporal turnover of common species in avian assemblages in North America. *J. Biogeogr*., 32, 1151–1160.

Lindenmayer, D.B., Likens, G.E., Andersen, A., Bowman, D., Bull, C.M., Burns, E., et al. (2012). Value of long-term ecological studies. *Austral Ecol*., 37, 745–75.

Magurran, A.E., & Henderson, P.A. (2003). Explaining the excess of rare species in natural species abundance distributions. *Nature*, 422, 714–716.

Magurran, A.E., & Henderson, P.A. (2010). Temporal turnover and the maintenance of diversity in ecological assemblages. *Philos. Trans. R. Soc. Lond. B Biol. Sci*., 365, 3611–3620.

Magurran, A.E., Dornelas, M., Moyes, F., Gotelli, N.J., & McGill, B. (2015). Rapid biotic homogenization of marine fish assemblages. *Nat. Commun*., 6, 8405.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

McCullagh, P. & Nelder, J.A. (1989). Generalized Linear Models: Monographs on Statistics and Applied Probability. 2nd edition. Chapman & Hall.CRC, London, UK. pp. 328

McGill, B.J., Dornelas, M., Gotelli, N.J., & Magurran, A.E. (2015). Fifteen forms of biodiversity trend in the Anthropocene. *Trends Ecol. Evol*., 30, 104–113.

Okin, G.S., D'Odorico, P., & Archer, S.R. (2009). Impact of feedbacks on Chihuahuan desert grasslands: Transience and metastability. *J. Geophys. Res. Biogeosciences*, 114, G01004.

Orrock, J.L., & Watling, J.I. (2010). Local community size mediates ecological drift and competition in metacommunities. *Proc. R. Soc. Lond. B Biol. Sci*., 277, 2185–2191.

Orrock, J.L., & Fletcher Jr., R.J. (2005). Changes in Community Size Affect the Outcome of Competition. *Am. Nat*., 166, 107–111.

Österblom, H., Hansson, S., Larsson, U., Hjerne, O., Wulff, F., Elmgren, R., et al. (2007). Human-induced Trophic Cascades and Ecological Regime Shifts in the Baltic Sea. *Ecosystems*, 10, 877–889.

Rebich-Hespanha S, Walker L, Donahue D, Fine L, French L, Hampton SE, et al. (2015). *VizSearch: Visualizing NCEAS Research Projects and Publications.* Available at: https://www.nceas.ucsb.edu/VizSearch. Last accessed 30 June 2017.

R Core Team (2016). R: A language and environment for statistical computing. R Foundation for statistical Computing, Vienna, Austria. URL: https://www.R-project.org/.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Rowe, R.J. (2007). Legacies of Land Use and Recent Climatic Change: The Small Mammal Fauna in the Mountains of Utah. *Am. Nat*., 170, 242–257.

Sasaki, T., & Lauenroth, W.K. (2011). Dominant species, rather than diversity, regulates temporal stability of plant communities. *Oecologia*, 166, 761–768.

Scheffer, M., & Carpenter, S.R. (2003). Catastrophic regime shifts in ecosystems: linking theory to observation. *Trends Ecol. Evol*., 18, 648–656.

Scheffer, M., Carpenter, S., Foley, J.A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, 413, 591–596.

Shea, K., & Chesson, P. (2002). Community ecology theory as a framework for biological invasions. *Trends Ecol. Evol*., 17, 170–176.

Smith, M.D. (2011). An ecological perspective on extreme climatic events: a synthetic definition and framework to guide future research. *J. Ecol*., 99, 656–663.

Smith, M.D., & Knapp, A.K. (2003). Dominant species maintain ecosystem function with non-random species loss. *Ecol. Lett*., 6, 509–517.

Smith, M.D., Knapp, A.K., & Collins, S.L. (2009). A framework for assessing ecosystem dynamics in response to chronic resource alterations induced by global change. *Ecology*, 90, 3279–3289.

Thibault, K.M., & Brown, J.H. (2008). Impact of an extreme climatic event on community

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

assembly. P*roc. Natl. Acad. Sci*., 105, 3410–3415.

Thibault, K.M., White, E.P., & Ernest, S.K.M. (2004). Temporal Dynamics in the Structure and Composition of a Desert Rodent Community. *Ecology*, 85, 2649–2655.

Tingley, M.W., Monahan, W.B., Beissinger, S.R., & Moritz, C. (2009). Birds track their Grinnellian niche through a century of climate change. *Proc. Natl. Acad. Sci*., 106, 19637–19643.

Valle, D., Baiser, B., Woodall, C.W., & Chazdon, R. (2014). Decomposing biodiversity data using the Latent Dirichlet Allocation model, a probabilistic multivariate statistical method. *Ecol. Lett*., 17, 1591–1601.

Valone, T.J., Brown, J.H., & Jacobi, C.L. (1995). Catastrophic Decline of a Desert Rodent, Dipodomys spectabilis: Insights from a Long-Term Study. *J. Mammal*., 76, 428–436.

Venables, W.N, & Ripley, B.D. (2002). *Modern Applied Statistics with S.* Fourth Edition. Springer, New York, New York. pp. 199-206.

Wainwright, M.J., & Jordan, M.I. (2008). Graphical Models, Exponential Families, and Variational Inference. *Found. Trends® Mach. Learn*. 1, 1–305.

Wallach, H.M., Murray, I., Salakhutdinov, R., and Mimno, D. (2009). Evaluation Methods for Topic Models. In Proceedings of the 26th Annual International Conference on Machine Learning, (New York, NY, USA: ACM), pp. 1105–1112.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Walther, G.-R. (2010). Community and ecosystem responses to recent climate change. *Philos. Trans. R. Soc. Lond. B Biol. Sci*., 365, 2019–2024.

Western, B., & Kleykamp, M. (2004). A Bayesian Change Point Model for Historical Time Series Analysis. *Political Analysis*, 12, 354-374. doi:10.1093/pan/mph023

White, E.P., Ernest, S.K.M., & Thibault, K.M.. (2004). Trade-offs in Community Properties through Time in a Desert Rodent Community. *Am. Nat.*, 164, 670–676.

Williams, J.W., Blois, J.L., & Shuman, B.N. (2011). Extrinsic and intrinsic forcing of abrupt ecological change: case studies from the late Quaternary. *J. Ecol*., 99, 664–677.

Wirth, C., Lichstein, J.W., Dushoff, J., Chen, A., & Chapin, F.S. (2008). White Spruce Meets Black Spruce: Dispersal, Postfire Establishment, and Growth in a Warming Climate. *Ecol. Monogr*., 78, 489–505.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

FIGURES

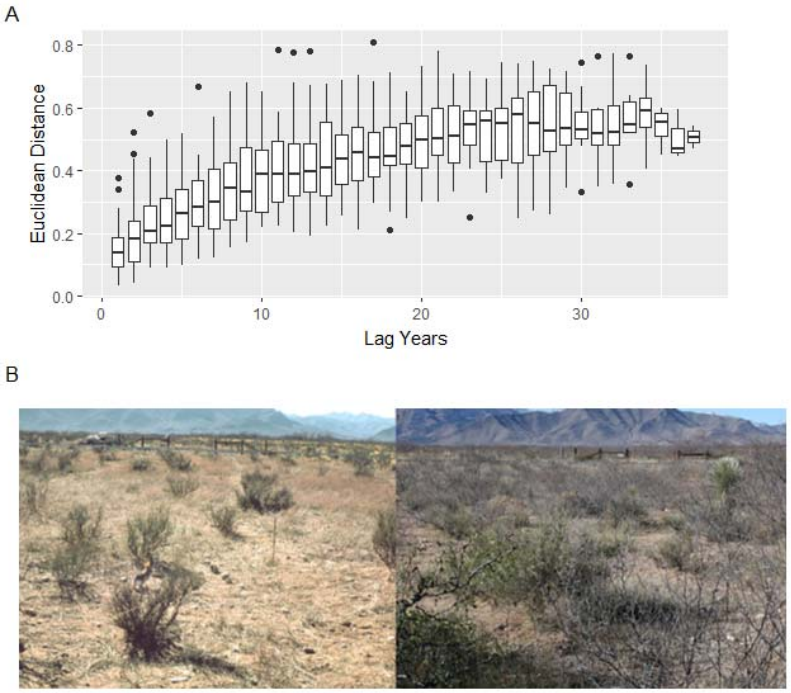

Figure 1

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

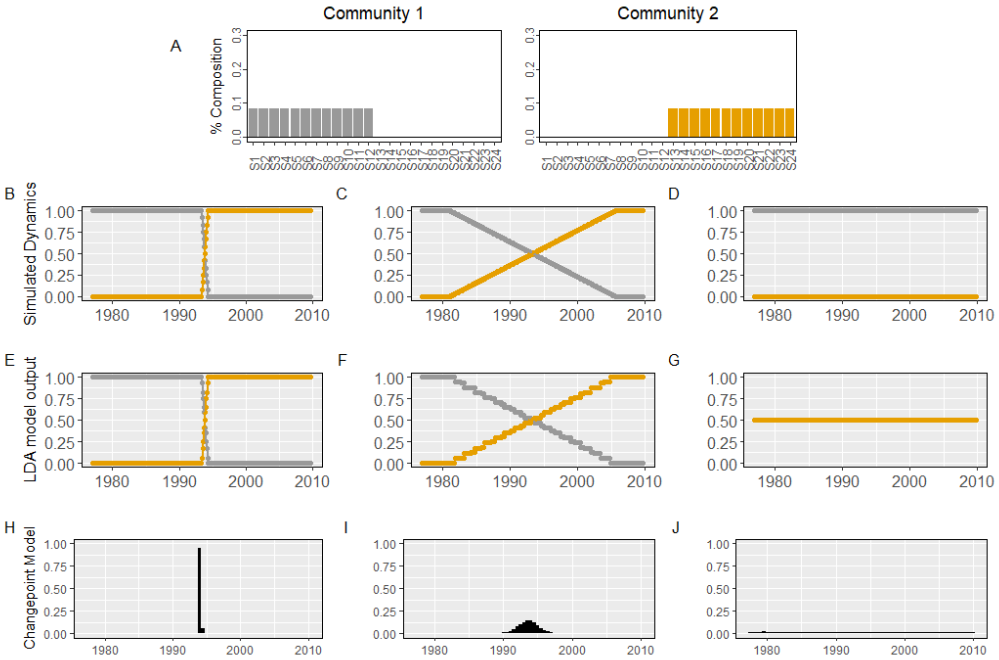

Figure 2

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

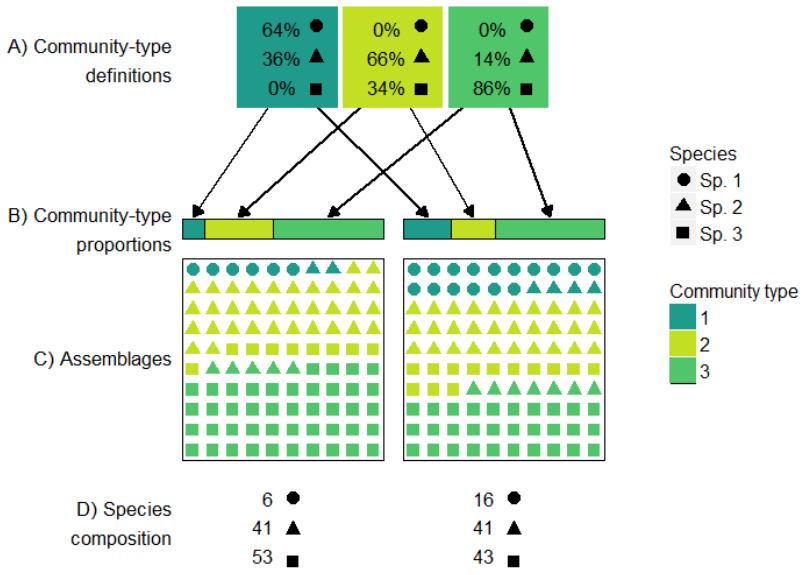

Figure 3

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

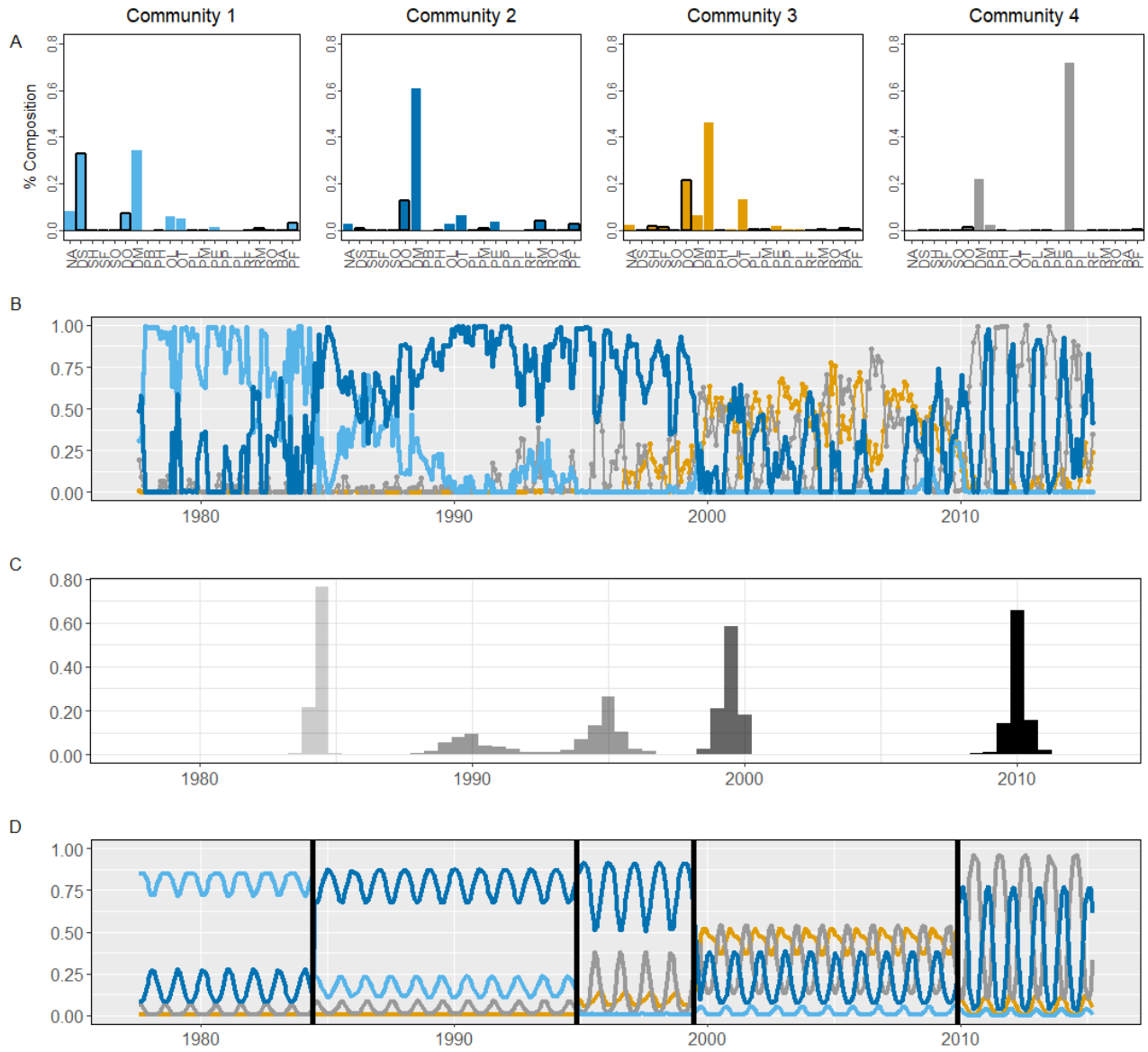

Figure 4

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

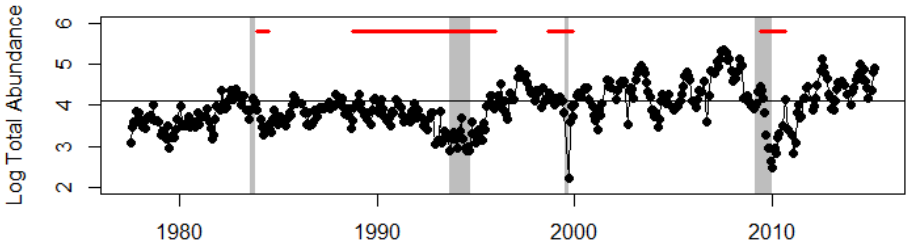

Figure 5

Figure 1. Example of a linear approach to calculating temporal trends in species composition, with data from a desert rodent community. **A.** Euclidean distance between all possible pairs of years (1977-2015) as a function of time lag (after Thibault et al 2004). The monotonically increasing difference in community composition with increasing time lag corresponds to a significant increase in shrub cover; **B.** paired photographs showing same perspective of Plot 1 in 1995 and 2015, demonstrating the increase in shrub cover during this time.

Figure 2. Demonstrating LDA and change-point model using simulated data. **A.** species composition of two community-types to be simulated; **B.** simulation of unchanging community dynamics; **C.** simulation of rapid transition from community-type 1 to community-type 2; **D.** simulation of gradual transition from community-type 1 to community-type 2; **E-G.** output of LDA model, trying to recover dynamics simulated in panel B-D; **H-J.** histograms showing the distribution of estimated change-point location.

bioRxiv preprint doi: https://doi.org/10.1101/163931; this version posted July 15, 2017. The copyright holder for this preprint (which was not certified by peer review) is the author/funder, who has granted bioRxiv a license to display the preprint in perpetuity. It is made available under aCC-BY-NC 4.0 International license.

Figure 3. Graphical representation of the LDA model. **A.** Community-type definitions: the relative species composition of each community-type; **B.** Community-type proportions: the relative frequency of each community-type contributing to assemblages; **C.** Assemblages: generated by randomly drawing species from the community-type pools in A according to the proportions in B; **D.** Species composition of assemblages, as would be seen in actual community data.

Figure 4. A. species composition of the four community-types produced by the LDA model, with species arranged on the x-axis by decreasing body size, and grassland-affiliate species emphasized by black boxes around the bars (see Table S1, Supplement); **B.** prevalence of the four community-types over time as estimated by the LDA model; **C.** histograms of four change-points representing the greatest changes in the prevalence of community-types from B; and **D.** the change-point model's estimate of how community-type prevalence changes before and after each transition point. Species codes in panel A: NA = *Neotoma albigula*, DS = *Dipodomys spectabilis*, SH = *Sigmodon hispidus*, SF = *Sigmodon fulviventer*, SO = *Sigmodon ochrognathus*, DO = *Dipodomys ordii*, DM = *Dipodomys merriami*, PB = *Chaetodipus baileyi*, PH = *Chaetodipus hispidus*, OL = *Onychomys leucogaster*, OT = *Onychomys torridus*, PL = *Peromyscus leucopus*, PM = *Peromyscus maniculatus*, PE = *Peromyscus eremicus*, PP = *Chaetodipus penicillatus*, PI = *Chaetodipus intermedius*, RF = *Reithrodontomys fulvescens*, RM = *Reithrodontomys megalotis*, RO = *Reithrodontomys montanus*, BA = *Baiomys taylori*, PF = *Perognathus flavus*.

Figure 5. Total rodent abundance per hectare over time. Solid horizontal line shows long-term mean. Grey bars show duration of the four extreme events. Red bars show the 95% confidence interval for each of the community transition events.