Regime Detection Measures for the Practical Ecologist

A Thesis
Presented to
The Division of
University of Nebraska-Lincoln

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Jessica L. Burnett

2019

Approved for the Division
(School of Natural Resources)

Craig R. Allen

Dirac Twidwell

I thank my supervisors, Craig Allen and Dirac Twidwell, for providing me with this amazing opportunity and for supporting my growth as an independent researcher. I thank my also committee members, Craig Allen, David Angeler, John De Long, Dirac Twidwell, and Drew Tyre for their support and advisement, but especially for their comprehensive examination–I found this process transformative (albeit very stress-inducing). I parrticularly thank Dirac for his examination questions—I never knew how much I didn't know until I studied your recommendations. I also thank Craig and Dirac for supporting my efforts to study and conduct research in Austria. Studying at the International Institute for Applied Systems Analysis was an amazing opportunity! I thank Brian Fath and Elena Rovenskaya for their advisement, members of the Applied Systems Analysis research group for their feedback on my research, and to the postdocs and YSSPers. I owe thanks to Craig Allen and Kevin Pope for entertaining my many hours of discussion (interrogation?) regarding federal employment. I would like to especially thank some of the amazing and brilliant **female scientists** in my life for their encouragement: Jane Anderson, Karen Bailey, Hannah Birge, Mary Bomberger Brown, Tori Donovan, Brittany Dueker, Allie Schiltmeyer, Katie Sieving, Erica Stuber, Becky Wilcox, Carissa Wonkka, and Lyndsie Wszola. I thank these women and others for their contributions to my professional development: David Angeler, Christie Bahlai, Mary Bomberger Brown, John Carroll, Jenny Dauer, John DeLong, Tarsha Eason, Brian Fath, Ahjond Garmestani, Chris Lepczyk, Frank La Sorte, Chai Molina, Zac Warren, Hao Ye. I also thank fellow graduate students with whom I forged long-lived personal and professional relationships: Hannah Birge, Tori Donovan, Caleb Roberts, Allie Schiltmeyer, and Lyndsie Wszola. It is also worth noting that I among those afflicted with mental health "disorders". I am first grafteful to one friend (H) who unknowingly destigmatized mental health in my mind and without whom I may have never sought treatment. I also applaud fellow students and faculty in SNR and UNL who have been active in promoting positive mental health. I am forever indebtted to my general practitioner and mental health advocate, Terry Thomas M.A., M.S.N., A.P.R.N. *Financial support.* This research was funded by the U.S. Department of Defense<80><99>s Strategic Environmental Research and Development Program (SERDP project ID: RC-2510). The University of Nebraska-Lincoln (UNL) has been highy supportive in my doctoral studies and reserach. I am grateful for the generous of donors to the University of Nebraska Foundation, which provided me with two prestigious supplemental fellowships: Fling and Othmer. I also thank the Nelson Family (Nelson Memorial Fellowship, UNL) for supporting my domestic and international travel to conferences and workshops, and the Institute of Agriculture and Natural Resources, who funded large portions of my academic and research-related travel. I thank the School of Natural Resources for their financial support in my conference travel. The U.S. National Academy of Sciences generously funded part of my travel to the International Institute for Applied Systems Analysis (IIASA). This financial support provided me not only with invaluabe opportunities to attend and present at national and international conferences and workshops, conduct research abroad, and network–this funding alleviated some financial pressures associated with graduate school which allowed a more refined focus on my dissertation research. The opportunities and experiences provided to me by each funding source were amazing, thank you. Finally, to my partner of eight years–Schultzie–thank you for everything. Just kidding, thank you, Nat Price, you are amazing.

Table of Contents

List of Tables

List of Figures

Abstract

Identifying abrupt changes in the structure and functioning of systems, or system regime shifts, in ecological and social-ecological systems leads to an understanding of relative and absolute system resilience. Resilience is an emergent phenomenon of complex social-ecological systems, and is the ability of a system to absorb disturbance without reorganizing into a new state, or regime. Resilience science provides a framework and methodology for quantitatively assessing the capacity of a system to maintain its current trajectory (or to stay within a certain, and often desirable regime). If and when a system<80><99>s resilience is exceeded, it crosses a threshold and enters into an alternate regime (or undergoes a regime shift).

I use Fisher Information to detect regime shifts in time and space using avian community data obtained from the North American Breeding Bird Survey within the area east of the Rockies and west of the Mississippi River. Fisher Information is a technique that captures the dynamic of a system, and this metric will be calculated about a suite of bird species abundances aggregated to the route level for all possible time periods. Transmutation (aggregation error) about inclusion or exclusion of certain bird species, functional groups, and guilds will be analyzed. Efforts have been made to develop early warning indicators of regime shifts in ecosystems, however, for most ecosystems there is great uncertainty in predicting the risk of a regime shift, regarding both when and how long it will take to happen and if it can be recognized early enough to be avoided when desired. We will complement the use of Fisher Information with multiple

discontinuity analyses about body mass distributions at the route-level to achieve the aim of identifying individual species that best serve as early-warning indicators of regime shifts. For those species found on the edges of body mass aggregations, we test the hypothesis that the background variance in their abundances (on Breeding Bird Survey routes) will increase more than those not observed at the edge of discontinuity aggregations. Identification of early-warning indicators of regime shifts in ecological systems allows management efforts to focus on a single or a small number of species that inform us about ecosystem resilience and trajectory.

These methods transcend the primary objective of the Breeding Bird Survey (to monitor population trends) and use this expansive dataset in such a way that information about ecosystem order, trajectory, and resilience emerge. Here, we utilize an expansive dataset (the Breeding Bird Survey) to make broad-scale estimations and predictions about ecosystem resilience, regime status and trajectory, and ecosystem sustainability. Identification of regime shifts and early-warning indicator species may afford us the ability to predict system regime shifts in time.

Table of Definitions

Research surrounding regime shifts, threshold identification, change-point detection, bifurcation theory, etc. is muddled with jargon. Here, I provide a table of definitions (Table 1) for terms and concepts that may either be unfamiliar to the practical ecologist, or may have multiple meanings among and within ecological researchers and practitioners. With this table, I aim to both improve the clarity of this dissertation *and* highlight one potential issue associated with regime detection methods in ecology: semantics.

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature.

Term	**Definition**	**Synonyms**
Abrupt	A relative value of the speed and/or intensity of the change; the time period over which the regime shift occurs relative to the time observed (or expected to have been) in a particular state.	big, fast, quick, large
Alternative Stable State	**Controversially can be distilled as one of either: the number of unique stable configurations that a system can adopt (see Lewontin 1969), or the impacts that processes or pressures can have on a system's state (see May 1977).**	
Attractor	The set of values towards which a system tends regardless of its initial (starting) vaules.	

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature. *(continued)*

Term	Definition	Synonyms
Basin-Boundary Collision	**The parameter values for a system that causes the system to shift between alternate attractors.**	**non-local bifurcation**
Catastrophe Theory	The study of abrupt changes within a dynamical system.	
Catastrophic Bifurcation	**A relatively abrupt jump to an alternate attractor due to initial attractor.**	
Change-Point	See also 'Regime Shift'. A term often used in computer science, climatology, data science; represents the point at which a state changes its configuration.	
Change-Point Detection	**A change point method which does not require supervision; identifies potential change points without a priori potential change points.**	
Change-Point Estimation	A change point method which DOES require supervision; identifies potential change points when given a set of potential change points; well-developed in computer science, statistics, data mining, etc.; although well-developed, still lacks with giving statistical significance of change-points.	
Chaos	**A system with extreme sensitivity to initial conditions.**	
Critical Slowing Down (CSD)	When the recovery rate (time to return) of a system decreases (approaches zero) as a system approaches a critical point (possibly a threshold or tipping point). A characteristic observed in some empirical systems data (e.g. nutrient loading in shallow lakes).	
Degrees of Freedom	**The number of system parameters or components which vary independently.**	

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature. *(continued)*

Term	**Definition**	**Synonyms**
Domain of Attraction	The range of values around which a system fluctuates.	zone of fluctuation, basin of attraction, stable point, attractor
Driver	**A widespread anthropogenic source of change which leads to one or more pressures (e.g., land-use change).**	
Driver-Threshold Regime Shift	When a rapid change in external driver indcues a rapid change in ecosystem state.	
Dynamical System	**A time-dependent system which can be described in state-space.**	
Dynamical Systems Theory	The study of complex systems theory; the study of time-dependent systems.	
Equilibrium	**The set of values around which a system revolves and does not change.**	
Exogeneous Process (Forcing, Driver)	An external process influencing the state of the dynamical system.	
First-Order Stationarity	**When the mean is constantant over the observations.**	
Fold Bifurcation	This occurs when a stable point collides with an unstable point; when crossing a tipping point induces hysteresis.	
Fractal Properties	**A measurement of geometrical self-similarity; when a system has similar structure regardless of the scale of observation.**	**ergodic**

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature. *(continued)*

Term	**Definition**	**Synonyms**
Hysteresis	A system which is state-dependent (e.g. magnets); when a tipping point or threshold is crossed such that the previous state cannot be achieved by reversing the conditions.	
Leading Indicators	**When the statistical properties of the fluctuations (of the data) approach a critical transition.**	
Lyapunov Exponent (and Stability	A value that conveys the average rate of trajectory divergence that is caused by an endogenous force; how quickly (if at all) a system will tend away from a stable point if it starts near the stable point.	
Measure Theory	**The study of measures and measurement (e.g. volume, mass, time).**	
Moving (Sliding) Window Analysis	When a subsample of the data $$X_t$$ is used in lieu of a single observation, $$x_t$$.	
Noise	**Processes manifested in data which are unaccounted for; sometimes referred to as meaningless; random variability.**	
Non-Stationarity of the Mean Value	Infers that a trend or a periodicity is present in the time series.	
Online	**Real-time updating of model parameters, predictions, etc. (c.f. offline).**	
Persistent	A relative value of the longevity of the observed change in values.	long-lasting
Phase Space	**A graphical representation of two or more trajectories where one axis is not time. In this representation an equilibrium is defined as a single point in the state space.**	

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature. *(continued)*

Term	**Definition**	**Synonyms**
Prediction	A temporal forecast. Is intrinsic when a model and paramters are used to make forecast, is realized when the prediction becomes the actual state of the system.	
Pressure	**A perturbation which negatively influences a system, and can be defined as pulse, press, or monotonic.**	
Red Noise	Noise having zero mean, constant variance, and serial autocorrelation; autocorrelated random variability.	
Regime	**A set of system values that define a particular system state. Not necessarily stable, but some state variables or outputs of the system remain relatively constant over a defined period of time.**	
Regime Shift	"abrupt" and "persistent" change in a system's structure or functioning.	
Second-Order Stationarity	**The nean is constant and the covariance is a function of a time lag, but not of time.**	
Self-Similarity	A system satisfied by power-law scaling.	
Stable Equilibrium	**An equilibrium is stable when small perturbations do not induce change.**	
State Space	The set of all possible configurations of a system.	
State-Threshold Regime Shift	**When a graduaal change in external driver induces a rapid change in ecosystem state (e.g,. System crosses a threshold).**	
Stationarity	When the probability density function of a system does not change with time.	

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature. *(continued)*

Term	**Definition**	**Synonyms**
Statistical Stationarity	**A system with statistical properties unchanging over time. This concept extends to periodic stationarity for systems exhibiting periodic behavior.**	
Strange Attractor	An attractor which has fractal structure (an observable fractal dimension).	
Supervised Machine Learning	**When classifiers are used to train the data a priori.**	
System State	The observed (current) instance of the system within a state space.	
Threshold	**A point where the system reacts to changing conditions.**	
Tipping Point	A point in a system's trajectory where a small change in an endogenous force induces a large change in sytem state or values; the point where a system can flip into an alternative state.	
Trajectory	**The path of an object or system through space-time.**	**orbit, path**
Transient	A behavior or phenomenon which is responsive to intial (starting) conditions, or its effect declines over time.	
Trend Smoothing	**Local averaging of values such that the non-systematic components of the system are washed out.**	
Unstable Equilibrium	An equilibrium is unstable when small perturbations induce change.	

Table 1: A table of definitions for terms, theories, and phrases often appearing in ecological regime shift literature. *(continued)*

Term	**Definition**	**Synonyms**
Unsupervised Maain Learning	**When no prior training of the data is required (i.e. no classifications necessary a priori) to classify it.**	
White Noise	Noise having zero mean, constant variance, and is not autocorrelated; uncorrelated random variability.	

Chapter 1

Introduction

Anthropogenic activity in the last few decades will continue to influence the interations within and among ecological systems worldwide. The complexity of and drivers of changes in coupled human-natural systems is consequently altered, further limiting our ability to detect and predict change and impacts of change (Liu et al., 2007; Scheffer, 2009). Early warning systems are developed to detect, and in some cases predict, abrupt changes in disparate systems [e.g. cyber security [@!!!!!], infrastructure [@!!!!], banking crises (Davis & Karim, 2008), and agricultural systems]. The need to develop and improve early warning systems for natural and coupled human-natural systems is exacerbated by the consequences of climate change and globalization, especially when the human-related stakes are high.

1.1 Forecasting abrupt changes in ecology

Forecasting undesirable change is, arguably, the holy grail of ecology. Paired with an understanding of system interactions, a forecast is ideal if it provides reliable forecasts with sufficient time to prevent or mitigate unwanted systemic change. Early warning systems (or early warning signals, or early warning indicators) have been developed and tested for some ecological systems data (especially marine fisheries time series and for nutrient loading in shallow lakes). Despite the quantitative methods proposed as early warning systems for ecological data (hereafter referred to as regime detection measures, RDMs), many are currently of limited practical utility. This paradox may be a consequence of existing ecological early warning systems (or other quantitative methods for identifying systemic change) having one or more of the following characteristics:

1. not generalizable across systems or system types (especially when it requires a model or a determinsitic function to describe the system)
2. require a large number of observations
3. difficult to implement
4. difficult or to interpret
5. requires an understanding of the drivers of change
6. performs poorly under uncertainty
7. give no uncertaintiy around estimates (tying into interpretation issues)
8. cannot handle noisy data
9. ignores or does not sufficiently account for observation error
10. no baseline with which to compare results
11. no application/testing on empirical systems data
12. systems are subjectively bounded (i.e., components are chosen)
13. being overshadowed by semantics
14. are based on two observations (e.g., before-and-after)
15. cannot link the shift to potential drivers (i.e. the method reduces the dimensionality such that it is unitless and/or loses all relevant information)

Research focusing on the above areas as they relate to RDMs will contribute to the advancement and improvement of existing early warning systems, and will, hopefully, highlight methods which are useful and which are not to practitioners and decision makers.

1.2 Dissertation aims

The overarching aim of this dissertation is to advance our understanding of the utility and limitations of select early warning systems. Specifically, I focus on RDMs capable of analyzing multi-varaible data, including temporally- and spatially-explicit. Although the most widely-applied RDMs proposed in the ecological literature are those deveoped for and tested on single-variable time series (e.g., temperature or fisheries stock time series), the utility of these methods in multi-variable systems (data) is limited. Regime detection metrics for tracking and identifying changes in multivariable systems data are of greater use than single-variable RDMs in systems within which a change manifests dynamically and across multiple variables (e.g., species). Multivariable RDMs may also prove advantageous when the drivers of systemic change are unknown. Further, ecological systems are noisy, and ecological systems data are messy. Although it's taken us many decades to produce realiable weather forecasts 5 days out (and climate is a low-number system), ecologists produce regime detection methods with the promise of predicting high-dimensional ecosystem change in advance. Many of these RDMs are not models, like the weather forecasting models which have taken years to refine.

1.3 Dissertation structure

1.3.1 Chapter overview

The dissertation comprises a brief introduction [Chapter 1], an overview of the myriad regime detection measures used or proposed for use with ecological data [Chapter 2], a detailed guide to Fisher Information as a RDM written for the lay ecologist [Chapter 3], an application of Fisher Information to spatially-explicit data [Chapter 4], introduction of a new regime detection measure, velocity [Chapter 5], a study of data quality and data loss on select regime detectiob measures [Chapter 6], an application of body mass discontinuity analysis to spatially explicit data [Chapter 7], and a conclusions chapter [Chapter 8].

1.3.2 Accompanying software (appendices)

This dissertation is accompanied by the vignettes for two software I created, which are publicly available for use (MIT use and distribution license). The first is `regimeDetectionMeasures` (Appendix 8.5), is an R package for calculting multiple regime detection measures, and the second, `bbsRDM` (Appendix 8.7), is a package which downloads and uses the North American Breeding Bird Survey data to calculate regime detection measures (using `regimeDetectionMeasures`).

Chapter 2

A Brief Overview of the Ecological Regime Detection Literature

2.1 Introduction

> *If a regime shift occurs and no one detects it–is it a regime shift at all?* **No**, if the regime shift is defined as a change in a system which negatively impacts humans. **Yes** if the regime shift is defined simply as a shift in the underlying strucutre of a system.

Long-lasting changes in the underlying structure or functioning of natural systems due to exogeneous forcings (also called regime shifts) is of interest to ecologists. The ability to identify and predict these shifts is particularly useful for systems which are actively managed, provide ecosystem services, or provide benefit to societiy. Despite the utility of identifying and refining the regime detection methods (or early warning signals or indicators), there exists a disparity among the number of methods proposed for detecting abrupt changes in ecological, oceanographic, and climatological systems and the studies evaluating these methods using empircal data (@ Hawkins, Bohn, &

Doncaster, 2015). Further, new methods continue to permeate the literature despite this disparity. Although reviews of regime shift detection methods exist (Andersen, Carstensen, Hernandez-Garcia, & Duarte, 2009; Boettiger, Ross, & Hastings, 2013; Clements & Ozgul, 2018; Dakos et al., 2015a, 2015b; deYoung et al., 2008; Filatova, Polhill, & Ewijk, 2016; Kefi et al., 2014; Litzow & Hunsicker, 2016; Mac Nally, Albano, & Fleishman, 2014; Mantua, 2004; Roberts et al., 2018; S. N. Rodionov, 2005; Scheffer, Carpenter, Dakos, & Nes, 2015), the most comprehensive presentation of available methods as they is outdated (S. N. Rodionov, 2005)*[1]

Perhaps given the sheer number of methods available, there is not currently a single, comprehensive source to which the practical ecologist can refer for identifying potential regime detection measures. Prior reviews of this literature vary in both the number and detail of the methods presented, often focusing on a single aspect of regime shift theory (Andersen et al., 2009), or relaying methods to disciplinary audiences (Roberts et al., 2018). Here I present a brief, systematic review of the methods proposed as what I will call regime detection methods (RDMS) in the ecological literature. I also synthesize the RDMs which are designed to identify ecological regime shifts under uncertainty. I.e., when the regime shift is not hypothesized *a priori*.

Methods proposed for detecting ecological regime shifts (RDMs) are not easily identified using systematic literature review techniques for several reasons. First, the terminology associated with regime shift detection methodologies is highly variable within and among fields (Andersen et al., 2009). For example, the terms, *regime shifts, regime changes* and *tipping points* are variably used in studies of ecological systems, whereas *inhomogeneities* is common in meterology and climatology and *structural change* is largely confined to econometrics. Although semantics vary both within and across disciplines (e.g., a regime shift vs. a structural change), many methods are shared

[1] I also refer the reader to Kefi et al. (2014) and Yin, Leroux, & He (2017) spatial methods, and to Ducré-Robitaille, Vincent, & Boulet (2003) select tests for homogeneity in climate data.

or concurrently applicable. Second, papers introducing a new method or approach to identifying regime shifts are not often proposed in publication outlets with aims of disemminating new quantitative methods (e.g., *Ecological Modelling*, *Methods in Ecology and Evolution*). Rather, many new methods are published in journals with refined (e.g., *Entropy*, *Progress in Oceanography*), as opposed to broader scope scopes (e.g., *Ecology* and *Nature*).

Some RDMs require the use of mechanistic models however some methods fall into the category of model-independent (or model-free), or they require only simple autoregressive (AR) models. In most situations, the practical ecologist will have insufficient data or a limited understanding of the system with which to parameterize even the simplest mechanistic models. The regime detection measures requiring only a limited or no understanding of the mechanisms generating the observed data, I synthesize the utility of these methods here. Further, I synthesize methods not requiring an *a priori* hypothesis about if and where the regime shift occurred.

2.2 Methods

To identify the extent to which these methods are not obvious to the practical ecologist, I conducted a systematic literature review. I attemped to identify original papers which introduce new, quantiative RDMs. Although the review method was to detect as many methodological papers as possible, most RDMs of which I was previously aware were not identified using a systematic technique. Therefore, while highlighting the literature search results, I also provide the missing methods.

2.2.1 1. Identifying regime detection methods

Candidate articles were identified for two reasons: 1) a bibliographic analysis of regime shift relevant papers in ecology and 2) to identify regime detection methods proposed in the literature. The data used for the latter (identify methods) are a subset of the data used for the former (bibliographic analysis).

I first queried the Thomson-ISI Web of Science (WoS) database (on 06 March 2019) to identify articles which mention terms related to regime shifts, or abrupt changes, using the following boolean:

> TS=((“regime shift” OR “regime shifts” OR “regime change” OR “regime changes” OR “catastrophic change” OR “catastrophic shift” OR “catastrophic changes” OR “catastrophic shifts” OR “sudden change” OR “sudden changes” OR “abrupt shift” OR “abrupt shifts” OR “abrupt change” OR “abrupt changes” OR bistab* OR threshol* OR hystere* OR “phase shift” OR “phase shifts” OR “phase change” OR “phase changes” OR “step change” OR “step changes” OR “stepped change” OR “stepped changes” OR “tipping point” OR “tipping points” OR “stable states” OR “stable state” OR “state change” OR “state changes” OR “stark shift” OR “stark change” OR “stark shifts” OR “stark changes” “structural change” OR “structural changes” OR “change-point” OR “change point” OR “change-points” OR “change point” OR “break point” OR “break points” OR “observational inhomogeneity” OR “observational inhomogeneities”) AND (“new method” OR “new approach” OR “novel method” OR “novel approach”))

where ‘*’ indicates a wildcard. Limiting the search to the Web of Science Categories *Ecology* and *Biodiversity Conservation* [adding ‘AND WC=(Ecology OR ’Biodiversity Conservation’) to the above boolean] excludes many methods used solely in climatology, physics, and data science/computer science literatures, where change-point analyses are abundant. Although additional methods could be identified by searching these fields, this dissertation focuses on using methods for analysing *multivariable* data. Consequently, many methods for analysing abrupt breaks in a single longitudinal data are excluded in this review. I further filtered the results to identify articles which propose a 'new' method by retaining papers which included at least one of the following phrases in the title and/or abstract: > 'new method', 'novel method', 'new approach', 'new practical method', 'new simple method', 'new multivariate', 'new tool', 'novel tool', 'novel multivarte', 'novel approach', 'new numerical', 'novel numerical', 'new quantitative', 'novel quantitative', 'i introduce', 'we introduce'

I removed articles from this query based on both prior knowledge and removed those highlighted in previous reviews related to regime detection measures (Andersen et al., 2009; Boettiger et al., 2013; Clements & Ozgul, 2018; Dakos et al., 2015a, 2015b; deYoung et al., 2008; Filatova et al., 2016; Kefi et al., 2014; Litzow & Hunsicker, 2016; Mantua, 2004; Roberts et al., 2018; S. N. Rodionov, 2005; Scheffer et al., 2015).

There appeared disparity among the number of methods of which I was previously aware and those identified in an initial Web of Science review. In an attempt to identify as many new methods as possible I conducted an informal search of the Google Scholar database, a database notoriously broader in scope than other academic dataabses. The length of boolean for the Google Scholar database is limited by the number of characters. Unfortunately, this, coupled with the wide breadth of Google Scholar's search boundaries, limits the capacity to which Google Scholar can be used to refine the literature to a manageable number of articles. For these reasons I arbitrarily skimmed the titles of the first 25 pages of the Google Scholar results (25 pages = 250 articles). It should be noted that the order of terms appearing in the boolean are regarded as the order of desired relevancy. I used the following boolean to identify these articles in Google Scholar: > ('regime shift' OR 'regime change' OR 'tipping point') AND ('new method' OR 'new approach' OR 'novel method' OR 'novel approach')

The candidate articles identified by Google Scholar and Web of Science contained numerous articles proposing a new framework for identifying regime shifts rather than new methods. As this chapter concerns the latter (new methods) I exluded these by removing articles proposing a "new" combination of previously-used methods (see Kong et al., 2017; Seddon, Froyd, Witkowski, & Willis, 2014; Vasilakopoulos, Raitsos, Tzanatos, & Maravelias, 2017). I also did not consider papers which made relatively minor adjustments or recommendations to existing methods (Zhou & Shumway, 2008; but see K. Nicholls et al., 2011 for an addition of variable optimization to the method in @nicholls_detection_2011 that was not included in the results) or articles proposing new methodologies in mathematical journals (Byrski & Byrski, 2016; Salehpour, Gustafsson, & Johansson, 2011) that have yet to be associated with or tested on ecological data, or suggested to be useful for empirical data.

2.2.2 2. Bibliographic analysis of ecological regime shift literature

The stil-vague definition of ecological regime shifts has led to a breadth of articles exploring systemic changes in nature. As such I conducted an exploratory bibliographic analysis of the ecological regime shift literature. This literature search differs from the methods review in a few ways. First, I did not restrict my boolean to include terms that would yield quantitative methods. Further,

I identified candidate articles related to regime shifts and abuprt changes in ecological research using the ISI Web of Science database. Candidate articles were retrieved using a boolean containing words relating to regime shift and those which were restricted to the Web of Science categories *Ecology* and *Biodiversity Conservation*. The boolean used was: > TS=("regime shift" OR "regime shifts" OR "regime change" OR "regime changes" OR "catastrophic change" OR "catastrophic shift" OR "catastrophic changes"

OR "catastrophic shifts" OR "sudden change" OR "sudden changes" OR "abrupt shift" OR "abrupt shifts" OR "abrupt change" OR "abrupt changes") AND WC=("Ecology" OR "Biodiversity Conservation")

I created co-citation and keyword co-occurrence metrics using network analysis to identify patterns and trends in the current state and development of the ecological regime shift literature. In an attempt to undertand the evolution of regime shift methodologies in the ecological (and biodiversity conservation) literature I focus results on keywords and concept themes, rather than citations and author dominance. I used the package R `bibliographix` (Aria & Cuccurullo, 2017) to construct the networks, which uses clustering algorithms to identify patterns in the bibliography.

2.3 Results

2.3.1 1. Identifying regime detection methods

The search boolean for WoS boolean *not* including restriction to fields (WC) 'Ecology' and 'Conservation Biology' yielded over 20,000 results. Restricting to the abovementioned fields created a manageable database from which to filter. This search yielded 2,776 articles. 654 of these papers included terms relating to 'regime shifts' (Figure 2.2), many appearing in the journal *Ecological Modelling* (Figure 2.1). The rate of publication of 'regime shift' articles is not strongly correlated with the rate of papers published in 'Ecology' and 'Biodiversity Conservation' fields (Figure 2.3). Filtering the Web of Science results by including only articles mentioning terms related to 'new method' yielded 202 articles. After removing prior knowledge, only 93 articles remained to be reviewed 'by hand' (i.e., reading the entire paper). Of those reviewed I identified 2 'new' methods (2.4). Similarly, of the 250 articles reviewed from the Google Scholar search, I retained only 3 methods. I was previously aware of an additional 68

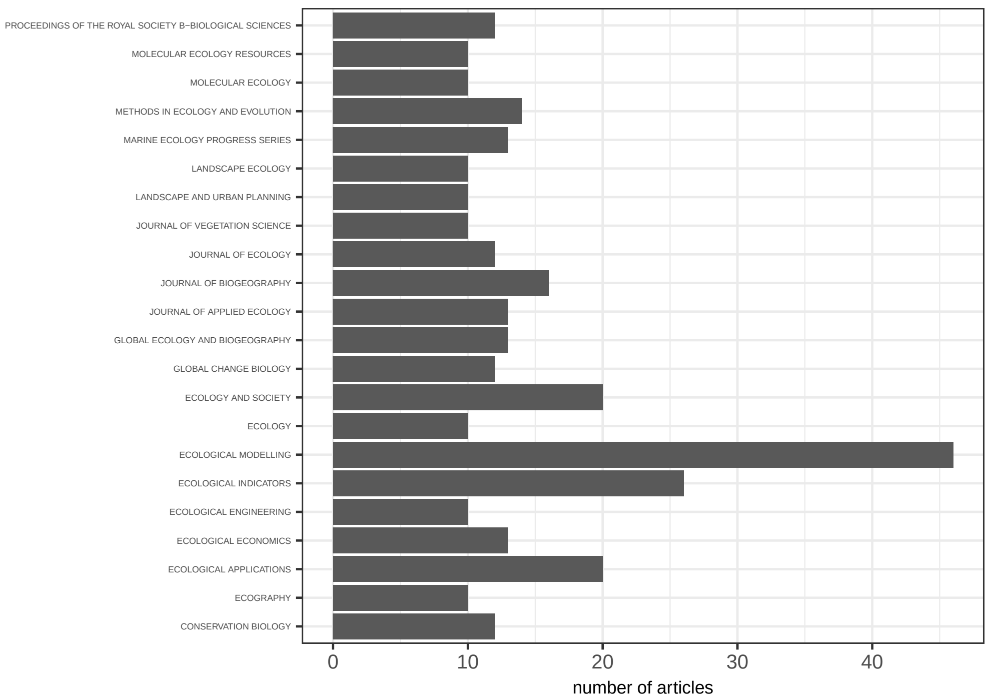


Figure 2.1: Distribution of the 'regime shift' articles for journals with at least 10 articles.

articles containing 'new' methods (2.4), approximately half of which were identified using the abovementioned techniques.

```
Warning in pandoc.table.return(...): Supplied relative values don't add up
to 100%. Reverting to default
```

Table 2.1: List of the regime detection methods identified in this review. (continued below)

Method	Metric type
Characteristic length scale (CLS) estimation	attractor reconstruction
Average standard deviates	metric
BDS test	metric
Coefficient of variation (CV)	metric
Conditional heteroskedasticity	metric

Method	Metric type
Cumulative deviation test (CUSUM)	metric
Degenerate Fingerprinting	metric
Degenerate Fingerprinting	metric
Downton-Katz test	metric
Fisher Information	metric
Intervention Analysis	metric
Inverse of AR(1) coefficient, variance, etc.	metric
Kurtosis	metric
LePage test	metric
Mann-Kendall test	metric
Mann-whitney U-test	metric
Moving detrended fluctuation analysis (MDFA)	metric
Nearest-neighbor statistics	metric
Nikiforiv method	metric
Oerleman's method	metric
Pettitt test	metric
Probability density function entropy method	metric
Quickest detection method (ShiryaevRoberts statistic)	metric
Rodionov method	metric
STARS	metric
Sequential tests/moving windows	metric

Method	Metric type
Signal-to-noise ratio	metric
Skewness	metric
Spectral density ratio indicator	metric
Spectrum indicator	metric
Stability Index of the Ecological Units	metric
Standard deviation (rising variance)	metric
Standard normal homoegeneity	metric
T-test	metric
Threshold Indicator Taxa ANalysis (TITAN)	metric
Variance Index	metric
Wilcoxon rank-sum	metric
dimension reduction techniques (e.g., PCA)	metric
NA	metric
NA	metric
NA	metric
two-phase regression	metric of a model
Zonal thresholding	metric*
Bayesian approaches	model
Convex model	model
Generalized model	model
Multivariable autoregressive models (MAR1)	model

Method	Metric type
Nonparametric drift-diffusion-jump model	model
Potential analysis	model
Regression-based models	model
Self-exciting threshold autoregressive state-space model SETARSS(p)	model
Smooth transition autoregressive model	model
shiftogram	model
Autocorrelation at-lag-1	model-based
Online dynamic linear modelling + time_varying autoregressive state_space models (TVARSS)	models
Clustering, various	NA
Degenerate Fingerprinting	NA
Fourier Analysis	NA
Free-knot splines & piecewise linear modelling	NA
Lanzante method	NA
MCMC	NA
Method 1-TBD	NA
Method 2-TBD	NA
Vector-autoregressive method	NA
Wavelet analysis (decomposition)	NA

Method	Metric type
method-fuzzy synthetic evaluation (FSE)	NA

Source
@NA
@ebbesmeyer19911976
@carpenterBrock2011early
@NA
@seekell2011conditional
@buishand1982some
@held2004detection
@livina2007modified
@karl1987approach
@fath_regime_2003
@francis1994decadal
@carpenter2008leading
@biggs2009turning
@yonetani1993detection
@goossens1987recognize
@mauget2003multidecadal
@he2008new
@pawlowski_identification_2008
@NA
@oerlemans1978objective
@pettitt1979non

Source
@pawlowski_identification_2008
@moustakides2009numerical
@rodionov_sequential_2005
@buishand1982some
@NA
@NA
@guttal2008changing
@biggs2009turning
@NA
@parparov2015quantifying
@carpenter2006rising
@alexandersson1986homogeneity
@ducre2003comparison
@baker2010new
@brock_variance_2006
@karl1987approach
@NA
@ives2003estimating
@NA
@andersen_ecological_2009,
@easterling1995new
@yin2017methods
@jo2016bayesian
@qi2016resilience
@lade2012early

Source
@ives2012detecting
@carpenter2011early
@ives2012detecting
@solow1987testing
@tong1990nonlinear
@see gal2010novel
@groger2011analyses
@vincent1998technique
@parparov2017quantifying
@NA
@kleinen2003potential
@carpenter2010early
@gal2010novel
@lanzante1996resistant
@NA
@manly2006two
@manly2006two
@solow_test_2005
@cazelles2008wavelet
@wang2011application

Using my prior knowledge of the relevant literature and by systematically searching the Web of Science and Google Scholar databases, I identified 66 unique regime detection measures (Figure 2.4; Table **??**).

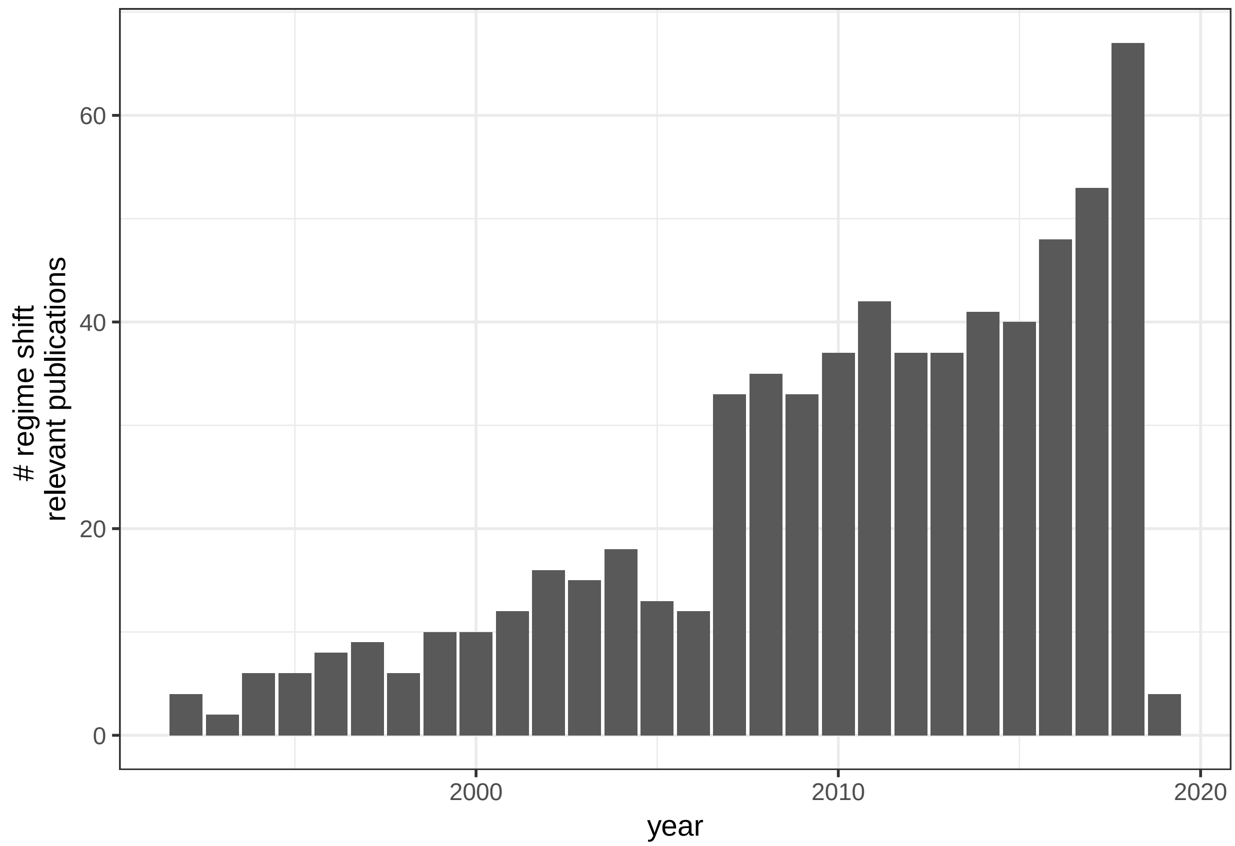


Figure 2.2: Number of publications by year in fields ‘Ecology’ and ‘Biodiversity Conservation’ which included terms related to ‘regime shift’ (total = 654).

2.3.2 2. Bibliographic analysis of ecological regime shift literature

A search of Web of Science for articles in Ecology and Biodiversity Conservation containing phrases related to ‘regime shifts’ yielded 1,636 original articles. These articles were not filtered in any fashion and as such all were considered in the bibliographic analysis.

I used the clustering algorithms of the bibliometrics package to produce a thematic map which uses a clustering algorithm to identify clusters (or themes) based on keywords associated with each article (Cobo, López-Herrera, Herrera-Viedma, & Herrera, 2011). Keywords are supplied both by the au-

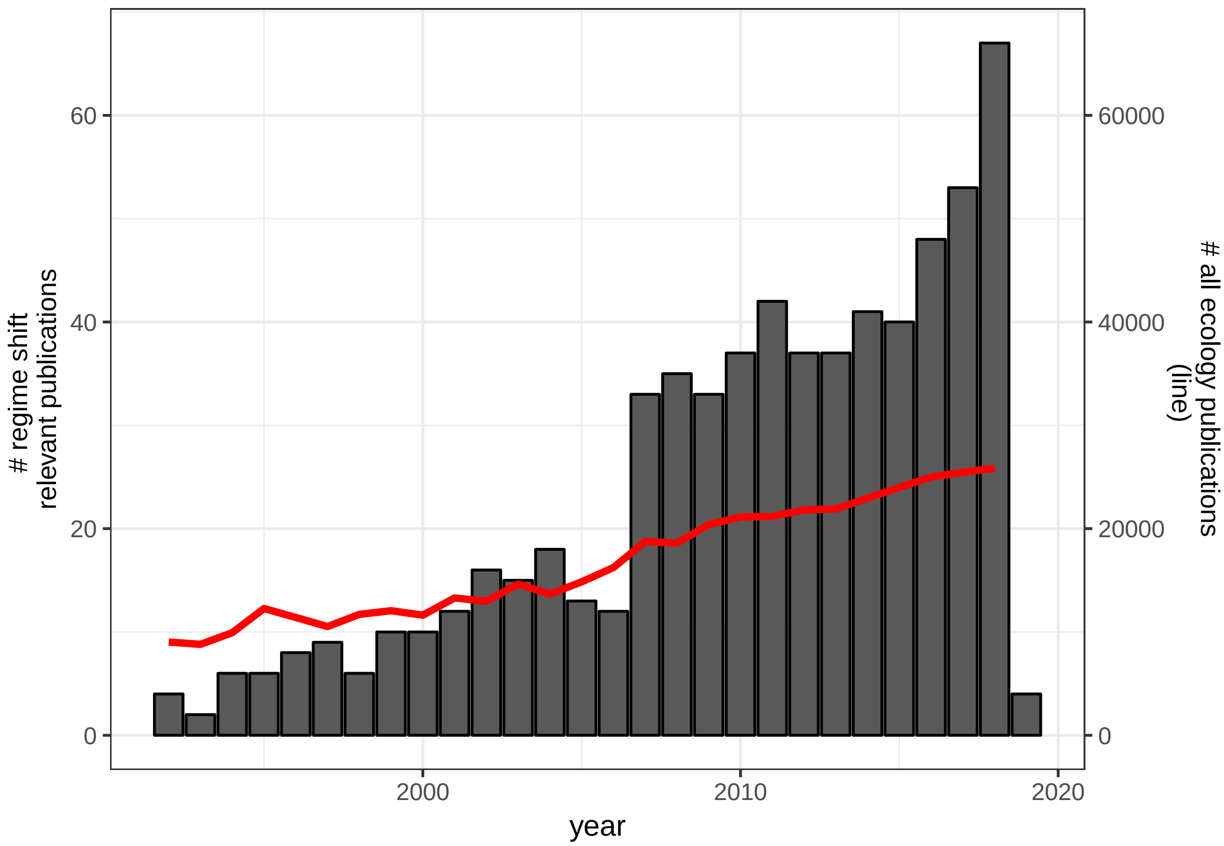


Figure 2.3: Number of publications by year in fields 'Ecology' and 'Biodiversity Conservation' which included terms related to 'regime shift' (total = 654).

thors and by the ISI Web of Science and appear to be used very differently among this literature (Figure @ref(fig:thematicMaps_keyword)). The clustering algorithm identified fewer clusters (themes) in the ISI-keywords (Figure @ref(fig:thematicMaps_keyword)a) than were identified among the author-supplied keywords (Figure @ref(fig:thematicMaps_keyword)b). This pattern is not surprising given the former keywords are restricted to pre-set themes while the authors can often supply any words. The themes identified in the ISI-keyword analysis were relatively consistent as the number of keywords analysed increasded (Figure @ref(fig:thematicMaps_keyword_isi)), but the themes varied drastically among the author-supplied keywords (Figure @ref(fig:thematicMaps_keyword_author)). For this reason I make inference on only the ISI-supplied keyword cluster analysis.

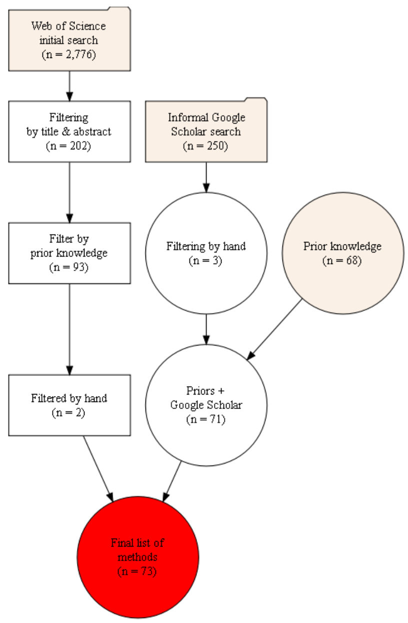


Figure 2.4: Flowchart of the litearture review process for identifying new regime detection methods. *Only the first ten pages (250 articles) of Google Scholar results were examined. Node shapes: folder = unfiltered articles; box = articles actively filtered; diamond = number of articles with new methods.

Four major themes were identified in the ISI keyword analysis and, interestingly, mostly fell within the two extreme quadrants, the first and the third (Figure @ref(fig:thematicMaps_keyword_isi)). The themes identified by ISI keywords were much larger in scope (e.g, dymamics, ecosystems, climate; (Figure @ref(fig:thematicMaps_keyword)a) than those identified in the author keywords (e.g., eutrophication, trophic cascade; Figure @ref(fig:thematicMaps_keyword)b). Regime shifts and ecosystems dynamics are usually have both high centrality and density (Figure @ref(fig:thematicMaps_keyword)b:d), suggesting these two themes are both important to the development of the field and still strongly influence the field. Although dynamics (i.e. non-linearity) plays a central role in the theory of ecological systems this is not reflected in many case studies of regime shifts in application (Litzow & Hunsicker, 2016). Litzow & Hunsicker (2016) found that ~ 50 of case studies using early warning indicators to identify regime shifts in time series actually

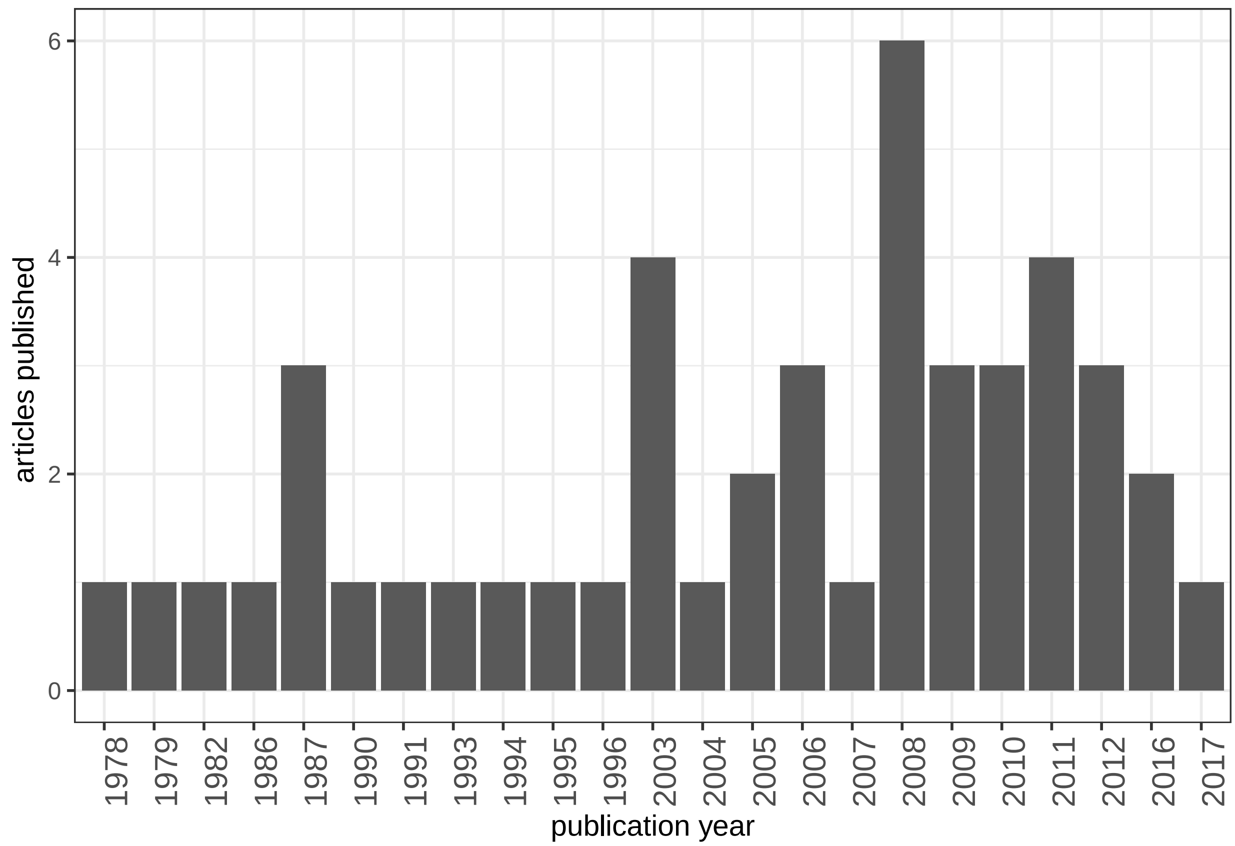


Figure 2.5: Number of methods published over time.

tested and/or accounted for non-linear dymamics in the data.

A few patterns appear in analyses of the intellectual structure of regime shift research in ecology (Figure 2.6). First, although the concept of stability, thresholds, and multiple stable states in ecological systems first appeared (and was well-receieved) in the literature in the 1970s [e.g., Holling (1973);@ may1977thresholds], the most important papers in this field appeared primarily in the early and mid 2000s (Scheffer & Carpenter, 2003; @ carpenter2006rising;@ folke2004regime; Walker, Holling, Carpenter, & Kinzig, 2004; Nes & Scheffer, 2005). The most recent major contributions to the field were concpetual works emphasizing planetery boundaries and tiping points and the impacts of not recognizing these shifts (Hughes, Carpenter, Rockström, Scheffer, & Walker, 2013). Finally, the "rise" of resilience theory (Folke et al., 2004; Walker et al., 2004), the first efforts of a search for early warning indicators of ecological regime shifts (Carpenter & Brock, 2006) and a spur of critique of regime shift detection methods [Andersen et al. (2009);@ contamin_indicators_2009] are recognized in the historiograph.

The clustering algorithm identified the most influential papers in the field (based solely on number of citations) as those published in the late 2000s (Fig 2.7), articles which are broad in-scope and are still used today to frame studies in the context of global change, planetary boundaries, and large-scale tipping points (Bennett, Peterson, & Gordon, 2009; Rockström et al., 2009; Smith & Schindler, 2009). Arguably, the papers that are equally influential include those which correspond with the observed rapid increase in the number of publications (in the early 2000s), Folke et al. (2004) and Scheffer & Carpenter (2003) (Fig 2.7). Numerous reviews of the regime shift

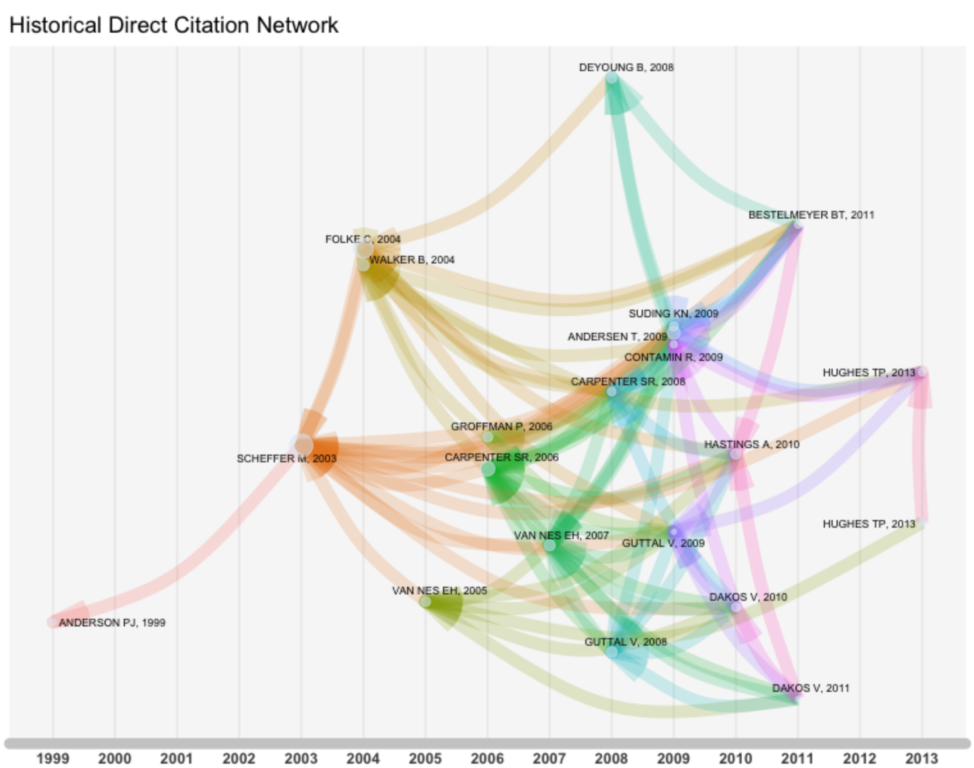


Figure 2.6: Chronological direct citation newtwork suggests the intellectual structure can be mapped to a few papers. This historiograph identifies important works explicitly in chronological, as opposed to absolute, order.

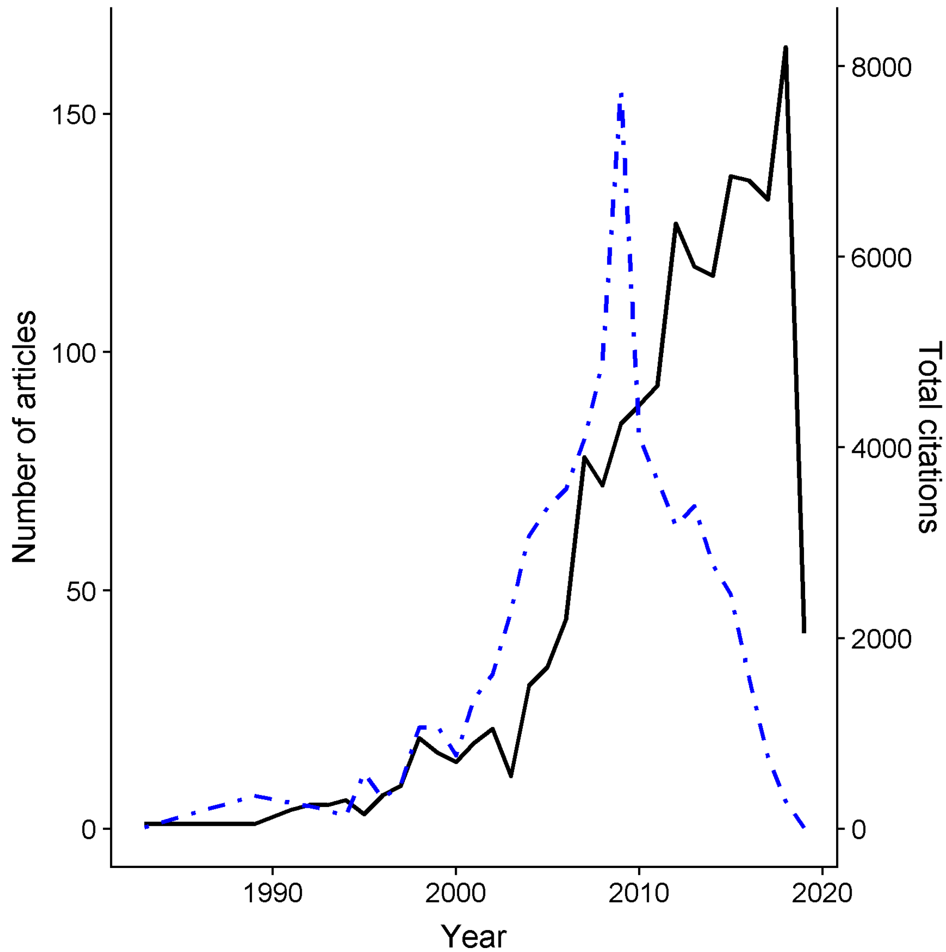


Figure 2.7: Total number of articles published and corresponding number of citations (for papers published that year). The most highly cited papers to-date are those published in the late 2000s.

literature exist, ranging from conceptual reviews of the state of regime shift theory in ecology and application [e.g., Bestelmeyer et al. (2011);@ mac2014scrutiny;@ andersen_ecological_2009], to studies of robustness of early warning indicators under various theoretical and practical conditions [e.g., Dutta, Sharma, & Abbott (2018); Perretti & Munch (2012); Lindegren et al. (2012); Hastings & Wysham (2010a); Figure 2.8]. Further, comprehensive reviews of the ecological regime shift detetion literautre are increasingly out-dated. A permanent and open-source database containing information critical to the testing, comparison, and implementation of RDMs may prove useful to the reader who is interested in applying RDMs but is lacking the

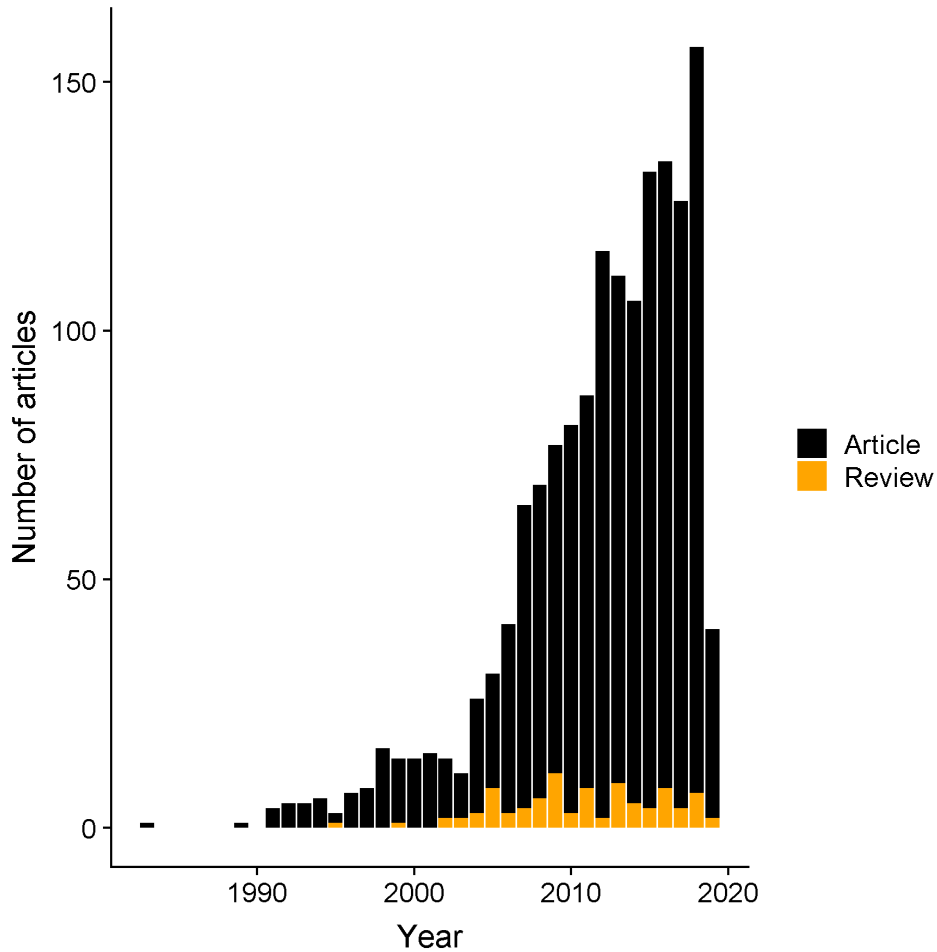


Figure 2.8: Total number of articles published per year by category as categorized by ISI. Book chapters, proceedings, editorials, and letters are excluded.

statistical or mathematical background to do so.

The early warning indicators that are often referred to as, “traditional early warning indicators” (variance, skewness, autocorrelation at lag-1) are fairly well-reviewed, and have been tested under a variety of conditions (Dakos, Carpenter, et al., 2012; @ ditlevsen2010tipping; Lindegren et al., 2012; Boettiger & Hastings, 2012; Dutta et al., 2018; Litzow & Hunsicker, 2016; Perretti & Munch, 2012; Sommer, Benthem, Fontaneto, & Ozgul, 2017). However, many of these works apply the traditional (and other) early warning indicators to simulated data, with only some (Contamin

& Ellison, 2009; Dutta et al., 2018; Guttal, Jayaprakash, & Tabbaa, 2013; Perretti & Munch, 2012) testing under varying conditions of noise and expected shift types. The utility and robustness of the traditional early warning indicators is not consistent across and within sytems, making them of limited utility in situations where the system cannot be reliably mathematically modelled, or where we have limited data [see also Ch. 6]. The authors of most reviews and comparative studies of early warning indicators suggest that no early warning indicator is reliable alone, or that work is needed to understand under what empirical conditions early warning indicators might fail (Clements & Ozgul, 2018; deYoung et al., 2008; Filatova et al., 2016; Kefi et al., 2014).

2.4 A synthesis of the methods available for the practical ecologist

Many of the methods identified in this review have yet to be tested on multiple, empirical data (see Table ??). I categorize the regime detection methods as one of either model-free or model-dependent. Model-free and model-dependent methods are those which do and do not require a mechanistic model to describe the system, respectively. Because many of the model-dependent methods are based on autoregressive modelling apporaches, this is highlighted in the model-dependent section.

2.4.1 Model-depdendent

Model-dependent require a mechanistic (mathematical) representation of the system, models which often carry strict assumptions that are easily violated by empirical systems (Abadi, Gimenez, Arlettaz, & Schaub, 2010). Model-dependent methods are usefully categorized are used under two contexts: differentiable systems of equations or autoregressive. The methods relying on mechanistic models include model descriptions of systems with many, dynamic and interacting components. For example, models are used to reconstruct trophic food webs where prey or predator collapse induces trophic regime shifts in freshwater lake systems (Carpenter et al., 2011).

2.4.2 Model-free

Model-free (or metric-based per Dakos, Carpenter, et al. (2012)) methods are those which do not require a mathematical represntation of the system. In fact, many require much less knowledge about the system component dynamics and their interactions to calculate a results. The utility of these methods vary with respect to the number of state variables encompassed in the method, and are therefore further categorized as either univariate (using a single dimesion) or multivariable (using but not necessarily requiring multiple dimensions).

The most widely used model-free univariate RDMs include descriptive statistics of individual system components (i.e. univariate), such as variance, skewness, and mean value (Andersen et al., 2009; Mantua, 2004; S. Rodionov & Overland, 2005). These univariate methods require only very simple calculations, however, their efficacy in empirical systems analysis is controversial. For example, variance (Carpenter & Brock, 2006) and skewness (of a single variable), oft referred to generally as 'leading indicators' or 'early-warning indicators' in the literature, has been applied to both theoretical and empirical systems data with varying results.

Hastings & Wysham (2010a) point out an important feature of using any methods for identifying regime shifts in empirical system data: we only have a single history within which we can compare AND these metrics which depend on the system exhibiting a change in variance or skewness around a mean value before and after a regime shift require hte system to have a smooth potential, rather than one which can manifest complex dynamics (i.e. non-smooth potential). If we are using RDMs to attempt to forecast and prevent non-smooth or abrupt changes, then there is little justification for relying upon these early warning indicators. Specifcally, these early-warning indicators may be most useful when the system is expected to undergo a transcritical or critical bifurcation before exiting a regime (Lenton, 2011).

Hastings & Wysham (2010a) aptly point out that any realisitic ecological model should include some degree of stochasticity, and when this stochasticity is introduced into the function, the funciton will likely not be differentiable at the point of the regime shift (Graham & Tél, 1984). If a function lacks a gradient along its range, then leading indicators will most likely not indicate the abrupt change in system dynamics alony any paramter.

2.5 Discussion

In this chapter I present an exhaustive list of the regime detection metrics proposed in the literature for analyzing ecological data. Although multiple reviews of regime detection measures exist, they are not comprehensive in their survey of the possible methods. Most reviews have summarized various aspects of regime detection measures. For example, Roberts et al. (2018) summarizes methods capable of handling multiple (c.f. a single) variable, and Dakos et al. (2015b) review only methods designed to detect the phenomenon of critical slowing down. Here, I did not discriminate–rather, I presented an exaustive list of the methods proposed for detecting ecological regime shifts, *sensu lato*, providing a much-needed update to collection provided by S. N. Rodionov (2005); and other review papers (Andersen et al., 2009; Boettiger et al., 2013; Clements & Ozgul, 2018; Dakos et al., 2015a, 2015b; deYoung et al., 2008; Filatova et al., 2016; Kefi et al., 2014; Litzow & Hunsicker, 2016; Mac Nally et al., 2014; Mantua, 2004; Roberts et al., 2018; S. N. Rodionov, 2005; Scheffer et al., 2015).

Leading indicators/regime detection measures which analyze only single variables (e.g., variance, autocorrelation at lag-1) are well-tested on both theoretical and empirical data (e.g. Burthe et al., 2016). Among the most widely used RDMs indicators applied to time-series data include an index of variance, moments around the grand mean (skewness and kurtosis), and critical slowing down (S. Carpenter & Brock, 2011, p. @carpenter2006rising). Although univariate indicators may provide insight into relatively simple systems, their, their reliability as indicators for complex systems is less certain (Burthe et al., 2016, pp. @dutta2018robustness, @perretti2012regime, @sommer2017generic, @bestelmeyer_analysis_2011). Leading indicators can be a reliable warning of impending shift (S. Carpenter & Brock, 2011), Some methods have been applied to early-warning indicators in whole systems (Carpenter et al. 2011), however, it is uncommon to have enough information to build reliable networks or food webs. Consequently, reliably measuring the ecological system at hand is often realistically (and financially) not possible. To be useful to practitioners it may be necessary to move beyond heuristic methods, to methods which supply statistical signifiances or probabilites. And although critiques of some RDMs exist, the rate at which they are rigorously tested do not exceed the proliferation of new methods in the literature. For any method to gain credible traction as a pragmatic tool in ecology, studies shouldl address the concerns of these critiques. These can be addressed using, e.g., bootstrapping, simulations.

In this review I restricted articles to those implying they introduced a 'new method'. Avoiding this potential barrier would have required I review the titles, abstracts, and bodies of over 22,000 articles (Figure 2.4). Alternatively, this may also be ameliorated by searching the relevant literature for *applications* of regime detection measures to ecological data, however, I suspect this would similarly yield a large number of articles to review. Also, only a handful of methods have been introduced to the mainstream methodological journals in ecology (e.g., *Ecological Modelling*, *Methods in Ecology and Evolution*; Figure 2.9). Although many mainstream publications (e.g., *Science*, *Ecology Letters*) include applications of some of the methods identified in this chapter (Table **??**), I argue that celebrity and 'new and shiny' (Steel, Kennedy, Cunningham, & Stanovick, 2013) methods may influence which methodological articles are printed in these popular journals. A critical survey of potetial and realized applications

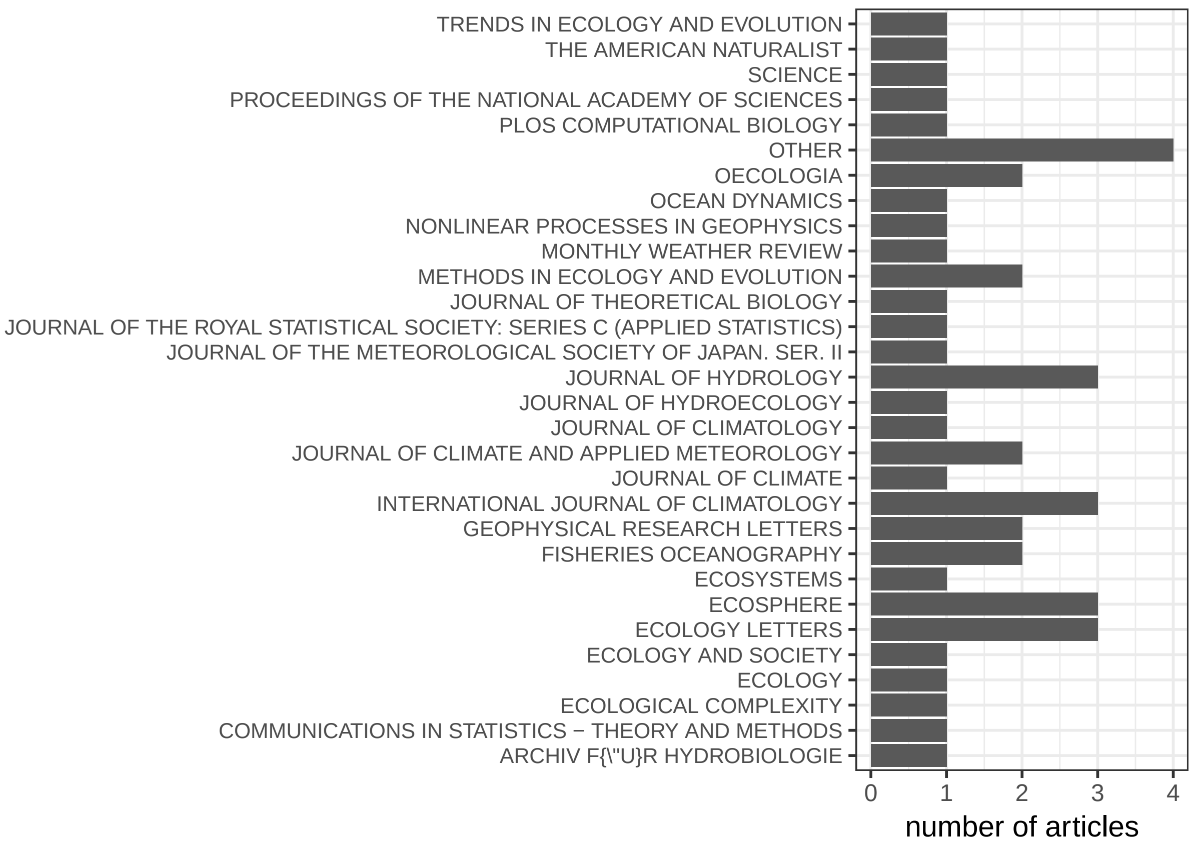


Figure 2.9: Distribution of identified methods across publications via the literature review.

of these methods would be useful for highlighting the needs of future research and methodological improvements. Many of the methods presented in Table **??** have either not been applied to empirical data at all, or were tested only once, often but not always in the article introducing or adapting the methodology (Hawkins et al., 2015). Some methods, especially those dubbed 'early warning indicators' (variance, autoregressive model coefficients) have become relativley mainstream in their application to empirical data, despite having been shown to be less robust in noisy and nonlinear systems (Burthe et al., 2016), in systems exhibiting lag-effects (Guttal et al., 2013), and in systems not exhibiting catstrophic shifts (Dutta et al., 2018). Unlike these early warning indicators, fewer efforts have been made to test robustness under these and more simple conditions (Dutta et al., 2018; c.f. Brock & Carpenter, 2010; Benedetti-Cecchi, Tamburello, Maggi, & Bulleri, 2015). In addition to the paucity of studies attempting to understand the limitations of these methods, this review suggests that simply identifying the suite of methods used in ecological regime shift detections may be difficult using traditional review methods. Many of the methods metnioned in this review were not identified using a systematic search process in Web of Science and Google Scholar–rather, they were methods of which I was either previously aware and/or highlighted in the few methods reviews (Andersen et al., 2009; Boettiger et al., 2013; Clements & Ozgul, 2018; Dakos et al., 2015a, 2015b; deYoung et al., 2008; Filatova et al., 2016; Kefi et al., 2014; Litzow & Hunsicker, 2016; Mantua, 2004; Roberts et al., 2018; S. N. Rodionov, 2005; Scheffer et al., 2015). To facilitate this process, an online, comprehensive database may prove useful to the practical ecologist.

Many of the methods identified in this review have yet to be tested on multiple, empirical datum (see Table **??**). I categorize the regime detection methods as one of either model-free or model-dependent. Model-free and model-dependent methods are those which do and do not require a mechanistic model to describe the system, respectively. Because many of the model-dependent methods are based on autoregressive modelling apporaches, this is highlighted in the model-dependent section (however most autoregessive models are non-specific).

Model-dependent require a mechanistic (mathematical) representation of the system, models which often carry strict assumptions that are easily violated by empirical systems (Abadi et al., 2010). Model-dependent methods are usefully categorized are used under two contexts: differentiable systems of equations or autoregressive. The methods relying on mechanistic models include model descriptions of systems with many, dynamic and interacting components. For example, models are used to reconstruct trophic food webs where prey or predator collapse induces trophic regime shifts in freshwater lake systems.

Model-free (or metric-based per Dakos, Carpenter, et al., 2012) methods are those which do not require a mathematical represntation of the system. In fact, many require much less knowledge about the system component dynamics and their interactions to calculate a results. The utility of these methods vary with respect to the number of state variables encompassed in the method, and are therefore further categorized as either univariate (using a single dimesion) or multivariable (using but not necessarily requiring multiple dimensions). The most widely used model-free univariate RDMs include descriptive statistics of individual system components (i.e. univariate), such as variance, skewness, and mean value (Andersen et al., 2009; Mantua, 2004; S. Rodionov & Overland, 2005). These univariate methods require only very simple calculations, however, their efficacy in empirical systems analysis is controversial. For example, variance (Carpenter & Brock, 2006) and skewness (of a single variable), oft referred to generally as ‘leading indicators’ or ‘early-warning indicators’ in the literature, has been applied to both theoretical and empirical systems data with varying results.

Hastings & Wysham (2010a) point out an important feature of using any methods for identifying regime shifts in empirical system data: we only have a single history within which we can compare AND these metrics which depend on the system exhibiting a change in variance or skewness around a mean value before and after a regime shift require hte system to have a smooth potential, rather than one which can manifest complex dynamics (i.e. non-smooth potential). If we are using RDMs to attempt to forecast and prevent non-smooth or abrupt changes, then there is little justification for relying upon these early warning indicators. Specifcally, these early-warning indicators may be most useful when the system is expected to undergo a transcritical or critical bifurcation before exiting a regime (Lenton, 2011). Hastings & Wysham (2010a) aptly point out that any realisitic ecological model should incorporate some degree of stochasticity, and when this stochasticity is introduced into the function, the function will likely not be differentiable at the point of the regime shift (Graham & Tél, 1984). In other words, most (if not all) ecological systems have non-smooth potentials, and many of the current methods for identifying regime shifts assume otherwise, often failing if the assumption is violated.

To make the regime detection measures more available and transparent to the practical ecologist, I recommend the following: 1. consitent use of fewer methods 1. persistent collection and maintenance of baseline data (reference data) 1. an on-line database of all methods - open-sourced - linked to the original sources (in ecology and statistics or mathematics) - linked to applications 1. a critical review of the current state of methods in ecology - including methodological advancements - especially highlighting where the method fails to perform - including historical tracking of specific methods to identify which may need to be retired, rather than resuscitated 1. more empirical applications of these methods (especially of those only tested on toy and experimental data) 1. relation of RDMs in ecology to other fields (computer science, data science, climatology and oceanography)

I suggest (Table 2.3) a suite of questions which may be useful in a critical review of the characteristics, rigor, and application potential of methods in the context of ecological regime shift detection.

Table 2.3: Potential questions for a comprehensive review of the ecological regime detection metrics literature.

Type	Questions
Methodological	Does the method assume smooth potential?
	Does the regime shift need to be identified *a priori*?
	What are the major assumptions about the distribution of the original data?
	Does the method explicitly assume the system/variable is stationary?
	Does the performance of the method change with non-stationarity?
	Can the method handle unstructured data?
	Can the method handle multiple regime shifts?
	What types of regime shifts can the method detect (e.g., stochastic resonance, slow-fast cycles, noise-induced transition)?
	Is it a model- or metric-based method?
	Does it have forecasting potential?
	Can the method handle uneven sampling?
	What are the minimum data requirements (resolution, extent, number of observations)?
	How does the method handle missing data (e.g., new invasions)?
	Does the method assume Eulerian or Lagrangian processes?
Ecological	Does the system *have* smooth potential?
	Has the method been tested on empirical data? If so, to what rigor?
	What is the impact of losing state variables on long-term predictions (e.g., species extinction)?
	Can the method identify drivers?
	What assumptions does the method make about the system?

	What types of regime shifts are possible in the system?
	Are regime shift(s) suspected *a priori*? What lag(s) exist in the data (system)? Would a positive forecast change management action? Do predictions translate to other systems? Can we interpolate data if necessary? If so, what does this mean for inference?
	In which discipline(s) beyond ecology has the method been tested?

Chapter 3

Decoupling the Calculation of Fisher Information

This chapter is intended for submission to the publication Methods in Ecology and Evolution.[1]

3.1 Abstract

Ecological regime shifts are increasingly prevalent in the Anthropocene. The number of methods proposed to detect these shifts are on the rise, yet few are capable detecting regime shifts without a priori knowledge of the shift, and fewer are capable of handling high-dimensional, multivariate and noisy data. A variation of Fisher Information has been proposed as a method for detecting changes in the "orderliness" of ecological systems data. Although this method is described and applied in numerous published studies, its calculation and the concepts behind its calculation are not clear. Here, I succinctly describe this calculation using a two-species predator-prey

[1] Co-authors include: N.B. Price, A.J. Tyre, D.G. Angeler, T. Eason, D. Twidwell, and C.R. Allen

model. Importantly, I demonstrate that the actual equation for calculating Fisher Information metric comprises fewer steps than was previously described, by decoupling the dimensionality-reduction component from the actual Fisher Information calculation component. I hope this work will serve as a reference for those seeking to understand Fisher Information in the context of ecological systems and regime shifts, and will stimulate further research of the efficacy of these composite regime shift detection metrics..

3.2 Introduction

Changes in the feedback(s) governing ecosystem processes can trigger unexpected and sometimes undesirable responses in environmental conditions (Scheffer, Carpenter, Foley, Folke, & Walker, 2001; Walther et al., 2002). Ecologists often refer to such changes as regime shifts, but this term is used interchangeably in the literature with state change, state transition, or alternative state (Andersen et al., 2009). Climate change and globalization are triggering novel and unexpected changes in ecosystems (Hughes, 1994; Parmesan, 2006; Scheffer et al., 2001; Walther et al., 2002), and the rapidity with which these changes occur make predictive modeling difficult. Although detecting regime shifts is increasingly difficult as we increase the extent and complexity of the system in question (Jorgensen & Svirezhev, 2004), advances in the collection and analysis of ecological data (La Sorte et al. 2018) may improve our ability to detect impending regime shifts in time for intervention (Carpenter et al., 2011; deYoung et al., 2008; Groffman et al., 2006; Jorgensen & Svirezhev, 2004; Sagarin & Pauchard, 2012; Wolkovich, Cook, McLauchlan, & Davies, 2014).

Numerous quantitative approaches have been proposed as regime shift detection methods (Clements & Ozgul, 2016 ; Mantua, 2004; S. Rodionov & Overland, 2005, p.

@andersen_ecological_2009), but few are consistently applied to terrestrial ecological data (deYoung et al., 2008). I broadly classify these methods as either model-based or model-free [Boettiger & Hastings (2012); Hastings & Wysham (2010b); Dakos, Carpenter, et al. (2012). Model-based methods use mathematical (mechanistic) representations of the system (Hefley, Tyre, & Blankenship, 2013), which often carrying strict assumptions that are easily violated by dynamic systems such as ecosystems (Abadi et al., 2010). Further, model misspecification may yield spurious results (Perretti, Munch, & Sugihara, 2013). Model-free (or metric-based, per Dakos, Carpenter, et al., 2012) regime detection methods require fewer assumptions to implement than do model-based methods, and typically require much less knowledge (if any) about system component interactions. The most widely used model-free methods include both descriptive statistics of individual system components, such as variance, skewness, and mean value (Andersen et al., 2009; Mantua, 2004; S. Rodionov & Overland, 2005) and composite measures of multiple variables, notably principal components analysis (Möllmann, Folke, Edwards, & Conversi, 2015; Petersen et al., 2008), clustering algorithms (Beaugrand, 2004), and variance index (Brock & Carpenter, 2006).

3.2.1 Fisher Information as a Regime Detection Method

A method which has been more recently applied in the analysis of ecological and social-ecological systems is Fisher Information (Cabezas & Fath, 2002; Karunanithi, Cabezas, Frieden, & Pawlowski, 2008). As a multivariate, model-free method, Fisher Information integrates the information present in the entire data of a system and distills this complexity into a single metric. This allows Fisher Information to capture ecosystem dynamics with higher accuracy than univariate-based metrics, which frequently fail to detect regime changes (Burthe et al., 2016). However, despite the potential of this method its mathematical underpinnings – specifically its calculation and the concepts behind its calculation– are not clear. In this paper, I address this knowledge gap. I first provide an overview of the method and highlight the need to account for scaling properties, an inherent feature in complex systems. I then succinctly describe the decoupling of the dimensionality-reduction component from the actual Fisher Information calculation component using a two-species predator-prey model. I finally discuss the results from a theoretical viewpoint and its practical utility for predicting regime shifts, an increasing concern motivated by current rates of fast ecological change.

3.2.2 The Sustainable Regimes Hypothesis

Fisher Information (hereafter, FI; Fisher, 1922) is a model-free, composite measure of any number of variables, and is proposed as an early warning signal for ecological regime shift detection and as a measure of system sustainability (Eason & Cabezas, 2012; Eason et al., 2014a; Karunanithi et al., 2008; Mayer, Pawlowski, Fath, & Cabezas, 2007). Three definitions of FI in this context exist: (i) a measure of the ability of the data to estimate a parameter, (ii) the amount of information extracted from a set of measurements (Frieden, 1990; Roy Frieden, 1998), and (iii) a measure representing the dynamic order/organization of a system (Cabezas & Fath, 2002). Although definitions (i) and (ii) are widely applied in the statistical and physical sciences, I focus on definition (iii) as it is gaining traction as a tool to analyze used in the context ofeco ecological systems analysisresponses to fast environmental change. The application of FI to complex ecological systems was posed as part of the "Sustainable Regimes Hypothesis," stating a system is sustainable, or is in a stable dynamic state, if over some period of time the average value of FI does not drastically change (Cabezas & Fath, 2002). This concept can be described using an ecological example. Consider the simple diffusion of a population released from a point source at $t = 0$. This process can be described by a bivariate normal distribution, $p(x, y|t)$. As the time since release, t, increases, the spread of the distribution, $p(x, y|t)$, disperses because the animals have moved further from the release location. As the animal moves away from the release location, the potential area within which it currently occupies will increase with time. In this example, FI will decrease in value as t increases because $p(x, y|t)$ contains less information (higher uncertainty) about where the animals will be located. If we assume constant dispersal, as $t \to \infty$ the animals will be relatively uniformly distributed across the environment and $p(x, y|t)$ will carry no information about the location of the animals. Consequently, as $t \to \infty$ FI approaches zero (no information). Per the Sustainable Regimes Hypothesis (Cabezas & Fath, 2002), this example system is not in a stable dynamic state over the range of t, since FI decreases with time.

Conversely, if a population following a simple logistic growth model, $\frac{dN}{dt} = rN(1 - \frac{N}{K})$, varies around some carrying capacity, K, and the average system parameters (r, K, and their variances σ_r, σ_k) are stationary, then the logarithm of the population size should follow a normal distribution, $N\ normal(\mu, \sigma)$. In this situation, the FI measured over any selected window of time will be relatively constant and, per the Sustainable Regimes Hypothesis, indicates the system is in a stable dynamic state. Further, this Hypothesis posits that a perturbation to N will also not affect FI so long as the perturbation occurs with a stationary probability distribution and if the perturbation does not change the distributions of r and K.

3.2.3 Fisher Information Requires Dimension Reduction

An important feature of the FI method is that it requires a complete reduction in dimensionality (i.e., from >1 to 1 system component). For example, a recent application of Fisher Information to empirical data condensed a species pool from

109 species time series into a 1-dimensional time series (Spanbauer et al., 2014). A reduction in dimensionality, i.e. condensing multivariate data into a single metric, of over two orders of magnitude likely involves a large loss of relevant information, raising the questions of what information is preserved during the dimensionality reduction step in calculating FI, what is lost, and whether this is important. Other dimension reduction techniques, e.g., principal component analysis (PCA) and redundancy analysis (RDA), attempt to preserve the variance of the data, and the number of components scales with the dimensionality of the data (i.e. they are scale explicit). In contrast, by reducing entirely the dimensionality of the data, the FI method does not identify which features of the data are preserved, and the dimensionality does not scale with the dimensionality of the original data.

3.2.4 Aims

The key contribution of this study is that I decouple the dimensionality reduction step of the FI method (Step 1) from the statistical analysis step (Step 2). By isolating the dimensionality reduction step, we can evaluate it based on its own merits and relate it to more well-known and established methods of dimensionality reduction. By isolating the statistical analysis step, one can better understand how Fisher Information is calculated on the single-dimensional data. I believe that this decoupled approach will eliminate some confusion regarding the calculation of FI, allowing interested researchers to readily evaluate the merits of this method. To facilitate our explanation of the method, I reproduce the predator-prey analysis used in (Fath, Cabezas, & Pawlowski, 2003; Mayer et al., 2007), then induce a "regime shift" into the model. I hope this work will serve as a useful explanation of the FI metric for those seeking to understand it in the ecological regime shift context and will stimulate research using this and other multivariate, model-free, and composite measures to understand ecological regime shifts.

3.3 Methods

3.3.1 Predator-Prey Model System

Our model system is a two-species predator-prey model (Eq. (3.1); Fath et al., 2003; Frieden & Gatenby, 2007; Mayer et al., 2007), hereafter referred to as the "model system":

$$\begin{aligned} dx_1 &= g_1 x_1(1 - \tfrac{x_1}{k}) - \tfrac{l_{12} x_1 x_2}{1+\beta x_1} \\ dx_2 &= \tfrac{g_{21} x_1 x_2}{1+\beta x_1} - m_2 x_2) \end{aligned} \tag{3.1}$$

The specified parameters for the model system are $g_1 = m_2 = 1$,$l_{12} = g_{12} = 0.01$,$k = 625$,and $\beta = 0.005$ (Fath et al., 2003; Frieden & Gatenby, 2007; Mayer et al., 2007). The initial conditions (predator and prey abundances,) for the model system were not provided in the original references (Fath et al., 2003; Mayer et al., 2007). I used package `deSolve` in Program R (version 3.3.2) to solve the model system (Eq. Eq. (3.1)), finding $x_1 = 277.781$5 and $x_2 = 174.551$ to provide reasonable results. A complete cycle of this system corresponds to 11.145 time units.

3.3.2 Inducing a Regime Shift

Mayer et al. (2007) calculated FI for a predator-prey system for several discrete values of carrying capacity of prey. The results of this study showed that FI was different for systems with different carrying capacities (K). However, this study did not address the central question of **FI behavior during a regime shift**. As an extension of the original study, I simulated a regime shift by modeling an abrupt decline in carrying capacity, k. I assume k is described by Eq. (3.2) where k_1 is the initial carrying capacity, k_2 is the final carrying capacity, t_shift is the time the regime shift occurred, and α is the parameter controlling the rate (slope) of the regime shift. The hyperbolic tangent function (see Eq. (3.2)) simulates a smooth and continuous change in k while still allowing for the regime shift to occur rapidly. I incorporate the change in k into our system of differential equations by defining the rate of change in k, $k'(t)$, given by (Eq. (3.2)). I assume $k_1 = 800$ and $k_2 = 625$, values corresponding to the range of carrying capacities explored by Mayer et al. (2007). I simulated a time series of 600 time units, introducing a regime change after 200 time units, and $\alpha = 0.05$.

$$\begin{aligned} k(t) &= k_1 - 0.5(k_1 - k_2)(\tanh(\alpha(t - t^*)) + 1) \\ k'(t) &= 0.5\alpha(k_1 - k_2)(\tanh(\alpha(t - t^*))^2 + 1) \end{aligned} \tag{3.2}$$

3.3.3 Decoupling the Steps for Calculating Fisher Information

Two methods exist for calculating Fisher Information (FI) as applied to ecological systems data to which I refer the “derivatives-based” method (first appearing in Cabezas & Fath (2002) and the binning" method (first appearing in Karunanithi et al. (2008)). Although the binning method is proposed as an alternative to the derivatives-based method for handling noisy and sparse data, our decoupling method reveals it may be an unnecessary method. Therefore, I focus on only the derivatives-based method for explaining the theoretical basis for the FI method. The general form of FI can be found in (Fath et al., 2003; Mayer et al., 2007) and I refer the reader to (Cabezas & Fath, 2002).

Step 1: Dimensionality Reduction. The key idea of the dimensionality reduction step is to calculate the Euclidean distance travelled in phase space. In phase space, each coordinate axis corresponds to a system state variable (e.g., number of predators and number of prey). The state of the model system over time describes a path or trajectory through phase space. The distance travelled represents the cumulative change in state relative to an arbitrary starting point in time. For the model system, the distance travelled in phase space can be obtained by solving the differential equation given by Eq. (5.3)

$$\frac{ds}{dt} = \sqrt{(\frac{dx_1}{dt}^2 + \frac{dx_2}{dt}^2)} \tag{3.3}$$

The original motivation for the dimensionality reduction step is that, under restrictive conditions, there is a one-to-one mapping between the state of the system (s), defined in a multidimensional phase space, and the distance travelled, a one-dimensional summary (Cabezas & Fath, 2002). To relate this abstract idea to a more familiar situation, we draw an analogy between the path traced by the system in phase space and the path of a car over the course of a trip. The distance travelled by the car over time is related to the position of the car. Given the route of the car, we could determine the location of the car at any point in time if we know how far it has travelled. However, the distance travelled provides no information about the proximity of locations (i.e., system states). For example, two points in phase space may be arbitrarily close, but the distance travelled would be different if these system states occur at different points in time. Moreover, if the system revisits the same state twice then the one-to-one mapping breaks down and a single state maps to potentially very different values of distance travelled.

What is preserved in the calculation of distance travelled is the rate of change of the system (e.g., the speed and acceleration of the car). The rate of change of the system is the first derivative of the distance travelled in phase space. This is an important point because the concept of a "regime shift" is often associated with the idea of a sudden change in system state. Therefore, it may not be unreasonable to employ a dimensionality reduction procedure that preserves these system dynamics.

Step 2: Statistical Analysis. The product of **Step 1** is a one-dimensional time series of what I call "distance travelled", s, (in phase space). The variable s is referred to as "Fisher variable s" and "represent[s] a particular state of phase space" in the FI literature (Mayer et al., 2007). I believe distance travelled (s) is more descriptive than "Fisher Variable s" and avoids confusing the state of the system, defined in multiple dimensionsby the multivariate data , with the one-dimensional summary. Using this measure, we next calculate the probability of observing the system in a particular state by assuming a one-to-one mapping between distance travelled and the system state. That is, we calculate the probability of observing the system at a particular distance, $p(s)$, along the trajectory for some period of time from 0 to t_end. The time at which we observe the system is assumed to be a random variable, $T_{obs} \sim Uniform(0, t_end)$. This approach assumes the system is deterministic and is observed without error. However, the observed distance travelled by the system, s, is a random variable because it is a function of the random observation time.

Importantly, the probability of observing the system at a particular value of s increases if the system is changing slowly at that point in time. That is $p(s)$ is inversely proportional to the system rate of change, s'. Mathematically, the distance travelled in phase space, s, is a monotonically increasing function of time and we assume it is differentiable. Therefore, the probability density function of the distance travelled is $p(s) = \frac{1}{T}\frac{1}{s'}$, where $s' = \frac{ds}{dt}$ is the speed (or velocity) of s, and T is the time interval over which the system was observed (t_{start}-t_{end}). We note that $p(s)$ is simply a constant multiplied by the inverse of the speed of the system.

The original motivation for the FI calculation as applied to ecological systems was the hypothesis that "since Fisher Information is a measure of the variation" it is also "an indicator of system order, and thus system sustainability" (Cabezas & Fath, 2002). Equation (3.4) is a general form of FI and Equation (4.4) is the form used in the derivative-based method for FI (see eq. 7.3b and 7.12 in Mayer et al., 2007). To better understand the FI calculation, note that Eq.(4.4) is, in part, a measure of the gradient content of the probability density function. As the probability density function becomes flatter, the FI value will decrease. In this way, the FI calculation is closely related to the variance. In fact, the FI value for a normal distribution calculated according to Eq. (4.4) is simply one over the variance. It is also important to note that FI is zero for a uniform distribution, as the probability density function is flat. Note also that FI goes approaches inf if the system is not changing over some period of time (Eq. (4.4)).

$$I = \int \frac{ds}{p(s)} \left[\frac{dp(s)}{ds} \right]^2 \tag{3.4}$$

$$I = \frac{1}{T} \int_0^T dt \left[\frac{s''^2}{s'^4} \right]^2 \tag{3.5}$$

##Results Distance travelled (s), speed ($\frac{ds}{dt}$), and acceleration ($\frac{d^2s}{dt^2}$) capture the dynamics of the model system [Eq. (3.1); Fig. @ref(fig:distSpeedAccel]. I simulated a regime shift in the carrying capacity of this model system, at approximately $t = 200$ (Fig. 3.2). The location of this regime shift with respect to the system trajectory in phase space over the entire simulated time period is shown in (Fig. 3.3). Although a slight change is captured by s (Figure 4) at the location of this regime shift, it is not pronounced. Although the distance travelled, s (Fig. 3.4) changes fairly smoothly around the location of the regime shift, the system exhibits a steep decline in speed ds/dt soon after the induced regime shift (Fig. 3.5).

I calculated FI for the distribution of s over a series of non-overlapping time windows. According to Mayer et al. (2007) the length of the time window should be equal to one system period such that FI is constant for a periodic system, however, the system periods are not identical before, during, and after the regime shift. Therefore, I performed two separate calculations of FI using window sizes corresponding to the initial (when $t < 200$) and the final ($t > 200$) periods of the system ($winsize = 13.061$ and 11.135 time units, respectively). Using these window sizes the drop in FI at the regime shift initiation is bigger than the magnitude of the fluctuations preceding it (Fig. 3.6).

3.4 Discussion

Part of the appeal of the FI method of regime shift detection is that it provides a 1-dimensional visual summary of system "orderliness". However, I have demonstrated that the dimensionality reduction step can be performed separately from the calculation of FI. The rate of change of the system (velocity, $\frac{ds}{dt}\}$), on which FI method is based, is also a 1-dimensional quantity. In the simple predator-prey example, calculating and plotting FI did not provide a clear benefit over simply plotting the system rate of change directly. I suggest that future research uncouple the dimensionality reduction step and the FI calculation step in order to better illustrate the benefits of the FI method relative to dimensionality reduction alone. In the predator-prey example, I assumed the data was free from observation error. Despite these ideal conditions, the estimated FI had high variation and the results depended on the size of the time window used in the calculation. This issue arises because the period of the cyclic system is changing during the regime shift such that it is difficult to find a single window size that works well for the entire time series. Mayer et al. (2007) describe this as a "confounding issue" related to "sorting out the FI signal of regime change from that originating from natural cycles" and suggest using a time window that is large enough to include several periods. However, in the absence of a quantitative decision rule defining what changes in FI indicate regime shifts, it is difficult to separate the signal in the FI metric from the noise due to fluctuations in the natural cycles.

Further research is needed to define quantitative decision rules for what changes in FI constitute a regime shift.

The example used in this study is unrealistic in that I assume no measurement error and therefore focus on the "derivatives-based" method of calculating FI. However, our analysis also has implications for the "binning" method of calculating FI that was later developed for high-dimension noisy data (Karunanithi et al. (2008)). Rather than attempting to estimate the rate of change of each system component (e.g., hundreds of species) and combining these estimates to get the total system rate of change, I suggest an approach where the dimensionality of the data is first reduced by calculating distance travelled in phases-pace and then only a single rate of change is estimated. The advantage of this approach is that for an n-dimensional system it only requires the estimation of one derivative rather than n-derivatives . The drawback to this approach is that noisy observations will likely introduce some bias into the estimate of the system rate of change. Nonetheless, I believe this approach is worth exploring due to its simplicity relative to the "binning" method. The Fisher Information of an n-dimensional system is a vector of unitless values which can only be compared within a dataset (e.g., within a single community time series) and interpreting FI is still largely a qualitative effort (Fath et al., 2003; Mantua, 2004), not unlike most regime detection methods [Ch. 2]. When the FI of a system is increasing, the system is said to be moving toward a more orderly state, and most studies of FI propose that sharp changes in FI, regardless of the directionality of the change, may indicate a regime shift (Cabezas & Fath, 2002; Karunanithi et al., 2008; Spanbauer et al., 2014). Although the aforementioned and numerous other works interpret FI in this context (e.g., Eason et al., 2014a; Eason & Cabezas, 2012), I suggest future work which clearly identifies the ecological significance of the Fisher Information metric and its significance within the ecological regime shift paradigm.

3.5 Acknowledgements

I thank H. Cabezas and B. Roy Frieden for early discussions regarding the development of Fisher Information, and T.J. Hefley for comments on an earlier draft. This work was funded by the U.S. Department of Defense's Strategic Environmental Research and Development Program (project ID: RC-2510).

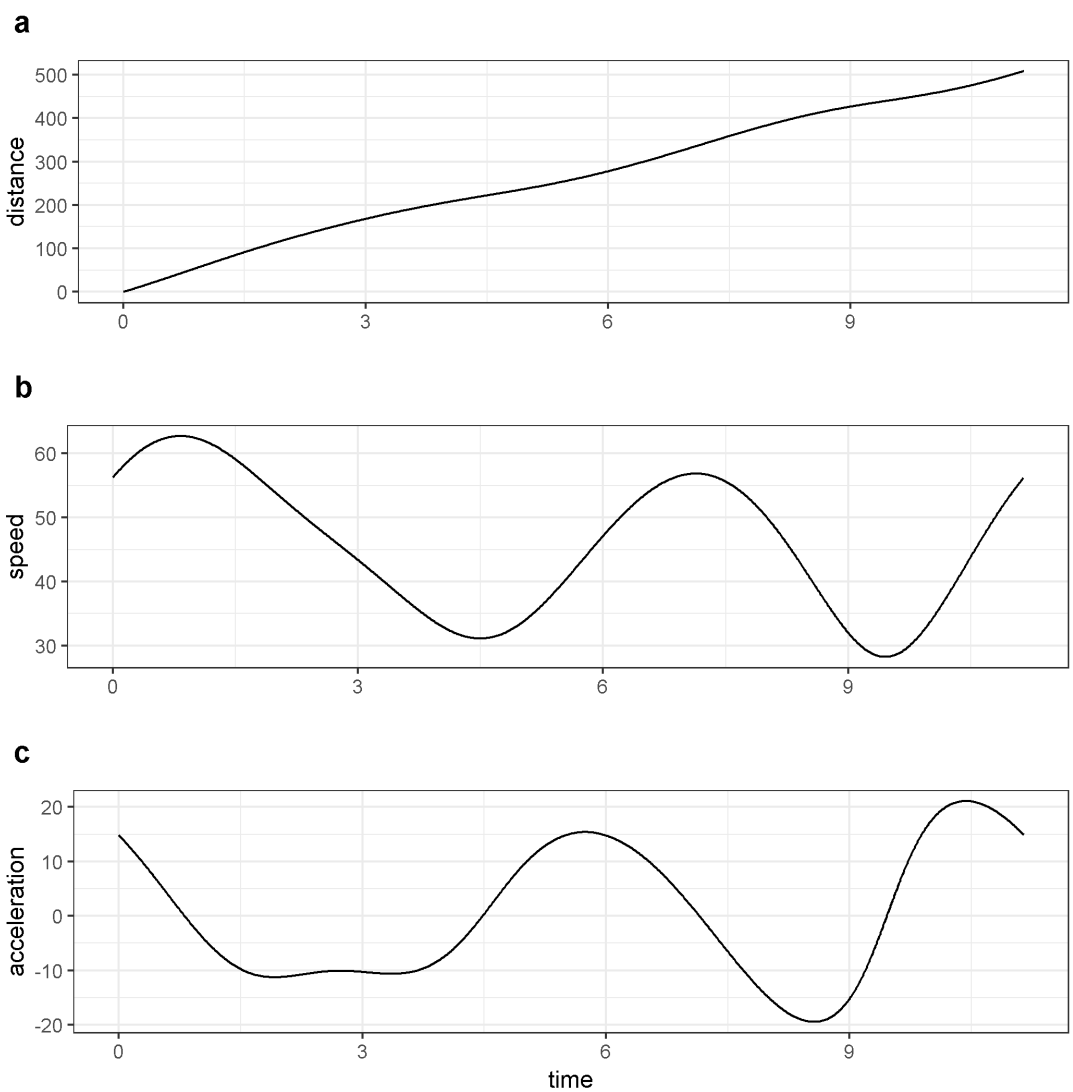


Figure 3.1: From top to bottom, distance traveled in phase space, speed tangential to system trajectory, acceleration tangential to system trajectory.

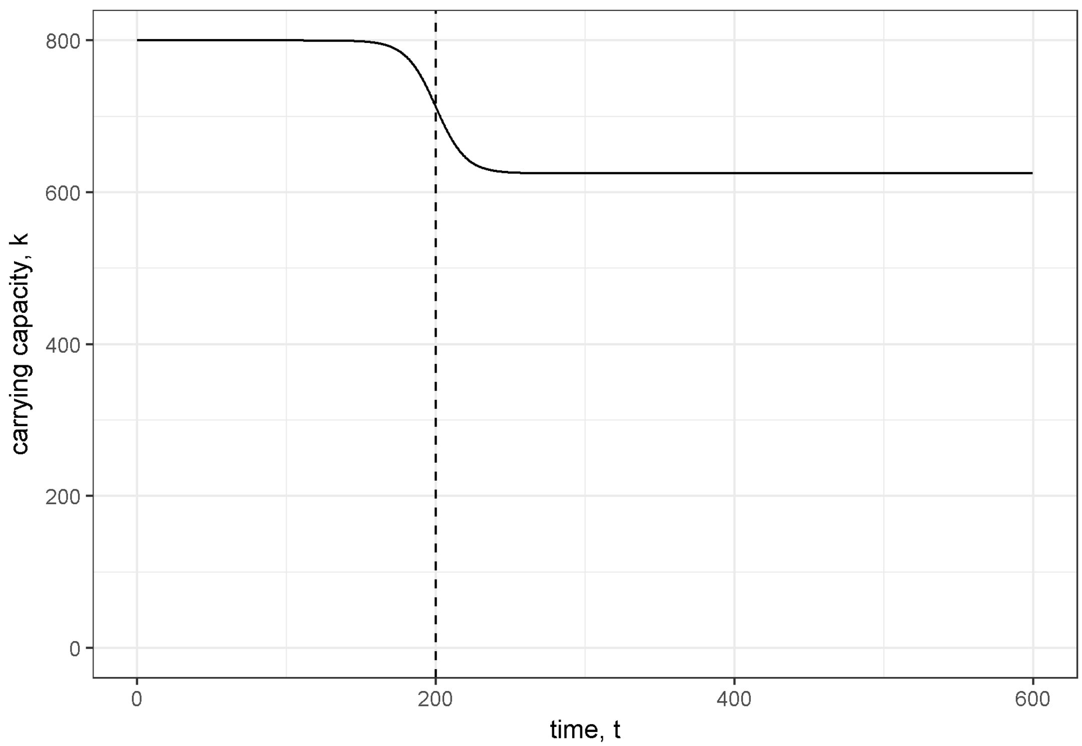


Figure 3.2: Carrying capacity over time with a regime shift occuring around time 200.

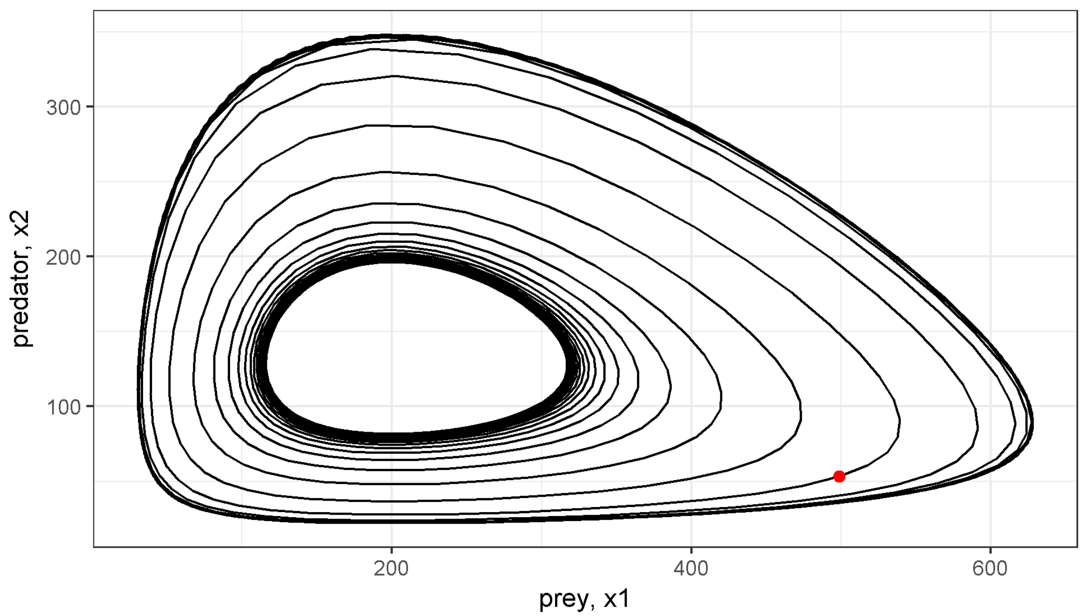


Figure 3.3: Phase space plot of system trajectories for different values of k

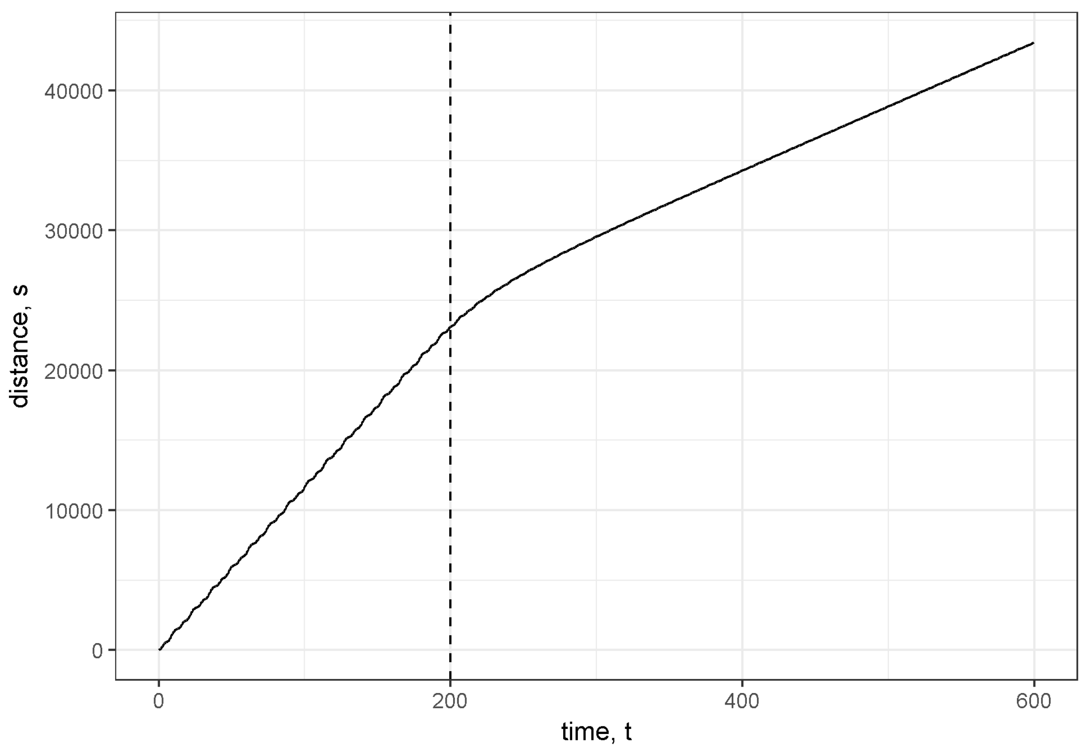


Figure 3.4: Distance travelled in phase space over time. Dashed vertical line at time 200 indicates location of regime shift.

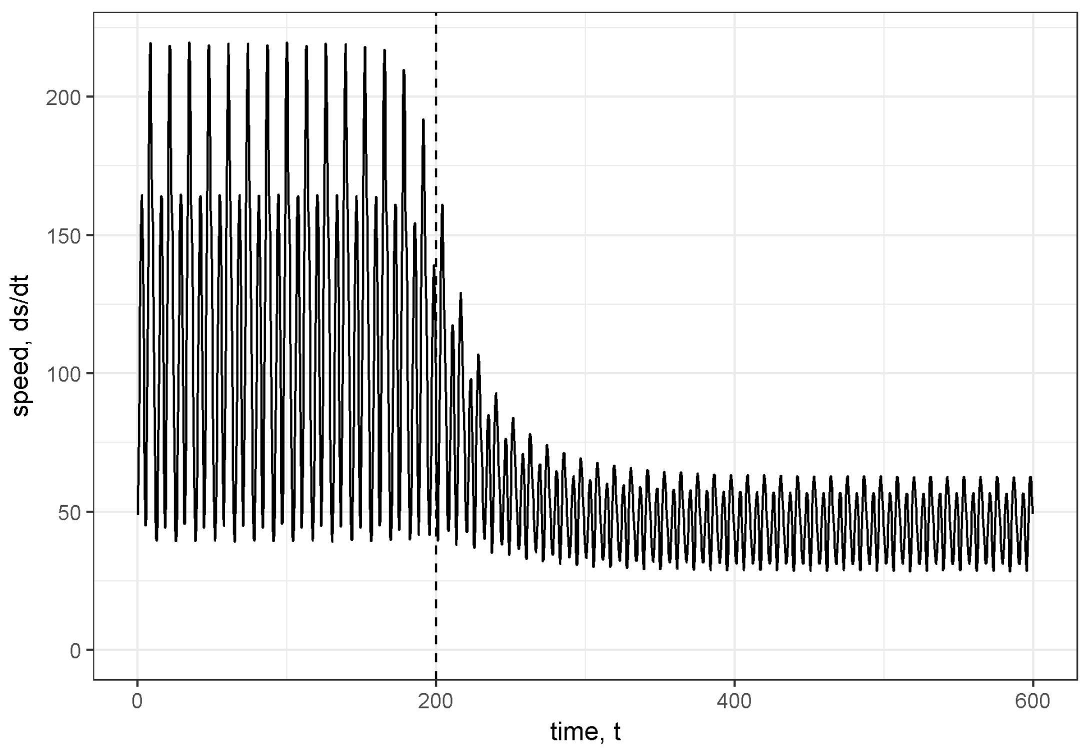


Figure 3.5: Speed of the system (rate of change, velocity) in phase space. Dashed vertical line at time 200 indicates location of regime shift.

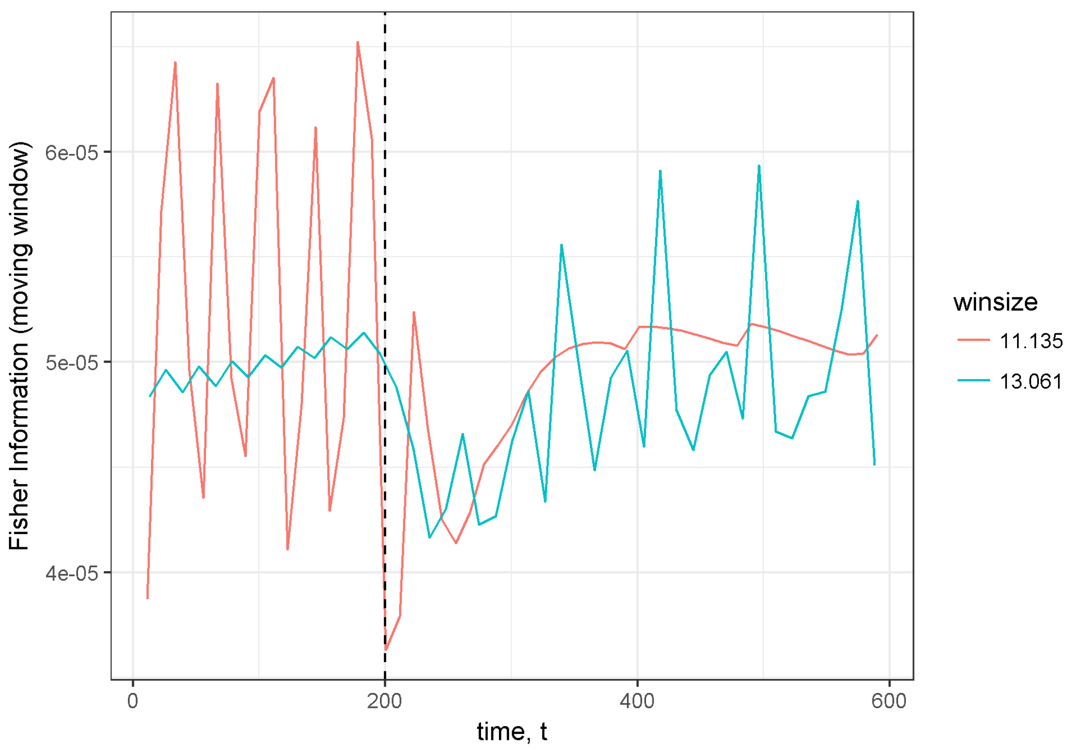


Figure 3.6: Fisher Information calculated for non-overlapping time windows. Two different window sizes were used as indicated by color. Dashed vertical line at time 200 indicates approximate location of regime shift.

Chapter 4

An application of Fisher Information to spatially-explicit avian community data

4.1 Introduction

Ecosystems are open, dynamical systems which arguably cannot be fully represented by deterministic models. Despite the complexity of most ecological systems, some patterns have emerged in certain statistical mechanics of ecological observations. An uptick in recent years of studies of **regime shifts** (1) in ecology has spurred an increase in the number of 'new' methods for detecting ecological regime shifts (2), some of which are proposed as indicators of 'spatial' regime shifts (Butitta, Carpenter, Loken, Pace, & Stanley, 2017, pp. @kefi2014early, @sundstrom2017detecting, @guttal2009spatial, @brock_variance_2006).

As defined in 1, a regime shift is largely considered an abrupt and persistent change in a system's structure or functioning. Following this definition and without any associated **pressures** 1, it is not yet clear whether identifying a 'spatial regime' using a snapshot of a system (a single or short period of time relative to the time scale of the pressure) is pragmatic. One spatial regime detection measure (hereafter, SRDM) is variance (Brock & Carpenter, 2006), despite its controversial applicability to temporal data (Bestelmeyer et al., 2011; Burthe et al., 2016; Dutta et al., 2018, p. @perretti2012regime; Sommer et al., 2017). By assuming that variance increases across space prior to a 'regime' shift, one can calculate the variability across a landscape.

Defining the spatial regime shift is important since observations of non-random spatial processes (e.g., land cover), could manifest as either rapid shift (e.g. an ecotone) or a gradual change (slow mixing along a gradient). Consequently, and because most RDMs signal abrupt change, only the former may be identified as "regime shifts" using SRDMs. For the concept of spatial regimes to be ecologically useful, potential pressures must be associated with system structure over space *and* time. Additionally and perhaps more importantly, the processes driving the observed information (drivers, pressures) should be such that a statistically identified regime shift will roughly correspond with the time scale on which the pressure(s) operate.

Although it is suggested that statistical and pragmatic models and methods are advanced more rapidly by bottom-up approaches, i.e. case studies (see DeAngelis & Yurek, 2017), to my knowledge no studies have yet to test the rigor of SRDMs using spatially-explicit empirical data. The objective of this chapter is to determine the utility of Fisher Information [Eq. (4.4)] as a spatial regime detection measure. This chapter is also supported by original software developed for implementation in Program R, which is publicly available [see Appendix 8.5].

4.2 Data and methods

4.2.1 Data: North American breeding bird communities

I use community abundance data from long-term monitoring programs to identify spatial and temporal regimes using the Fisher Information (FI) derivatives method (see Eq. (4.4)). The NABBS trains citizen scientist volunteers to annually collect data using a standardized roadside, single observer point count protocol and has been collecting data regularly across North America (4.1) since 1966. The roadside surveys consist of 50 point counts (by sight and sound) along an approximately 24.5 mile stretch of road. Due to strict reliane on volunteers, some routes are not covered every year. Additionally, some routes are moved or discontinued, and some routes are not sampled in a given year. Route-year combinations which are missing years but are not discontinued are treated as missing data. Although NABBS volunteers identify all species as possible, persistent biases exist in this protocol. To reduce the influence of potential sampling bias, I removed waterfowl, waders, and shore species (AOU species codes 0000 through 2880).

4.2.2 Study area

Although the NABBS conducts surveys throughout much of North America, I limited analyses to the continental United States and parts of southern Canada. NABBS coverage of the boreal forests of Canada are sparse in space, and many routes in Mexico have fewer than 25 years of observations.

Focal military base

The Mission of the US Department of Defense is to provide military forces to deter war and protect the security of the country, and a primary objective of individual military bases is to maintain military readiness. To maintain readiness, military bases strictly monitor and manage their natural resources. Military bases vary in size and nature, and are heterogeneously distributed across the continental United States (See Fig. 4.2). The spread of these bases (Fig. 4.3), coupled with the top-down management of base-level natural resources presumably influences the inherent difficulties associated with collaborative management within and across military bases and other natural resource management groups (e.g., state management agencies, non-profit environmental groups.

Much like other actively managed landscapes, miltiary bases are typically surrounded by non- or improperly-managed lands. Natural resource managers of military bases face environmental pressures within and surrounding their properties, yet their primary objectives are very different. Natural resource managers of military bases, whose primary objective is to maintain military readiness, are especially concerned with if and how broad-scale external forcings might influence their lands. Prominent concerns include invasive species, wildlife disease, and federally protected species (personal communication with Department of Defense natural resource managers at Eglin Air Force and Fort Riley military bases). For these reasons, natural resource managers attempt to create buffers along their perimeters (e.g., live fire/ammunitions suppression, wide fire breaks). Identifying the proximity of military bases to historic and modern ecological shifts may provide insight into the effectiveness of their natural resource management efforts. The NABBS routes chosen for analyses in this Chapter lie within or near Fort Riley military base (located at approximately 39.110474°, −96.809677°; Kansas, USA). Fort Riley (Fig. 4.4) is a useful reference site for this study. Woody encroachment of the Central Great Plains over the last century has triggered shifts in dominant vegetative cover and diversity (Ratajczak et al. 2012) in the area surrounding Fort Riley military base (e.g., Van Auken 2009). This phenomena should present itself as a regime boundary should Fisher Information be a robust regime shift detection method.

Spatial sampling grid

To my knowledge, Sundstrom et al. (2017) is the only study to use the Fisher Information on spatially-referenced data. The authors of this study hand-picked NABBS routes to be included in their samples such that their metrics should detect 'regime changes' when adjacent sampling points represented different ecoregions (broad-scale vegetation classification system). The authors also suggest each ecoregion is similarly represented, having a similar number of NABBS routes within each ecoregion in the analysis. However, this method of handpicking routes resulted in a transect which was neither North-South nor East-West running (see Sundstrom et al. (2017)), but rather zigzagged across a midwestern region. I constructed a gridded system across the continental United States and parts of Canada. The gridded system comprises East-West running transects transects running in either North-South or East-West directions. This method ameliorates some sampling bias, as I have arbitrarily defined sampling transects, rather than hand-picking sites to include in the analysis. Additionally, this approach allows for raster stacking, or layering data layers (e.g., vegetation, LIDAR, weather) on top of the sampling grid and results, allowing one to identify potential relationships with large-scale drivers. This method also provides a simple vector for visualizing changes in the Fisher Information over space-time, using animations and still figures. For brevity, I present visual results of only three, spatially-adjacent, East-West running transects (Fig. 4.5) at multiple time periods.

4.2.3 Calculating Fisher Information (FI)

Fisher Information, $I(\theta)$, was developed in 1922 by Ronald Fisher as a measure of the amount of information that an observable variable, X, reveals about an unknown parameter, θ. Fisher Information is a measure of indeterminacy (Fisher 1922) and is defined as,

$$I(\theta) = \int \frac{dy}{p(y|\theta)} \left[\frac{dp(y|\theta)}{d\theta}\right]^2 \tag{4.1}$$

where $p(y|\theta)$ is the probability density of obtaining the data in presence of θ. The Fisher Information measure (FIM) is used to calculate the covariance matrix associated with the likelihood, $p(y|\theta)$. Fisher Information is described as Extreme Physical Information (EPI; Frieden and Soffer 1995, Kibble 1999, Frieden et al. 2002), a measure that has been used to track the complexity of systems in many scientific disciplines including, physics, cancer research, electrical engineering, and, recently, complex systems theory and ecology

Fisher Information as gathered from observational data provides insight as to the dynamic order of a system, where an orderly system is one with constant (i.e., unchanging) observation points, and one whose nature is highly predictable. A disorderly system is just the opposite, where each next data point is statistically unpredictable. In ecological systems, patterns are assumed to be a realization of ecosystem order; therefore, oneshould expect orderliness in a system with relatively stable processes and feedbacks. Orderliness, however, does not necessarily infer long-term predictability. Equation (4.1) is next adapted to estimate the dynamic order of an entire system, s, as

$$I = \int \frac{ds}{p(s)} \left[\frac{dp(s)}{ds}\right]^2 \tag{4.2}$$

where $p(s)$ is the probability density for s. Here, a relatively high Fisher Information value (I) infers higher dynamic order, whereas a lower value (approaching zero) infers less orderliness. To limit the potential values of I in real data, we can calculate the amount of Fisher Information by re-expressing it in terms of a probability amplitude function $q(s)$ (Fath et al. 2003, Mayer et al. 2007, eq. 7.3):

$$I = 4 \int ds \left[\frac{dq(s)}{ds} \right]^2 \tag{4.3}$$

A form specific to the pdf of distance travelled by the entire system, which I call the 'derivatives' method, is defined as (Mayer et al., 2007, eq. 7.12):

$$I = \frac{1}{T} \int_0^T dt \left[\frac{s''^2}{s'^4} \right]^2 \tag{4.4}$$

where T is the number of equally spaced time points over which the data are integrated. Numerical calculation of I using the binning method (Eq. (4.3) and (4.4)) each incorporate a moving-window procedure for calculating the probability of the system, $p(s)$, as being in one of an unidentified number of states (s). Although previously applied to spatially-explicit terrestrial community data,the binning method requires multiple parameters to be defined *a priori*, which have been shown to influence inference based on the metric. I therefore calculated FI using the derivatives equation [see Chapter 3].

The binning procedure allows for a single point in time or space to be categorized into more than one state, which violating the properties of alternative stable states theory. The size of states (see Eason and Cabezas 2012) measure is required to construct $p(s)$. In the case of high dimensional data, a univariate binning procedure of $p(s)$ is not intuitive (i.e., reducing a multivariable system to a single probability distribution rather than constructing a multivariate probability distribution). Importantly, when using community or abundance data, rare or highly abundant species can influence the size of states criterion, thus influencing the assignment of each point into states.

Finally, Eq. (4.3) assumes equal spacing (in space or time) between sampling points. Each of these violations can be avoided by using Eq. (4.4); Cabezas and Fath 2002, Fath et al. 2003) to calculate the Fisher Information measure. The derivatives method (Eq. (4.4)) estimates the trajectory of the system's state by calculating the integral of the ratio of the system's acceleration and speed in state space (Fath et al., 2003). I calculated Fisher Information using Equation (4.4) for all East-West transects (see Fig. 4.5) for years 1980, 1990, 2000, and 2010.

4.2.4 Interpreting and comparing Fisher Information across spatial transects

Interpreting Fisher Information values

Here I define a potential regime change as a point(s) having a non-zero derivative, and at which relatively large changes (sharp increase or decrease) in the Fisher Information measure occur. Regime shifts are identified as data changing from one state to another, thus, rapid shifts in the value of FI should indicate the points, in time or space, at which the system undergoes reorganization. Spatial and temporal Fisher Information calculation does not vary, but interpretation of either differ in that a spatial analysis will identify a spatial regime boundary (Sundstrom et al., 2017) in space within a single time period, whereas analysis of temporal data will identify a point(s) in time at which a system in a specific location undergoes a regime shift. I follow the methods outlined in the relevant literature for interpreting the Fisher Information (e.g., Karunanithi et al., 2008; Eason & Cabezas, 2012).

Increases in FI is proposed as an indicator of system orderliness, where periods of relatively high values of FI indicate the system is in an "orderly" state, or is fluctuating around a single attractor. A rapid change in FI is supposed to indicated the system is no longer orderly and may be undergoing a reorganization phase. Whether Fisher Information can identify a switch among basins of attraction within a single, stable state (or around a single attractor) remains unknown, as does the number of states which a system can occupy. When a system occurs within any number of states equally, i.e., $p(s)$ is equal for each state, both the derivative, $(\frac{dq(s)}{ds}$, and I are zero. As $(\frac{dq(s)}{ds} \to \infty)$, we infer the system is approaching a stable state, and as $\frac{dq(s)}{ds} \to 0$ the system is showing no preference for a single stable state and is on an unpredictable trajectory. Eq. (4.3) bounds the potential values of Fisher Information at $[0, 8]$, whereas Eq. (4.1), Eq. (4.2), and Eq. (4.4) are positively unbounded $[0, \infty)$. If the Fisher Information is assumed to represent the probability of the system being observed in some state, s, then the absolute value of the Fisher Information index is relative within a single datum (here a single datum is a spatial transect). It follows that Fisher Information should be interpreted relatively, but not absolutely.

Interpolating results across spatial transects

Because the BBS routes are not regularly spaced, pairwise correlations of adjacent transects are not possible without either binning the Fisher Information calculations using a moving-window analysis, or interpolating the results to regularly-spaced positions in space. To avoid potential biases associated with the former option, I linearally interpolated Fisher Information within each spatial transect (Fig. 4.5) at 50 points along the longitudinal axis. The 50 longitudinal points at which I interpolated were the same across each spatial transect. I used the function *stats::approx()* to linearly approximate the Fisher Information. I did not interpolate values beyond the longitudinal range of the original data (using argument *rule=1* in package *approx*).

Spatial correlation of Fisher Information

If Fisher Information captures and reduces information regarding abrupt changes in community structure across the landscape, then the values of FI should be spatially autocorrelated. That is, the correlation of FI values should increase as the distance between points decreases. Fisher Information values calculated using Eq. (4.4) are **not** relatively comparable outside of our spatial transects, because the possible values are unbounded (can take on any value between $-\infty$ and ∞. However, because FI is directly comparabe **within** each spatial transect (e.g., 4.6), we can use use pairwise correlations among two transects (e.g., 4.6) to determine whether values of FI are consistent across space. I calculate the pairwise correlation (Pearson's) among each pair of adjacent spatial transects (e.g., Fig. 4.7). I removed a pair of points if at least one point was missing an estimate for Fisher Information. This occurred when the original longitudinal range of one transect exceeded its pair's range, since I did not interpolate beyond the original longitudinal range.

4.3 Results

4.3.1 Fisher Information across spatial transects

Interpreting the Fisher Information is currently a qualitative effort. As suggested earlier, rapid increases or decreases in FI are posited indicate a change in system orderliness, potentially suggesting the location of a regime shift. Using this method yields inconclusive results regarding the location of 'spatial regimes' (Fig. 4.8). Of the three spatial transects analyzed in this chapter (see Fig. 4.5), Figure 4.8 is representative of the lack of pattern observed in the Fisher Information values aross transects. I identified no clear pattern within or among spatial transects. Logtransforming the Fisher Information metric suppresses some of the extreme values, but still does not clearly identify sharp changes in the Fisher Information values.

4.3.2 Spatial correlation of Fisher Information

In addition to failing to identfiy clear geological boundaries across large swaths of our study area, (Fig 4.10) I also did not identify spatial correlation of Fisher Information among adjacent spatial transects (Fig. 4.11)[1]. For spatially-adjacent transects (e.g, transects 11 and 12, or 12 and 13 in Fig. 4.11), we should expect high and positive correlation values, and these values shoudl stay consistent across time *unless* the spatial transects were separated by an East-West running physical or functional boundary. This is not, however, what I expect in our East-West running transects (Fig. **??**), as the spatial soft-boundaries limiting the distribution and functional potential of avian communities are largely North-South (Fig. @ref(ewRoutes_ecoRegions)). Note spatial transects in Fig. @ref(fig:ewRoutes_ecoRegions) overlap multiple, large spatial ecoregion boundaries, such that we should expect our data to identify these points (boundaries). Upon initial investigation, there are no obvious signs of broad-scale patterns in FI across space (Fig. 4.13)[2]. If Fisher Information is an indicator of spatial regime boundaries, we should expect to see large changes in its value (in either direction) near the edges of functional spatial boundaries (e.g., at the boundaries of ecoregions). No clear regime changes appeared in areas where we might expect rapid changes (e.g., along the 105th meridian West, where a sharp change in altitude occurs).

[1] Pairs were compared (column) at select sampling years (rows), and pair-wise correlations among paired transects are presented. Large, positive correlations indicate Fisher Information signals similarly at adjacent spatial transects.

[2] Size indicates value of Fisher Information (values are scaled and centered within transects). Red box (in top panel) indicates extent of bottom panel.

Numerical investigation of the spatial correlation among adjacent transects also yielded no clear patterns. I did not identify any obvious correlation with changes in FI values and functional potential (using Omernick Ecoregion Level 2; see Fig. 4.13). Rather than abrupt changes in Fisher Information I found graduaal changes (e.g., see results for years 2000 and 2010 in Figs. 4.14,4.13).

4.4 Discussion

The Fisher Information measure was introduced as a method to avoid some analytical issues related to complex and noisy ecological data (Karunanithi et al., 2008), and has also been suggested as an indicator of *spatial* regimes (Sundstrom et al., 2017). I found no evidence suggesting Fisher Information [Eq. (4.4)] can identify 'spatial regimes'. Further, the absence of autocorrelation among spatially adjacent transects suggests Fisher Information may not be a realiable indicator of changes in bird community structure.

Although the Fisher Information equation [Eq. (4.4)] used in this study is a relatively straightforward and fairly inexpensive computational calculation, extreme care should be taken when applying this index to ecological data. Fisher Information is capable of handling an infinite number of inputs (variables), and given sufficiently low window size paramters, can technically calculate an index value for only two observations. It is important that the user understands the assumptions of identifying 'regime shifts; using Fisher Information, since the efficacy of this method has not been yet subjected to rigorous tests (but see 6). There are three primary assumptions required when using Fisher Information to estimate relative orderliness within ecological data (Mayer et al., 2007):

1. the order or state(s) (s) of the system is observable, 1. any observable change in

the information observed in the data represents reality and the variables used in the analyses will not produce false negatives, and 1. changes in I presumed to be regime shifts do not represent the peaks of cyclic (periodic) patterns.

The first assumption is one of philosophical debate and is thus not controllable. To attempt to control for false negatives, the user should take caution in her choice of input variables. In the the case of a high dimensional data, relativization and/or variable reduction measures may be useful (Rodionov 2005). However, Fisher Information does not convey information on how specific variables relate to the calculated index. Finally, we can take measures to account for cyclic behavior in the data by ensuring integration periods capture at one full cycle of the system and, given sufficently high number of observations, increasing the integration period may also alleviate some issues related to irreducible error (white noise).

The lack of patterns identified using Fisher Information may be influenced by one or more of the following: (1) the Breeding Bird Survey data collection scheme was designed to estimate and track **species** trends and not changes in entire communities; (2) these data consist of < 50 time points, and for some BBS routes much fewer. Ecological processes affecting large regions in this study area (e.g., the Central Great Plains) operate on larger time scales (i.e., » 50 points). A mismatch among the ecologically relevant scales and the temporal resolution and extent of our data may influence the ability of this index to capture large-scale changes in whole bird communities.

Aside from the typical biases associated with the BBS data (e.g., species detection probability, observer bias), there are additional considerations to be made when using these data to identify ‘spatial regimes’. Breeding Bird Survey routes are spaced apart so as to reduce the probability of observing the same individuals, but birds which fly (especially in large flocks) overhead to foraging or roosting sites have a higher probability of being detected on multiple routes. We have, however, removed these species (waders, shorebirds, waterfowl, herons) from analysis. Regardless, this study assumes there is potential for each unique BBS route to represent its own state. If routes were closer together, it is more probable that the same type adn number of species would be identified on adjacent routes. Therefore, if this method does not detect slight changes in nearby routes which occupy the same 'regime', then it follows that the method is sensitive to loss or inclusion of new species, which are spatially bounded by geological and vegetative characteristics. What new information does this give us about the system? Fisher Information reduces and removes the dimensionality of these middle-numbered systems, which omits critical information.

Effective regime detection measures should provide sufficient evidence of the drivers and/or pressures associated with the identified regime shifts (Mac Nally et al., 2014). The Fisher Information index collapses a wealth of data into a single metric, thereby foregoing the ability to relate state variables to the observed changes in Fisher Information, unlike other dimension reduction techniques. For example, loadings, or the relative influence of variables on the ordinated axes, can be derived from a Principal Components Analysis–this cannot be achieved using Fisher Information. If Fisher Information clearly suggested a spatial regime boundary or shift, a before-and-after post-hoc analysis of the regional community dynamics might confirm the regime shift occurrence.

4.4.1 Efficacy of Fisher Information as a spatial RDM

This study found no evidence suggesting Fisher Information accurately and consistently detects spatial boundaries of avian communities. Rapid changes in either direction of Fisher Information is suggested to indicate of a regime shift (Mayer, Pawlowski, & Cabezas, 2006, p. @eason_evaluating_2012). Although this interpretation has been applied to multiple case studies of Fisher Information, there is yet a statistical indicator to objectively identify these abrupt changes. After calculating the Fisher Information for each spatial transect (Fig. 4.5) during each sampling year, I used pairwise correlation to determine whether spatial autocorrelation existed among pairs of spatial transects. If some set of points are close in space and are *not* separated by some physical or functional boundary (e.g., an ecotone, high altitude rock formations), then the Fisher Infomration calculate should exhibit a relatively high degree of spatial autocorrelation that is consistent over time. It follows that the correlation coefficient of spatially adjacent transects should be similar, diverging only as the distance beteween the transects differs and/or a functional or physical boundary separates them.

Several questions remain regarding the efficacy of Fisher Information as a regime detection measure in both spatial and temporal data. If signals of regime shifts do exist, it is clearly not possible to identify them using visual interpretation. I also did not find evidence to suggest spatial autocorrelation of the calculations. I suggest future studies of Fisher Infomration focuses on temporal, rather than spatial data. Potential areas of research and questions include:

1. Relationship of Fisher Information to likelihood ratio-based unsupervised change-point detection algorithms (e.g., ChangeFinder; Liu, Yamada, Collier, & Sugiyama, 2013).

2. Sensitivity of Fisher Information to data quality and quantity [this is explored in Chapter 6].
3. What, if any, advantages does FI have over other density estimation techniques?
4. Does FI provide signals in addition to or different than geophysical and vegetative (e.g. LIDAR) observations (data)?

Figure 4.2: A single East-West transect of Breeding Bird Survey routes used to calculate the Fisher Information.

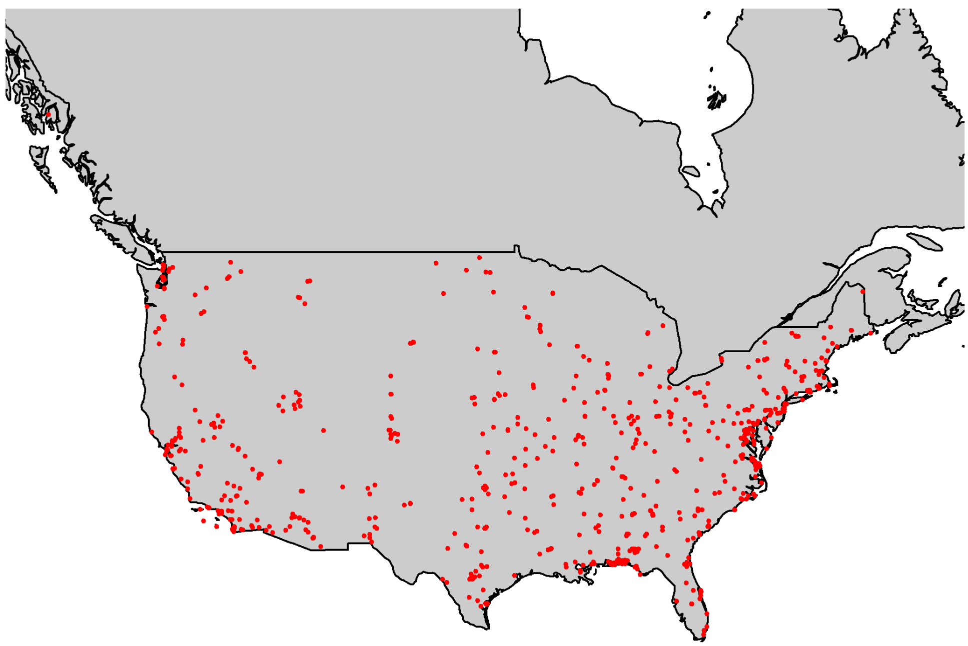


Figure 4.3: Locations of U.S. military bases in our study area.

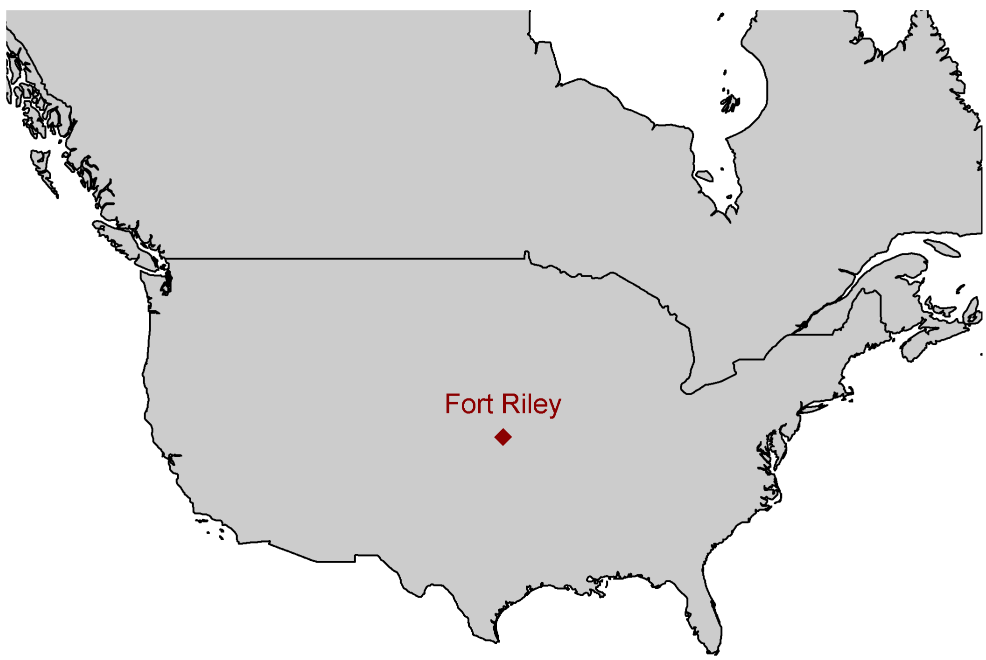


Figure 4.4: Locations of focal U.S. military bases, Eglin Air Force Base (AFB) and Fort Riley Military Base.

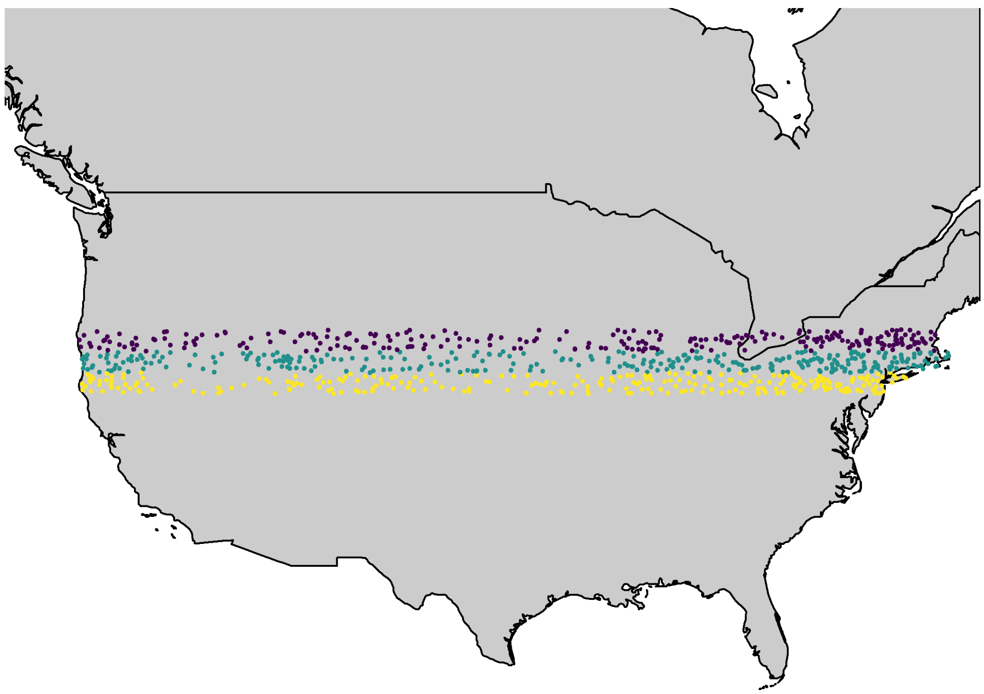


Figure 4.5: The three East-West running transects used to visualize results in this chapter.

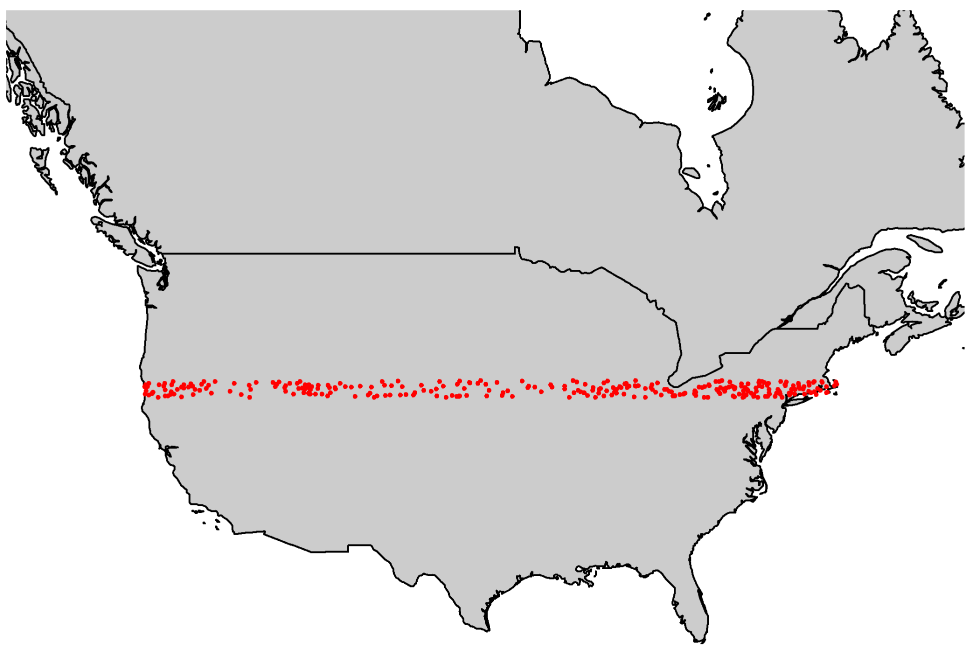


Figure 4.6: An example of two adjacent spatial transects within my sampling grid.

Figure 4.7: An example of two adjacent spatial transects (12, 13) within my sampling grid.

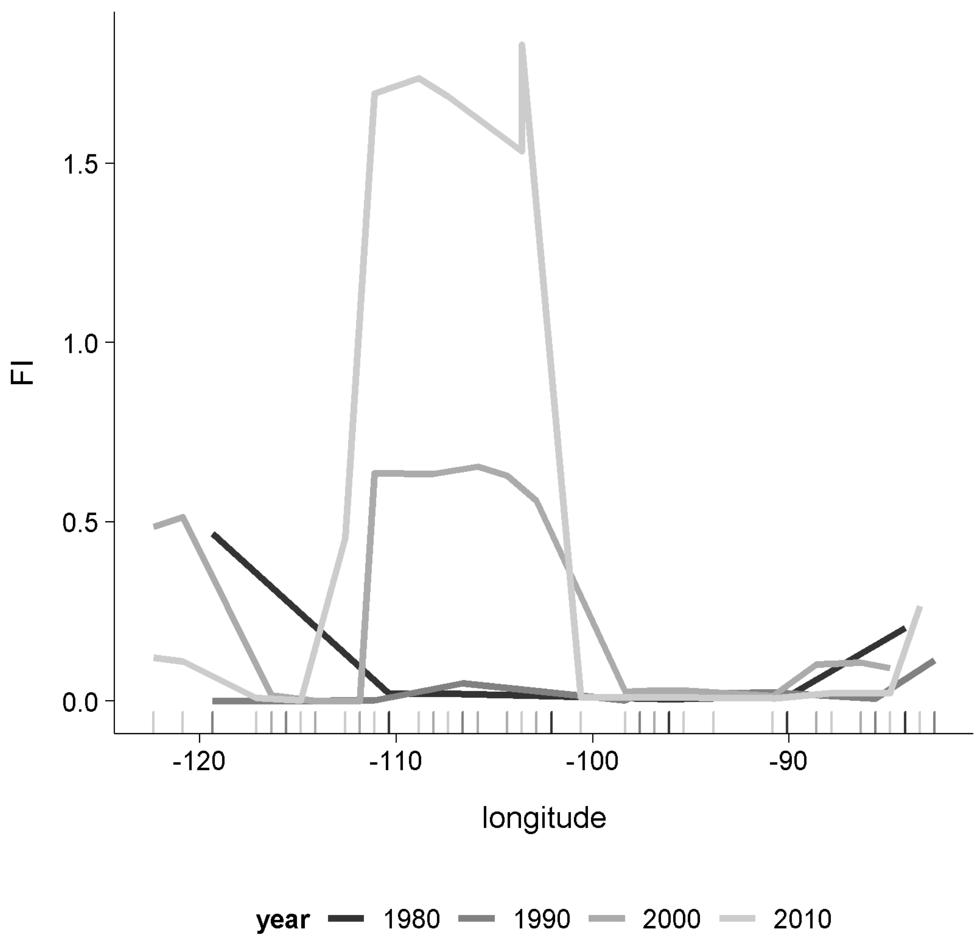


Figure 4.8: Fisher Information calculated for a single transect over time.

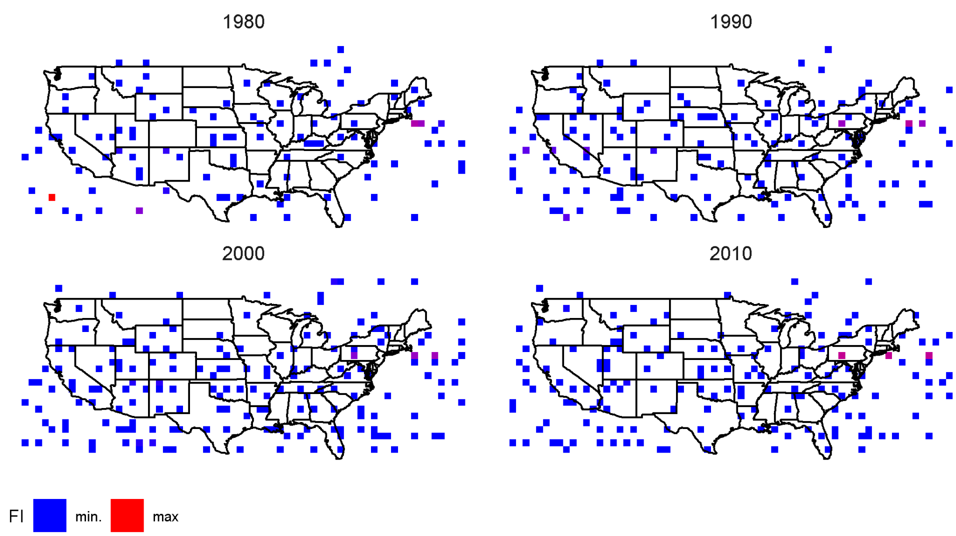


Figure 4.9: Fisher Information of 5 East-West spatial transects over time.

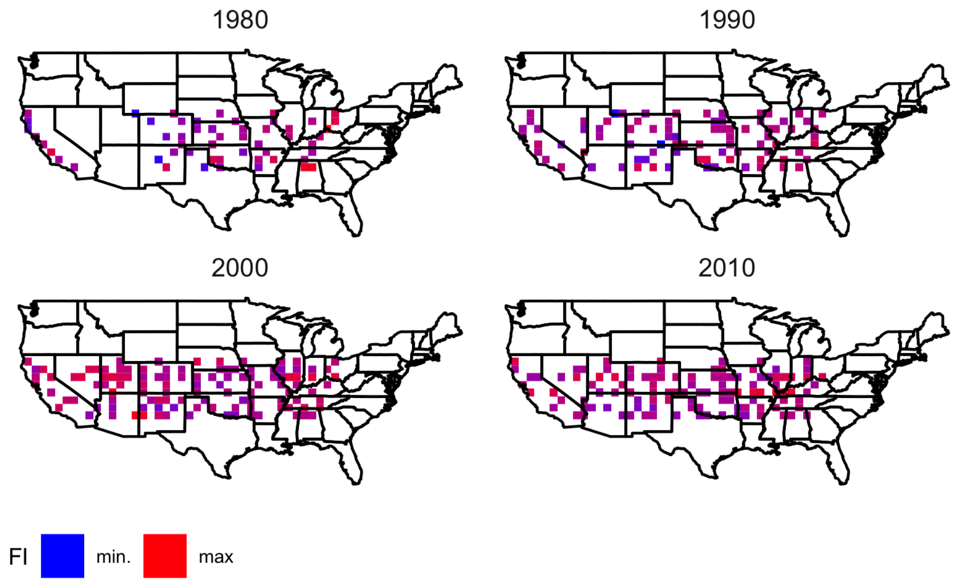
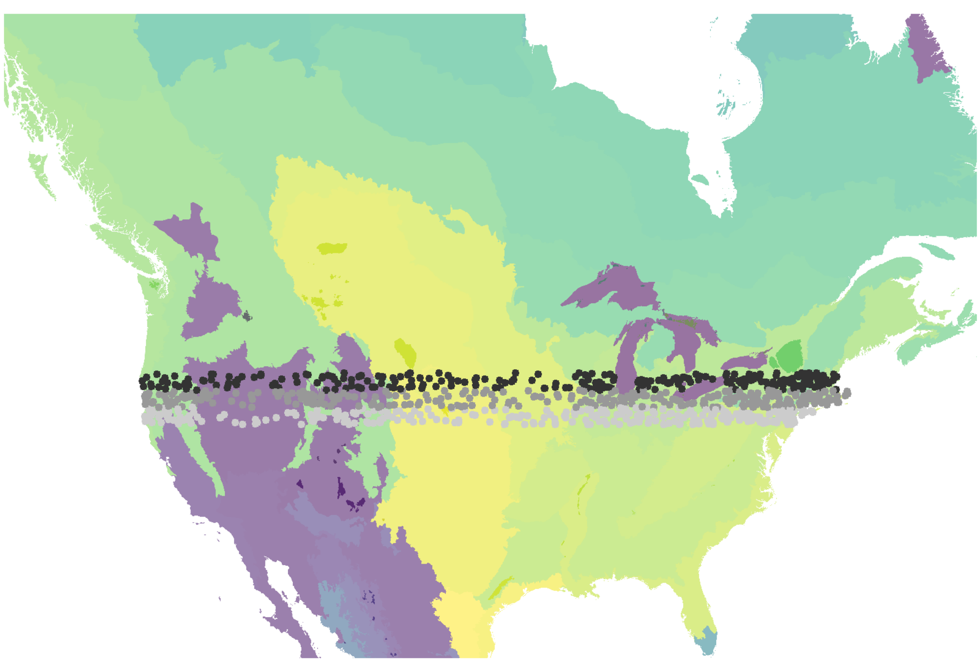


Figure 4.10: Fisher Information of 5 East-West spatial transects over time.



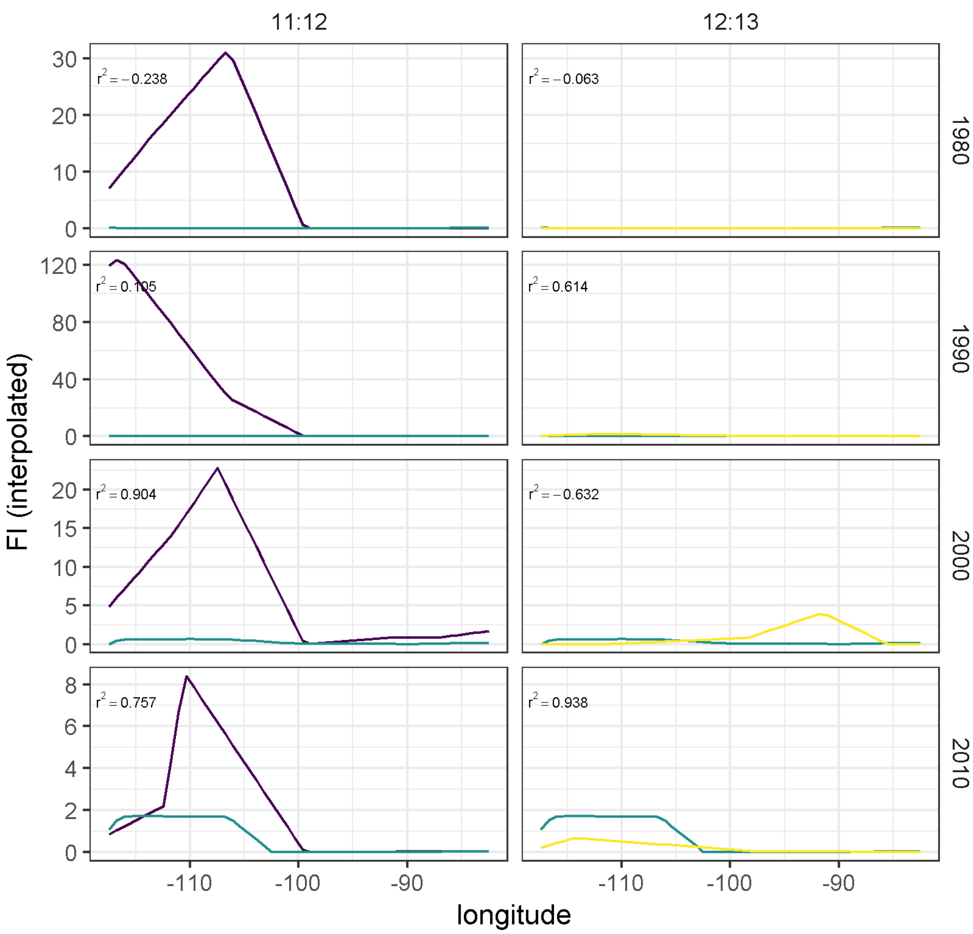


Figure 4.11: Pairwise relationships of Fisher Information (interpolated values) of spatially adjacent transects over time.

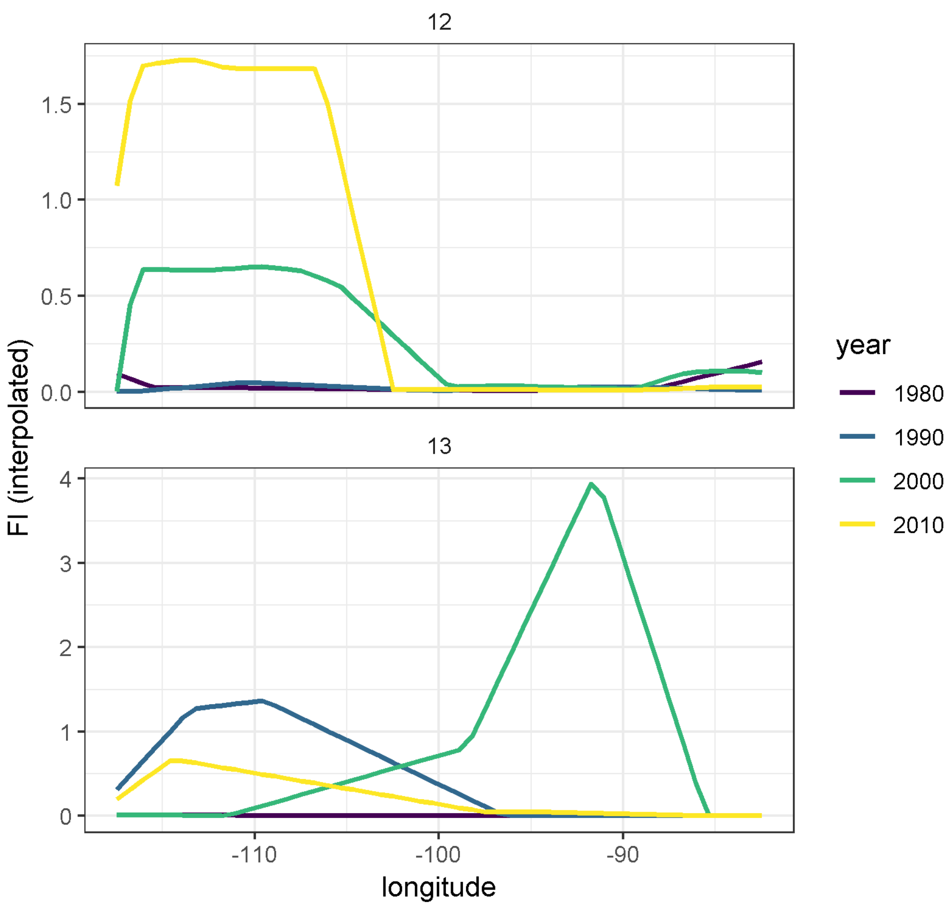


Figure 4.12: Fisher Information of two transect pairs over time.

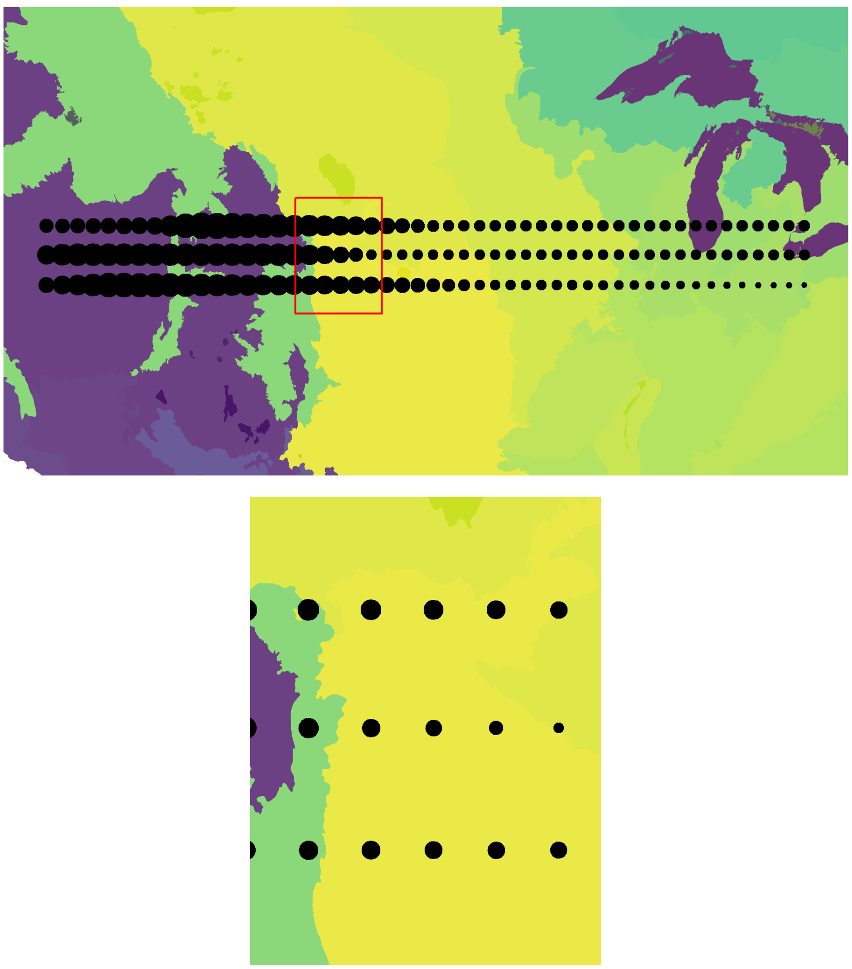


Figure 4.13: No patterns of abrupt change detected in Fisher Information along three transects in year 2010

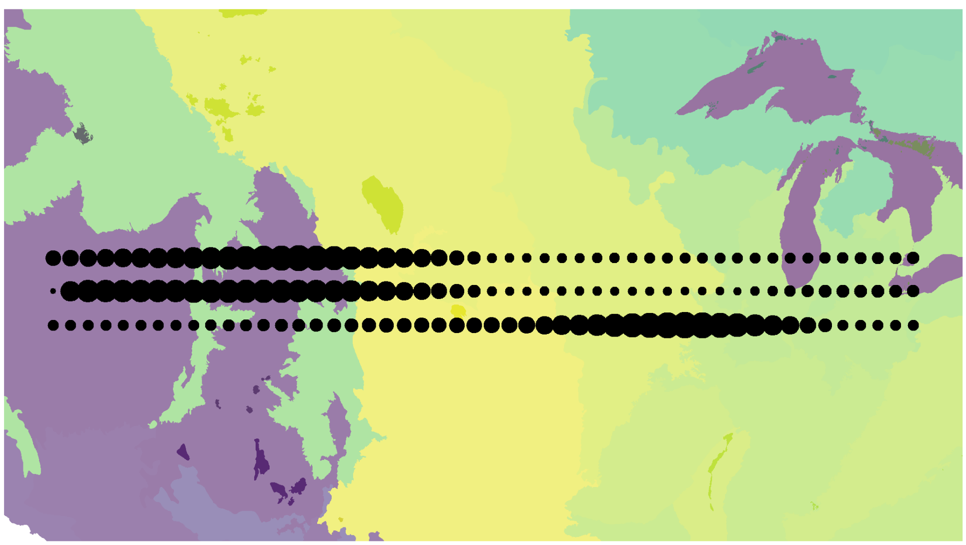


Figure 4.14: Fisher Information (scaled and centered; point size positively correlated with value) against ecoregion boundaries (EPA Level 2).

Chapter 5

Velocity (v): using rate-of-change of system trajectory to identify abrupt changes

5.1 Introduction

When, how and why ecological systems exhibit abrupt changes is a hallmark of modern ecological research, and changes which are unexpected and undesirable can have undesirable downstream consequences on, e.g., ecosytem services, biodiversity, and human well-being. Quantitatively detecting and forecasting these changes, however, has yet to be accomplished for most ecological systems (Chapter 2; Ratajczak et al., 2018). Moving from abrupt change methods requiring highly descriptive models and *a priori* assumptions of the state variable responses to drivers to methods requiring few, if any, *a priori* assumptions or knowledge is increasingly necessary for forecasting and managing complex ecosystems under an era of intensifying anthropogenic pressures.

A few broad classes of quantitative approaches exist for quantitatively identifying abrupt changes in complex ecosystems. First, one can use simple mathematical models to describe the system and statistically test for discontinuities in the observed variables (e.g., in coral reefs, Mumby, Steneck, & Hastings, 2013). Although mathematical representations are ideal, very rarely are ecological systems easily and well-described by them and often fail to meet the assumptions of the model. Second, we can track changes in the mean or variance of state variables to identify departures from the norm (e.g., early-warning indicators such as variance and variance index, Brock & Carpenter, 2006). Much like the mathematical modelling approach, these early-warning indicators have shown to be useful in some simple driver-response systems (e.g., lake eutrophication Carpenter, Brock, Cole, Kitchell, & Pace, 2008), but are unreliable in other empirical systems (e.g., Perretti & Munch, 2012; Dakos et al., 2012; Dutta et al., 2018). The last type of approach is the model-free approach [Dakos, Carpenter, et al. (2012); Ch. 2]. This group of abrupt change indicators can incorporate multiple state variables, and ideally requires no *a priori* assumptions about the expected driver-response relationships, or even about the drivers at all. It is this class of abrupt change indicators to which this chapter contributes.

5.1.1 Tracking ecosystem trajectory through time to explore system dynamics

```
knitr::include_graphics(here::here("chapterFiles/velocity/figsCalledInDiss/lorenz3
```

```
knitr::include_graphics(here::here("chapterFiles/velocity/figsCalledInDiss/lorenz3
```

A classic example of state-switching by a system is demonstrated in the Lorenz (‘butterfly’) attractor (Fig. 5.1; Takens, 1981). This phase plot (Fig. 5.1) provides an informative visual of the behavior of a chaotic system manifesting two attractors.

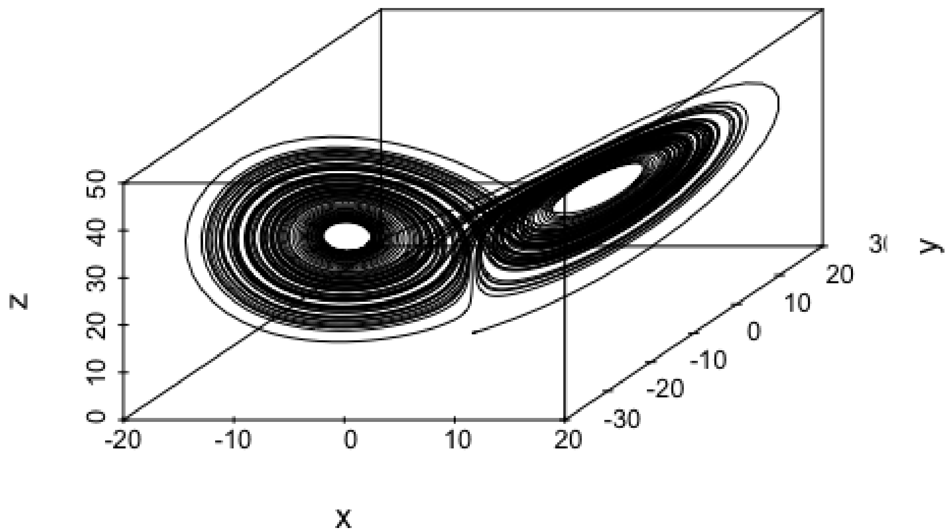


Figure 5.1: An example solution of the Lorenz ('butterfly') represented in 3-dimensional phase-space. Phase plots are typically used to visualize stable areas within a system's trajectory but reconstruction requires the difference models to be known and parameterized.

Although the periodic, attractor behaviors are made clear when examining the time series of each dimension (Fig 5.2), identifying such behaviors in additional dimensions becomes increasingly difficult.

System behavior/trajectory in phase space are used often in dynamical systems theory and systems ecology to make inference regarding system behavior and dynamics, but phase space (trajectory) dynamics are not commonly applied outside theoretical studies as a tool for ecological data analysis (c.f. Sugihara et al., 2012 for an example of phase-space reconstruction using Taken's theorem of ecological time series). Some methods of attractor reconstruction have been applied to environmental data (e.g., individual time series of fisheries stocks, climate, stock market; Sugihara et al., 2012; Ye et al., 2015), yet they **do not incorporate the dynamics of whole-systems**. Model-free methods for exploring and describing the dynamics of whole (i.e. >1 variable) ecological systems are restricted to the commonly-applied dimnesion reduction techniques and clustering algorithms (e.g., Principal Components Analysis,

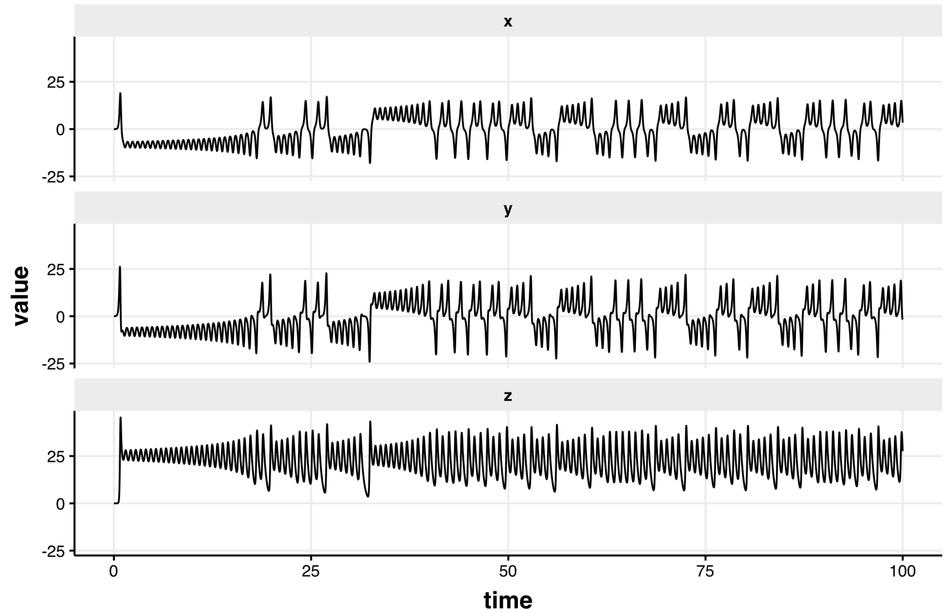


Figure 5.2: An example solution of the Lorenz ('butterfly') represented in individual system compoents.

K-means clustering). In fact, this is true of many abrupt change and regime shift indicators.

5.1.2 Rate of change as an indicator of abrupt change in the system trajectory

How quickly a system switches states [e.g., moving from attractor to another; 5.1] may yield insights into the responses of ecological systems to perturbations (e.g., anthropogenically induced pressures such as climate change, urbanization) and community shifts (e.g., species introductions or extinctions, shifts in dominance). For example, Beck et al. (2018) tracked rate of change using chord distances—a data transformation for positive values and which is suitable prior to ordination analysis—to capture abrupt changes in community composition of a temperate, paleodiatom community. Chord distance, however, is greatest when the observations among data rows (e.g., time, location) have no species in common. In other words, this measurement may be most useful in high community turnover conditions. Identifying alternative numerical methods for estimating system rates of change may be when the system does not exhibit, for example, high degrees of turnover.

Rate of change (ROC, often represented as Δ) is a term used for various measures which describe the relationship among to variables, measuring the change in one variable relative to another. As a refresher ROC is represented as **speed** (**S**) or **velocity** (**V**), where (**S**) is the adirectional magnitude (i.e. it is a scalar) of the displacement of an object over unit time and **V** describes both the direction and magnitude (i.e. it is a vector) of the object's movement in spacetime. **S** is a scalar taking values of ≥ 0 and **V** can take any value between $-\infty$ and ∞. For example, consider a car travelling at a constant speed of $50\frac{km}{h}$ around along a hilly landscape, where it is ascending and descending hills. Although **S** is constant, **V** changes in a sunusoidal fashion, where **V** is $\mathbf{V} > 0$ when ascending, $\mathbf{V} < 0$ when descending, and $\mathbf{V} \approx 0$ at in the valleys and at the peaks of the hills. Although **S** is useful when estimating other scalar quantities (e.g., $\frac{miles}{gallon}$), given a starting and/or final position in space, **S** is not informative of its the path travelled.

5.1.3 Aims

Here, I propose a method which simply describes the rate of change behavior of system dynamics in phase space: **velocity**, V. An alternative to other complicated, model-free approaches (e.g., Fisher Information; Cabezas & Fath, 2002), the velocity metric allows one to examine the behavior of an entire system along its trajectory (through space or time) without having to reconstruct the pahse space. The ability to handle noisy and high-dimensional data and the lack of subjective parameters in calculating the metric makes this method an ideal alternative to existing early warning indicators and phase-space reconstruction methods.

I first describe the steps for calculating this new metric (v), as both a dimension reduction technique and abrupt change indicator. Although this is the first instance of this calculation to, alone, be suggested as a regime detection metric, it has been used as part of a larger series of calculations of the Fisher Information metric [see Ch. 3], first introduced in Fath et al. (2003). I use this theoretical system to present baseline estimates of the expected behavior of v under various scenarios of changing mean and variability in a theoretical, discussing the contexts under which this metric may signal abrupt changes. Finally, I explore the utility of this metric in identifying known regime shifts in an empirical paleoecological time series data.

5.1.4 Analytical approach

I first describe the steps for calculating velocity by constructing a simple, two-variable system which exhibits only a rapid, discontinuous change in the means of the state variables. I next vary the mean and variance of the state variables of this system to demonstrate baseline expectations for the behavior of velocity under a simple rapid shift scenario. Next, I construct a second model system similar to the first, but one which exhibits a non-discontinuous rapid change in the state variables. The purpose of this section is three-fold. First, I demonstrate how velocity behaves when the system undergoes varying degrees of change (e.g., slow change versus nearly discontinuous, rapid). Second, I concurrently identify baseline expectations of velocity under varying conditions of mean and variability of the state varilbes before and after a shift. Third, by introducing a smoothing function to the rapid shift, we gain an understanding of how process variability (noise) impacts the shift detectability by the velocity metric. Finally, I calculate the velocity of an empirical, paleolitic freshwater diatom community time series to demonstrate the utility of the velocity metric in highly noisy, high dimensional, and irregularly-sampled data.

5.2 Steps for Calculating velocity, v

In this section, I first demonstrate the calculations of velocity using a very simple, two-variable toy system. The first system exhibits a rapid shift at a single point in time, where mean and variance are constant before and after the shift point. I demonstrate the signals achieved with and the variability within the v calculation by exploring a number of scenarios of this simple system. For the examples in this section, observations of x_i are randomly drawn from distribution $x_i \sim Normal(\mu, \sigma)$, where μ is the mean and σ is the standard deviation. Consider a system (Fig. 5.3)

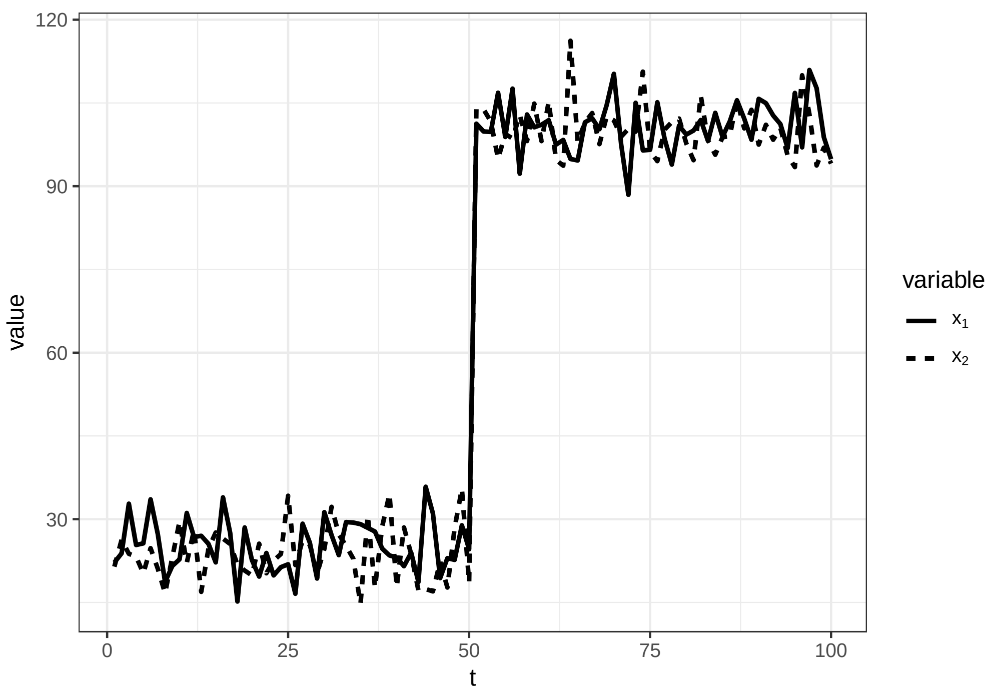


Figure 5.3: The 2-variable discrete time toy system used to demonstrate steps for calculating system velocity. Each variable, x, is drawn from a normal distribution with means that change at $t = 50$. State variables have constant standard deviation, $\sigma = 5$.

with N state variables (x_i), with observations taken at time points, t. System velocity is calculated as the cumulative sum over time period t_0 to t_j, as the total change in all state variables, $\{x_1...x_N\}$, between two adjacent time points, e.g., t_j and t_{j+1}, denoted $t_{j,j+1}$. I use this simple, two-variable system to demonstrate how *velocity* is calculated. The system comprises variables x_1 and x_2, with observations occurring at each time point $t = 1, 2, 3, ...100$. First, we calculate the change in each state variable, x_i, between two adjacent points in time, t_j and t_{j+1}, such that the difference, $x_{t_{j+1}} - x_{t_j}$ is assigned to the latter time point, t_{j+1}. For example, in our toy data, we use observations at time points $t = 1$ & $t = 2$ (Fig. 5.4). For all examples in this chapter, the state variables x_1 and x_2 were drawn from a normal distribution (using function *rnorm*), with parameters $\bar{x}_i$ (mean) and σ_i (sd) for 100 time steps, t. The regime shift in this system occurs at $t = 50$, where a shift in either or both $\bar{x}_i$ or σ_i.

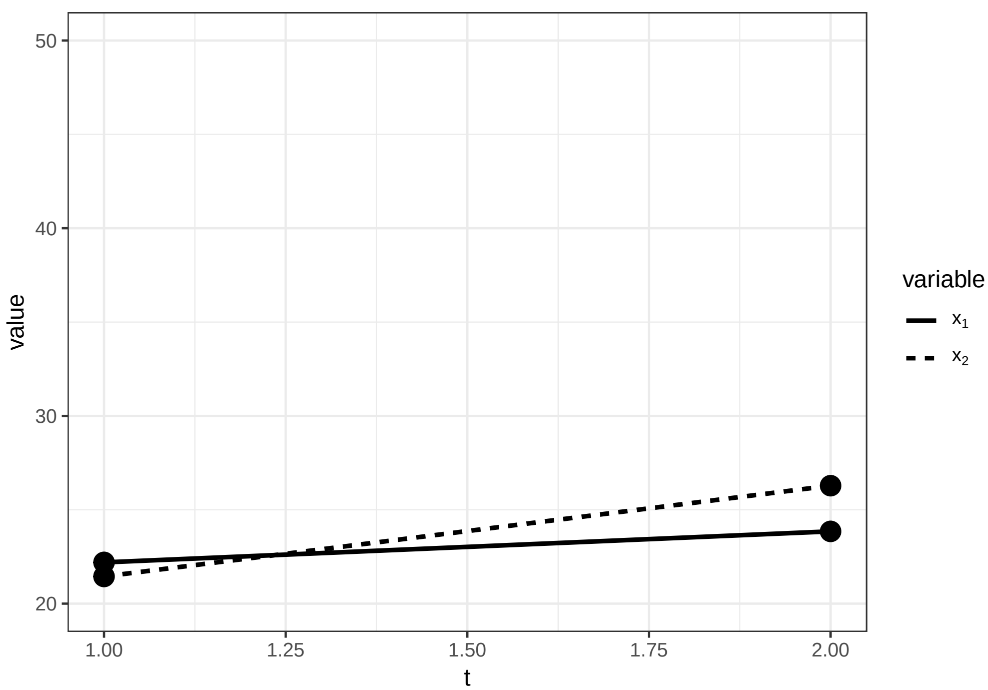


Figure 5.4: Data used to calculate velocity at the first two time points, t_1 and t_2.

5.2.1 Steps for calculating v

Step 1: Calculate Δx_i

The first step is to calculate the change in values for each state variables, x_i, between two consecutive time points [e.g., from time t to $t+1$ for the discrete-time system; Fig. 5.4; Eq. (5.1)]:

$$\Delta x_i = x_{i(t+1)} - x_{it} \tag{5.1}$$

Note that Δx_i can take any value between $-\infty$ and ∞.

Step 2: Calculate distance travelled, s

Next, we calculate the total change in the multivariable system as a function of the change in all state variables x_i. First, we calculate Δs as the square root of the sum of squares of the changes in all state variables per Pythagora's theorem (Eq. (5.2)):

$$\Delta s = \sqrt{\sum \Delta x_i^2} \tag{5.2}$$

Although Δs represents the absolute change in the system between consecutive points in time, this measure is not yet relative along the system's trajectory. To create a relative value we next calculate the total distance travelled along the system trajectory, s, as the cumulative sum of Δs (Eq. (5.2)) since the first observation, such that a cumulative sum is calculated for every t over the interval $[0, T]$ (Eq. (5.3)):

$$s_T = \Sigma_{t=0}^{T} \Delta s) \tag{5.3}$$

We now have a single measure, s_T [hereafter referred to as s; Eq. (5.3)] at each discrete point in time in our N-dimensional system (Fig. 5.5). It should be noted that s(Fig. 5.5) is monotonically increasing since the value of Δs (Eq. (5.2)) is a sum of

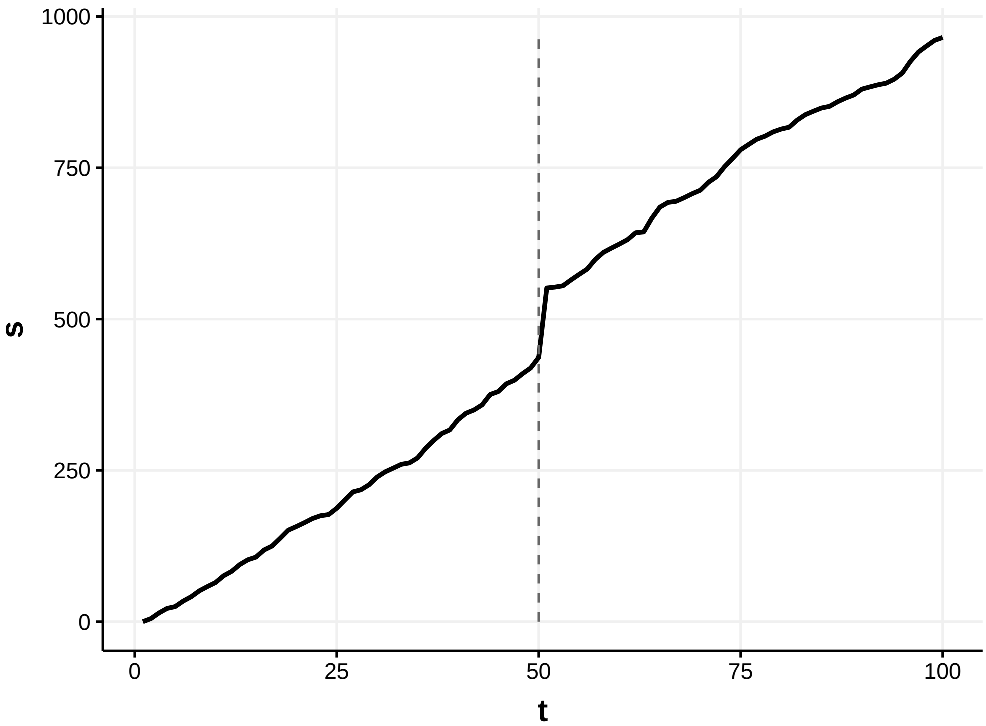


Figure 5.5: Distance travelled, s, for the 2-species toy system.

squares. Although discussed in a later section, it is important to note that s is not unitless—that is, s has units of the state variables, x_i. For example, if our 2-variable toy system represents biomass, then the units of s represents the cumulative absolute change in biomass of the entire system.

Step 3: Calculate velocity, v (or $\frac{\Delta s}{\Delta t}$)

Finally, we calculate the **system velocity**, v (or $\frac{\Delta s}{\Delta t}$), by first calculating the change in s (Eq. (5.3)), and then divide by the total time elapsed between consecutive sampling points:

$$v = \frac{s_{t+1} - s_t}{\Delta t} \tag{5.4}$$

The numerical results for each step in the calculation of velocity [Eq. (5.4)] is demonstrated using the first five time points of our toy system (Fig. 5.3) in Table 5.1.

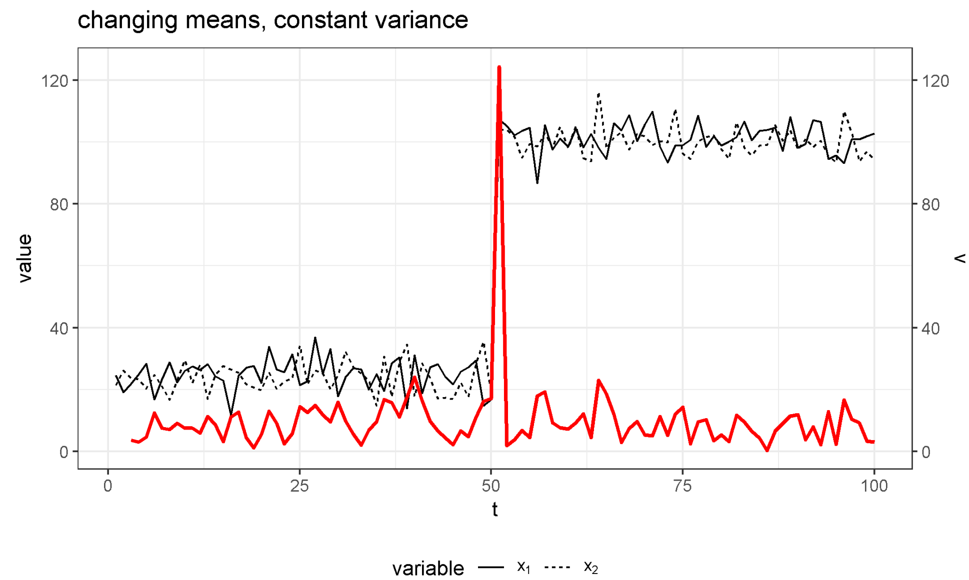


Figure 5.6: System change (s) and velocity (v) of the model system over the time period. Constant means ($\bar{x}_{pre} = 25$, $\bar{x}_{post} = 10$) and sharp change in variance for both state variables, $\sigma = 5$.

Table 5.1: Steps outlined for calculating system velocity, v, using the 2-variable toy data as an example.

t	x_1	x_2	Δx_1	Δx_2	Δt	$\sqrt{(\sum_{i=1}^{N} \Delta x_i^2)}$	s	v
1	22.198	21.448						
2	23.849	26.284	1.651	4.836	1	5.111	5.111	
3	32.794	23.767	8.944	-2.518	1	9.292	14.403	9.292
4	25.353	23.262	-7.441	-0.504	1	7.458	21.861	7.458
5	25.646	20.242	0.294	-3.020	1	3.035	24.895	3.035

5.3 Velocity v performance under a discontinuous transition

I used simulation techniques to determine the baseline expectations of the performance of velocity v under varying degrees of rapid shifts in the mean and variance of the toy system. The toy system in this section undergoes a discontinuous shift at $t = 50$ (see ??). If the system undergoes a rapid and discontinuous change in one or more state variables, the velocity, beacuse it is a rate of change, may $\rightarrow \infty$ as $\Delta t \rightarrow 0$. Therefore, it is important to understand the degree to which velocity can detect very sudden changes in mean values, despite effect sizes. Here, I varied each of the following system parameters at the regime shift location ($t = 50$): $\bar{x}_1$, increase in the mean value of x_1 and σ_1, the change in variance of x_1.

Simulations consisted of 10,000 random samples drawn from the normal distribution for each paramter, I randomly drew the toy system samples 10,000 times under increasing values of $\bar{x}_1$ and σ_1. To identify patterns in the influence of paramter values on velocity, I present the mean values of v across all simulations, with confidence intervals of ± 2 standard deviations. As mentioned above, the state variables x_1 and x_2 were drawn from a normal distribution (using function *rnorm*), with parameters $\bar{x}_i$ (mean) and σ_i (sd) for 50 time steps, t.

Velocity under varying degrees of Varying post-shift mean

I examined the influence of the magnitude of change in x_1 in the period before (pre; $t < 50$) and after (post; $t \geq 50$) by varying the mean parameter, $\bar{x}_1$ in the set $W = \{25, 30, 35, ...100\}$ (Figs. 5.7,**??**). As expected, the magnitude of v increases linearly as the total difference between $\bar{x}_{1_{pre}}$ and $\bar{x}_{1_{post}}$ increases (Fig. 5.8). This is not surprising because s increases as the total change in abundance across the entire sytem increases (Eq. (5.3)). Consequently the potential of v also increases with total state variable values (e.g. abundance, biomass). The linear relationship among v and total state variable values indicates that while v is capable of identifying large shifts in data structure, it may fail to identify subtle changes (i.e. lower effect sizes).

Varying post-shift variance

In the previous example, variance was constant before and after the abrupt shift at $t = 50$. To determine whether the signal emitted by v at the regime shift is lost or

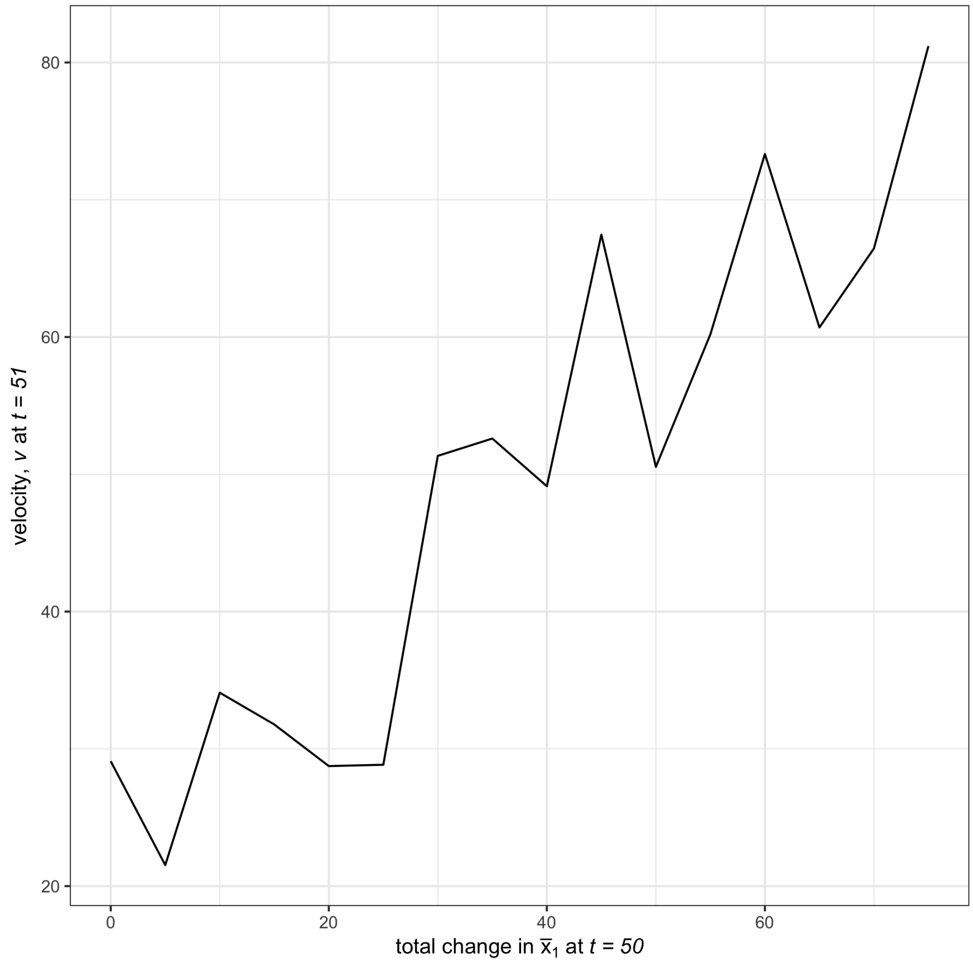


Figure 5.7: Velocity (v) generally increases as the total change in the mean value of $\bar{x}_{1_{t=50}}$ increases in a single iteration of our toy system ($N_{iter} = 1$, seed = 123). This 2-variable system exhibits a regime shift at $t = 50$, where variance is constant $\sigma = 5$, $\bar{x}_1 = 25$ when $t < 50$, $\bar{x}_2 = 50$ when $t \geq 50$, $\bar{x}_1 = 25$ when $t < 50$.

dampened when increasing variance I varied the variance parameter, σ_1 along the set $W = \{1, 2, 3, ...25\}$. The variance for both state variables (x_1, x_2) prior to the regime shift, σ_{x_1} and σ_{x_2}, was 5, with the change occurring in σ_{x1post}. Sytem velocity v appears senstive to increases in the variance at the point of the regime shift (Figs. **??**, **??**). Again, because velocity is a function of the total change in the state variables, the variability of v will increase with both mu and σ (Fig. **??**). Further, it is unsurpirsing that v exhibits sensitivity to changes in σ_{post} because, without prior smoothing of the

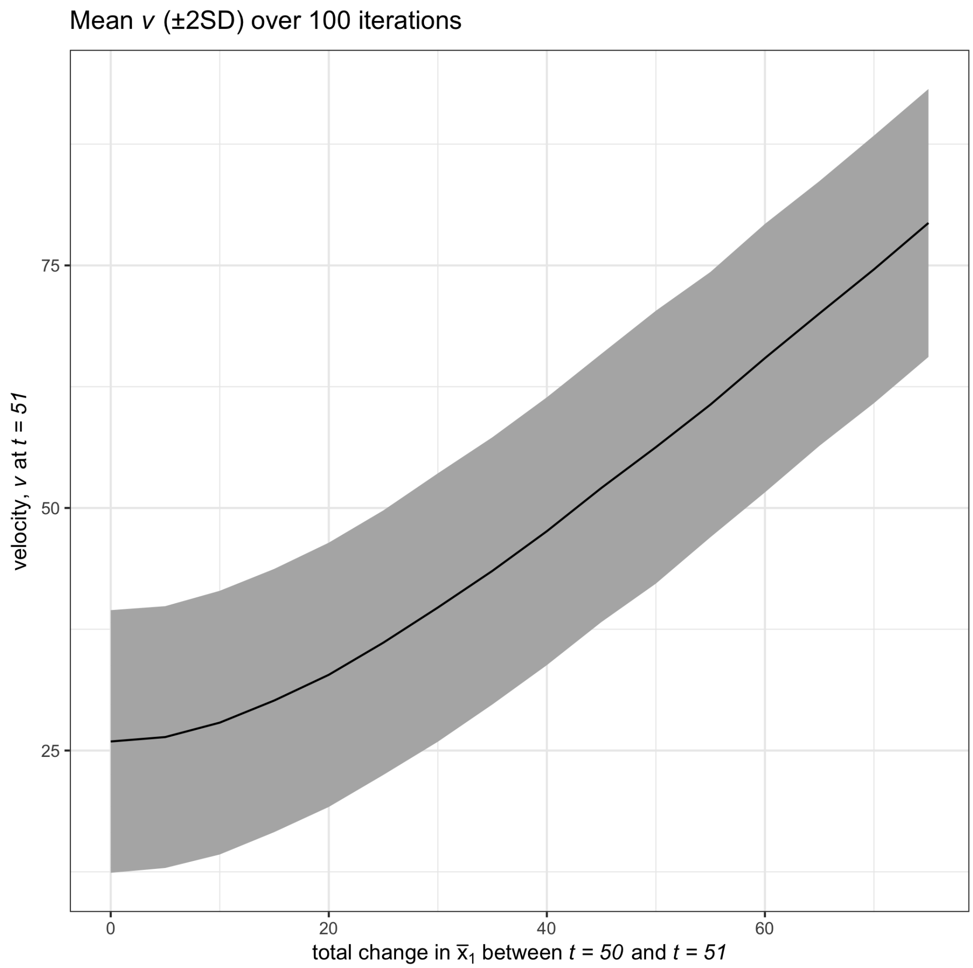


Figure 5.8: Change in velocity (v) as the total change in the mean value of $\bar{x}_{2_{t=50}}$ over 10,000 simulations. A regime shift was induced at $t = 50$ with constant varoance $\sigma = 5$, $\bar{x}_2 = 25$ when $t < 50$, and changes in variable mean values, $\bar{x}_2 = 50$ when $t \geq 50$, $\bar{x}_1 = 25$ when $t < 50$.

data, the tangential speed of a 'noisy' variable will always be noisy itself (see Figs. 5.6, 5.30, 5.31, 5.32).

Smoothing the data prior to calculating *v*

To determine whether process or observational noise influences the signal in v, I used linear approximation techniques to smooth the data prior to calculating the derivatives. I used the function `stats::approx` which linearaly interpolates the original data, x_1

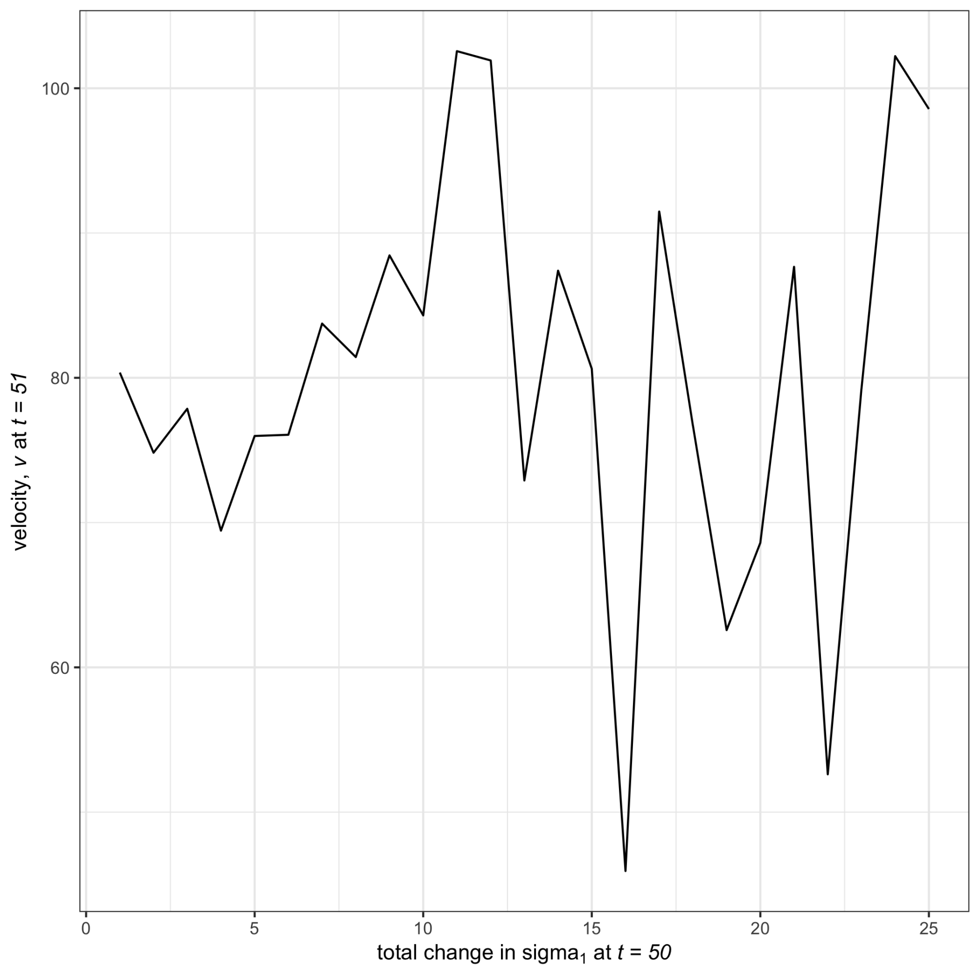


Figure 5.9: High variance of velocity (v) in a single iteration ($N_{iter} = 1$, seed $= 123$) of simulations as we increase σ_1 at $t = 50$.

and x_2, to regularly-spaced time points along the set $t = \{1 : 100\}$. I then calculated v as described in (Eqs. (5.1):(5.4)). Increasing the number of points (t) at which the original state variables were smoothed (i.e., t) did not influence the amount of noise surrounding the signal of the regime shift (at $t = 50$) in system velocity, v (Fig. 5.11).

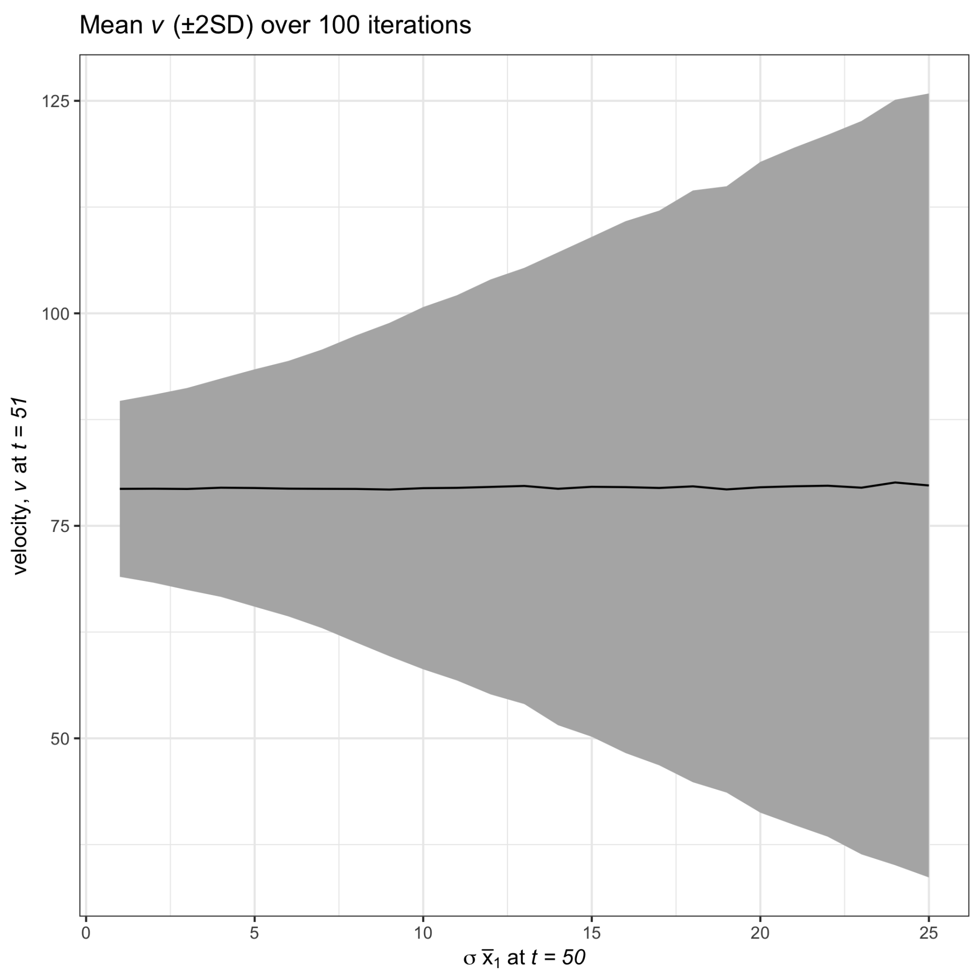


Figure 5.10: Average (± 2 SD) velocity (v) worsens as the variance of $\bar{x}_{2_{t=50(post)}}$ (post shift) increases. $\bar{x}_{1_{pre}} = 25$, $\bar{x}_{1_{post}} = 100$, $\bar{x}_{2_{pre}} = 25$, $\bar{x}_{2_{post}} = 50$, $\sigma_{1_{pre}} = 5$, $\sigma_{2_{pre,post}} = 5$

5.4 Velocity performance under a smooth transition

In the previous section I presented expectations for velocity signals under a discontinuous transition in a discrete-time system. Given velocity is a measure of the rate of a change in a system and the range of transition speeds ecological systems exhibit (e.g., slow driver-response or threshold dynamics), it is important to understand if and when the velocity signal is dampened under varying degrees of transition speeds.

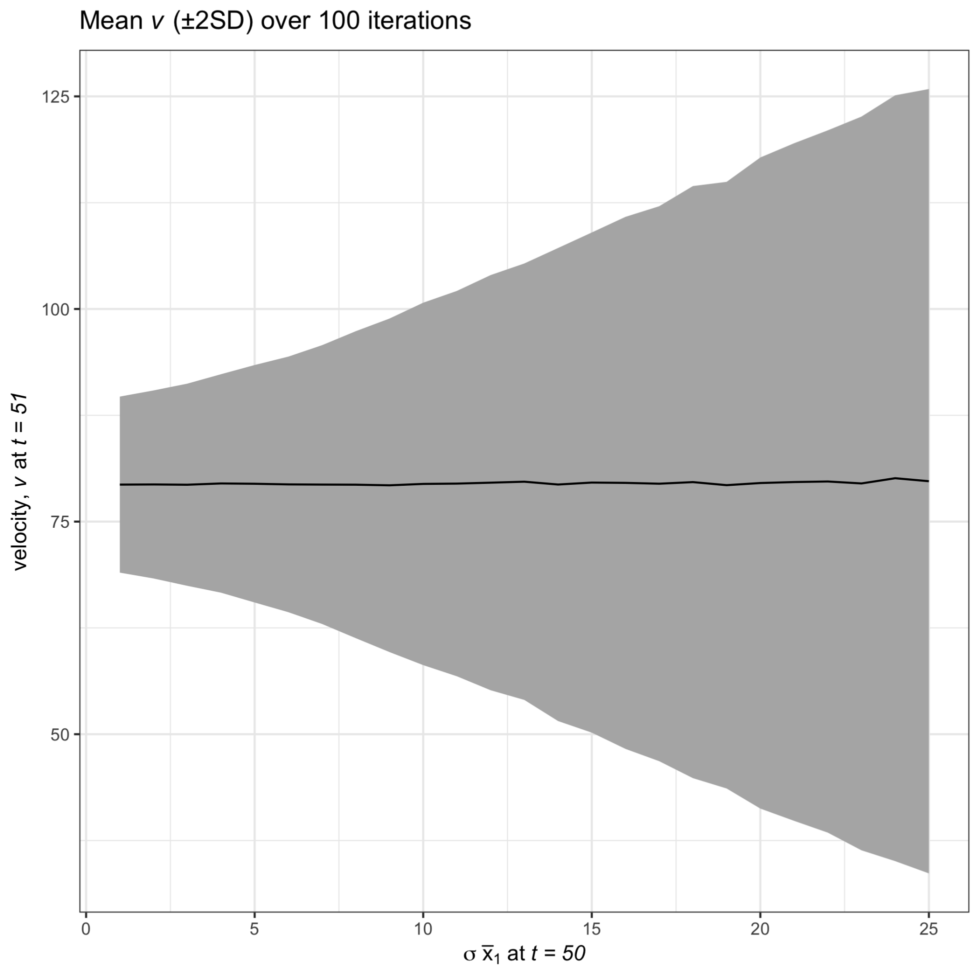


Figure 5.11: The noise in system velocity (v) is not obviously reduced in this system as the original data (x_1, x_2) is increasingly smoothed.

In this section I use a similar toy system, to demonstrate the expectations of velocity under a smooth shift and under varying degrees of rapidity.

Although the data constructed in this section are similar to that used in the previous section in that we are manipulating the mean and variance of two state variables before and/or after an abrupt shift, this section introduces a component of process noise into the shift itself. This is important because the derivative of a nearly discontinuous function is infinite. Although we are interested in identifying rapid shifts in systems, velocity will appraoch infinity as the rate of change in the shift increases and the sampling intervals decrease. In other words, if the system exhibits turnover in e.g. 25%

Table 5.2: Conditions for generating various scenarios of the hyperbolic tangent-induced abrupt change. σ_i represents the standard deviation of μ_{x_i} as the percent of μ_{x_i}, μ_{x_i} is the mean of the state variable, x_i, and pre and post represent the periods before and after the regime shift at $t = 50$, respectively.

conditions	$\sigma_{x1_{pre}}$	$\sigma_{x1_{post}}$	$\sigma_{x2_{pre}}$	$\sigma_{x2_{post}}$	$\mu_{x1_{pre}}$	$\mu_{x1_{post}}$	$\mu_{x2_{pre}}$	$\mu_{x2_{post}}$
μ_{x_1}, μ_{x_2}, σ_{x_1}, σ_{x_2}	0.05	0.10	0.05	0.10	10	55	15	44
μ_{x_1}, σ_{x_1}	0.05	0.10	0.05	0.05	10	55	15	15
μ_{x_1}, μ_{x_2}	0.05	0.05	0.05	0.05	10	55	15	44
μ_{x_1}	0.05	0.05	0.05	0.05	10	55	15	15
σ_{x_1}, σ_{x_2}	0.05	0.10	0.05	0.10	10	10	15	15
σ_{x_1}	0.05	0.10	0.05	0.05	10	10	15	15

of the state variables, we expect the value of velocity to be similar to that of a turnover in e.g. 75% of the variables. Removing the possibility of infinite values provides more relative measures within the community time series.

5.4.1 Generating the data

Here we consider a two-variable system over the time interval $[1, 100]$ with state variables x_1 and x_2 which exhibits abrupt shifts in mean and/or variance of one or both variables at time $t = 50$. I generated species observations for the true process and the true process with process variability. The true process data were created using the paramters for μ and σ for each of the conditions in described in Table 5.2 (random seed in Program R was 12345).

True process model

Data were generated from a normal distribution and an abrupt shift in the mean was incorporated using a hyperbolic tangent function. The true process for each state variable, x_i, was generated from Eq. ((5.5) (see Fig. 5.13)):

$$\begin{aligned}
\mu_{xipre} &\sim Normal(\mu_{xipre}, \sigma_{xipre}) \\
\mu_{xipost} &\sim Normal(\mu_{xipre}, \sigma xipost) \\
\mu_{x_i}(t) &= \mu_{xi}pre - 0.5(\mu_{xi}pre - \mu_{xipost})(\tanh(\alpha(t - t_{shift})) + 1) \\
&f
\end{aligned} \tag{5.5}$$

where $\mu_{x_i}(t)$ is the mean value of x_i at time t and *pre* and *post* are the periods before and after the abrupt shift (t_{shift}), respectively. The parameter α in Eq. (5.5) controls for the rate of change at the point of the abrupt change, t_{shift}, where higher values of α correspond with a higher slope at t_{shift}. I simulated a single iteration (dataset) for various conditions of changing μ_{xi} and σ_{xi} (see Table 5.2), for two state variables x_1 & x_2 at intervals of $t = 1$ along the temporal interval $t = [1, 100]$.

Observed process data

I generated observations by imputing noise into the true process model (Eq. (5.5)) through random sampling of σ_{xi} from a normal distribution (Eq. (5.6); Fig. 5.13):

$$\begin{aligned}
\mu_{xipre} &\sim Normal(\mu_{xipre}, \sigma_{xipre}) \\
\sigma_{xipre} &\sim Normal(0, \sigma_{Xipre}\mu_{Xipre}) \\
\mu_{xipost} &\sim Normal(\mu_{ipost}, \sigma_{ipost}) \\
\sigma_{xipost} &\sim Normal(0, \sigma_{Xipost}\mu_{xipost}) \\
\mu_{xi}(t) &= \mu_{xi}pre - 0.5(\mu_{xipre} - \mu_{xipost})(\tanh(\alpha(t - t_{shift})) + 1)
\end{aligned} \tag{5.6}$$

where σ_{xi} is the observed error around μ_{xi}, and σ_{Xi} is $X\%$ of μ_{xi} under various sampling conditions (as described in Table 5.2). I generated the error as a percent of the mean as this scaling relationship is commonly observed in ecological data (Taylor, 1961).

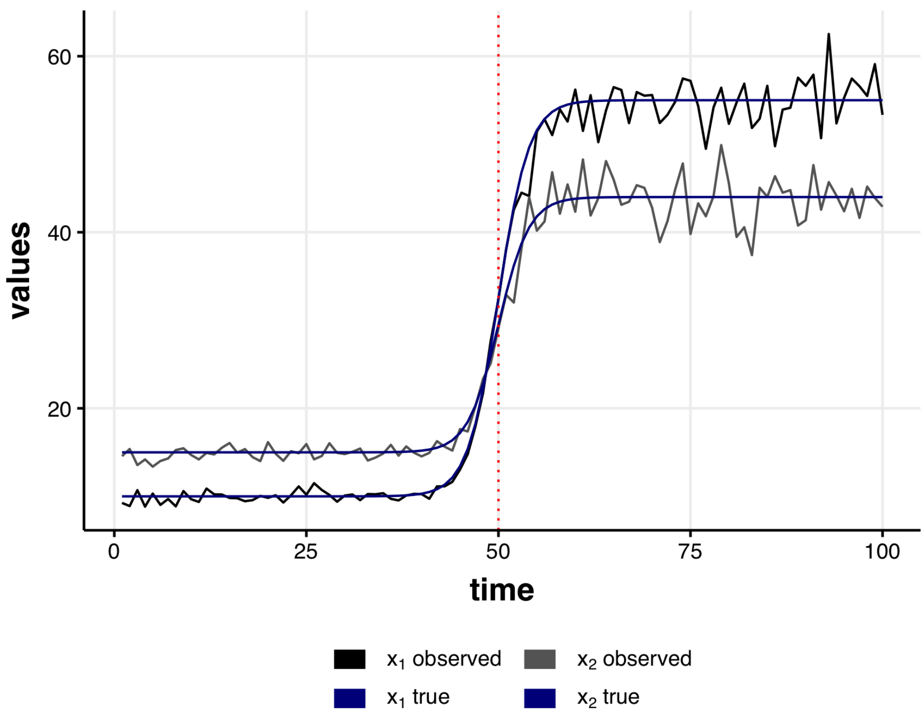


Figure 5.12: An example of the data generated by the true process model. In this example the mean values (μ_{xi}), but not the percent standard deviation (σ_{xi}), are varied before and after the transition point. The observed data are plotted against the true-process model for each state variable, x_i. Panels represent different degrees of the smoothing parameter, α (top: $\alpha = 0.25$, bottom:$\alpha = 1.00$).

5.4.2 Evaluating velocity performance under conditions of changing means and/or variance

I simulated a single dataset (R seed 12345) by randomly drawing a single realisation (observed data) of the hyperbolic tangent process model with additive process noise (Eq. (5.6)). I then calculated the distance travelled, s, and the velocity of the distance travelled, v (also referred to as $\frac{ds}{dt}$) using first differences. The first differences approach is a simple alternative to numerical integration tecniques, requiring only simple algebraic techniques. This method is ideal for discrete time data, or where computational power woudl not suffice for numerical integration. When using the first differences method, however, v will demonstrate high variability, depending on the

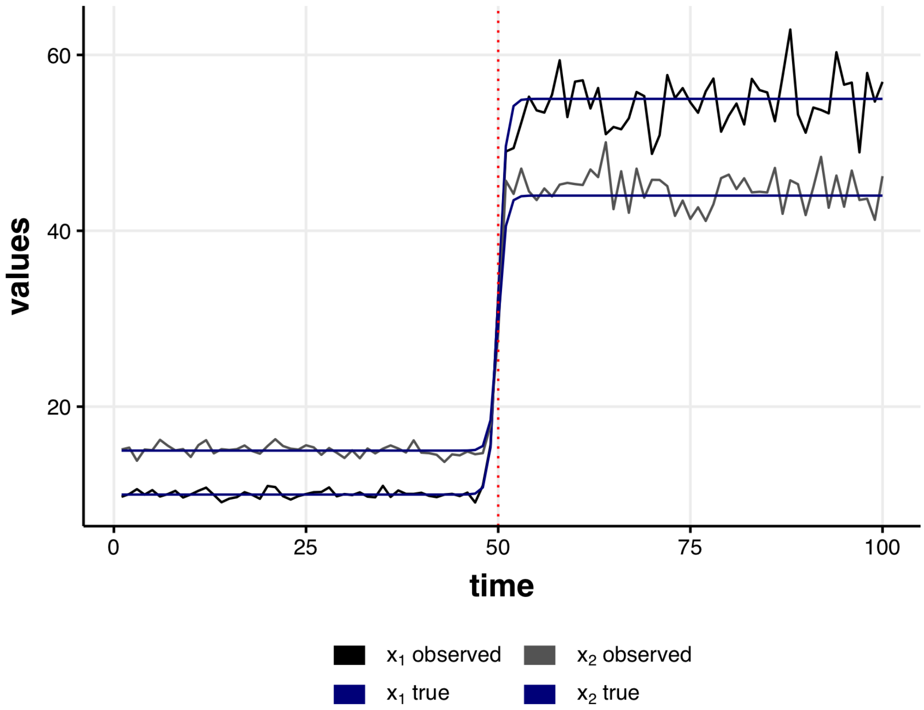


Figure 5.13: An example of the data generated by the true process model. In this example the mean values (μ_{xi}), but not the percent standard deviation (σ_{xi}), are varied before and after the transition point. The observed data are plotted against the true-process model for each state variable, x_i. Panels represent different degrees of the smoothing parameter, α (top: $\alpha = 0.25$, bottom:$\alpha = 1.00$).

amount of time between samples (i.e. as the intervals of $t - t + 1$ increase).

I also calculated v using a numerical integration method for non-smooth, noisy data, called total variation regularized differentiation (Chartrand, 2011). I used the R package `tvdiff` (Price & Burnett, 2019) to perform numerically integrate the distance travelled, s. The regularized differentiation method in this package (function `tvRegDiff`; descried fully in Chartrand, 2011) provides a numerical solution for calculating non-noisy derivatives of noisy, non-smooth data. Using this smooth-derivative estimation technique may be an ideal supplement to the velocity method in cases where process and observational error generate noisy observational data. Although not possible in most ecological systems data, here we can compare the fit of the smooth-derivative to the derivative of the true process, allowing us to determine the usefulness of calculating a smooth-derivative. There are two tuning paramters required to be chosen by the analyst when implementing the total-variation regularized differentiation, each of which influence the amount of noise smoothed out in the resulting derivative: α and the number of iterations. I implemented this numerical differentiation over 1,000 iterations, and selected α by comparing the antidifferentiated distance travelled, s, to the true process values of s (e.g., see Figure 5.14). For most conditions and smoothness I found the tuning parameter for `tvdiff` $\alpha = 0.50$ provided a good fit of s (Fig. 5.14), however, when the hyperbolic tangent smoothing paramter, α was low (i.e. $\alpha_{tanh} = 0.25$) higher values of α_{tvdiff} yielded more abrupt changes in the derivative.

Smooth changes in the mean

As discussed earlier, the velocity of the distance travelled, v, is a measure of how quickly the sum of the squared system variables change between observations (i.e. time). Consequently, as the total change in state variables grows, so will the maximum potential of the velocity, v. Following this logic, we should expect to see a spike in the derivative of the distance travelled when the system changes quickly. I tested this hypothesis under two conditions of changing means, where either one or both variables underwent mean shifts (see Table 5.2), and under varying degrees of transition smooothness (i.e. $\alpha_{tanh} = 0.25, 0.50, 0.75, 1.00$).

When the hyperbolic tangent smooth transition function is less steep (Fig. 5.15) the observed velocity signal is dampened. This signal, however, is quickly recovered when the transition function becomes more abrupt (Figs. 5.16,@ref(fig:mu1varpt75,5.18; $\alpha_{tanh} = 0.5, 0.75, and 1.00$, respectively). Unsurprisingly, when we shift the means of both state variables while holding the relative variance constant, the velocity signal

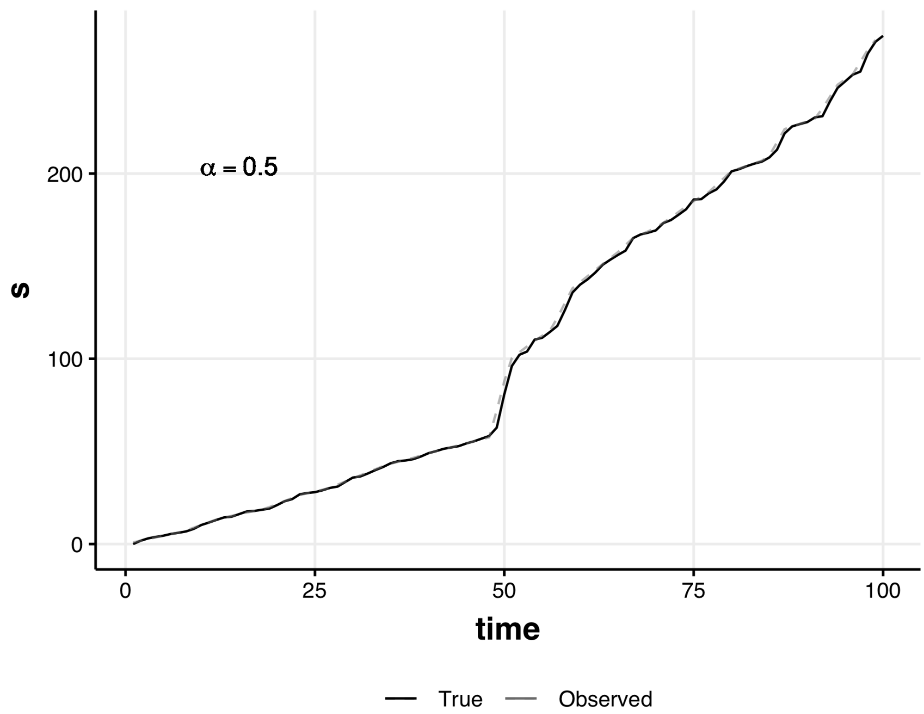


Figure 5.14: Antidifferentiated values ('observed') of the distance travelled, s, to the true process values of s ('true) provides a method for identifying the best values of the smooothing parameter, α. Under most conditions $\alpha \ll$ sufficed. Here, we compare the true and antidifferentiated values of s under the condition of changing μ_{x1} when the hyperbolic tangent function is most rapid ($\alpha_{tanh} = 1$) for the 'tvdiff' $\alpha = 0.50$. Not pictured: the antidifferentiated values of s (observed) is increasingly smoothed as α increases.

changes more abruptly (Fig. 5.19) than when only a single variable shifts mean value (compare with Fig. 5.15). Figure 5.20 is representative of the increasing signal in velocity as α_{tanh} increases.

Smooth changes in variance

Abrupt changes sometimes manifest first as a change in the variability, rather than the mean value, of the state variables. This condition manifests in the velocity signal when both variables experience a shift in relative variance (Fig. 5.21), however, does not signal change when only one variable exhibits a shift in variance (Fig. 5.22).

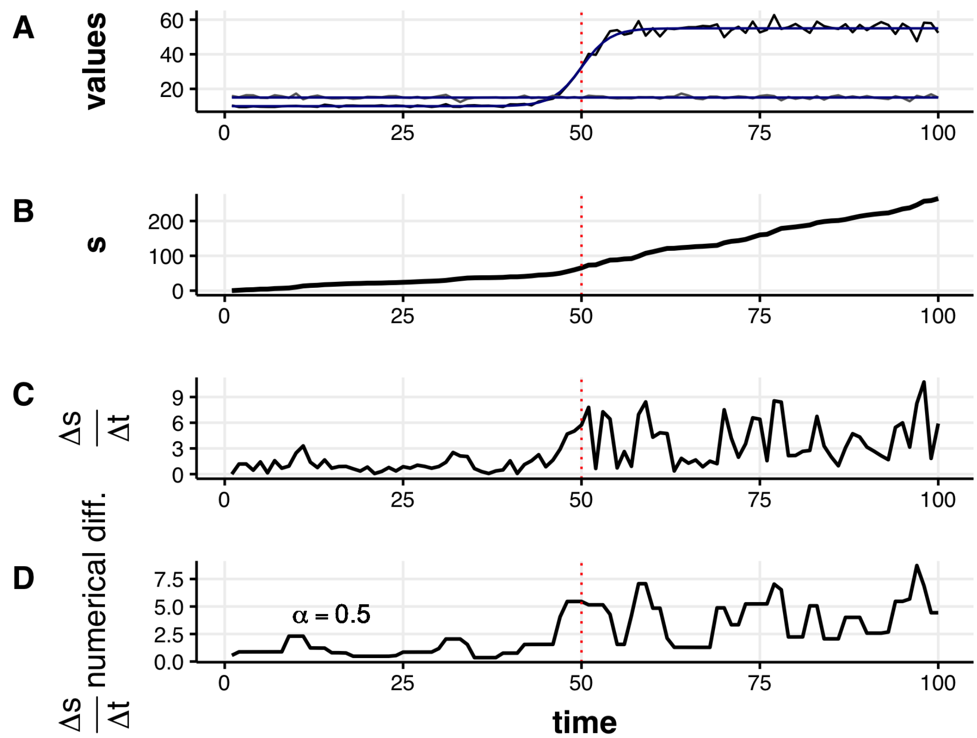


Figure 5.15: The velocity signal is muted when the hyperbolic smoothing parameter, α, is low (0.25). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

Again, given the total magnitude of change influences the distance travelled, s, and the derivative of s, v, it is not surprising that the velocity signal is greater around the transition point when both, compared to a single, state variable exhibits increased variability about the mean. In these scenarios I shifted the variability in the state variables x_i from only $\sim 5\%$ to $\sim 10\%$ (see Table 5.2)—this percent variability is low relative to most empirical observational ecological datasets. As such, I expect the velocity signal to be more pronounced when empirical systems undergo shifts in variance in at least one state variable.

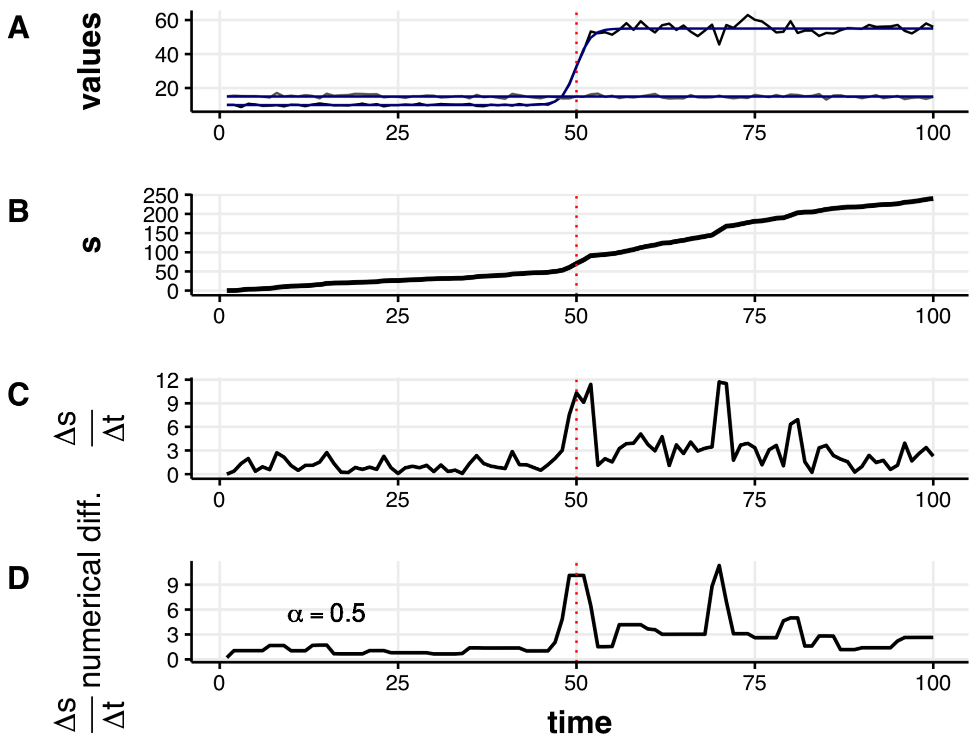


Figure 5.16: The velocity signal is muted when the hyperbolic smoothing parameter, α, is moderate (0.50). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

Smooth changes in the mean and variance

Given the signals identified in the velocity when one or both state varialbes exhibits a shift in mean and/or variance, it is unsurprising that even under smooth transitions (when $\alpha_{tanh} = 0.25$), velocity manifests as a signal of change (Fig. 5.23). This signal is most pronounced when the shift is abrupt (Fig. 5.24).

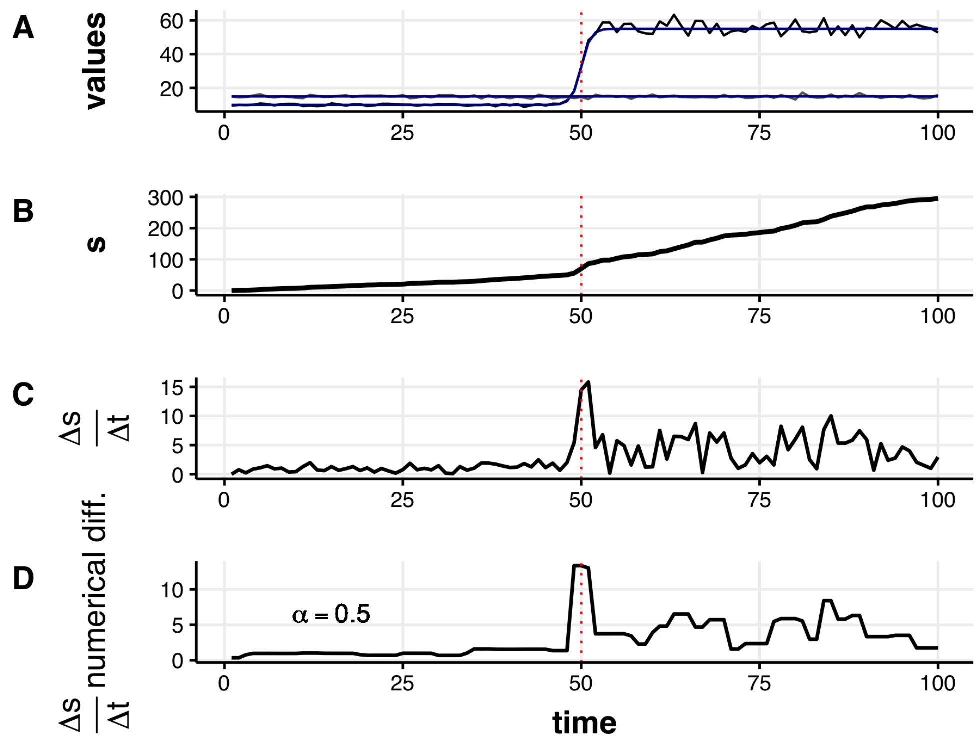


Figure 5.17: The velocity signal is muted when the hyperbolic smoothing parameter, α, is moderate (0.50). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

5.5 Velocity performance under empirical transitions: paleolithic freshwater diatom community

To gather baseline information on the use of velocity in empirical systems data, I calculated velocity for the paleodiatom system described in Chapter 6 (see also Appendix **??**. Briefly, the paleodiatom community comprises 109 time series over a period of approximately 6936 years (Fig. 5.25). As elaborated in Spanbauer et al. (2014), the paleodiatom community is suggested to have undergone regime shifts at multiple points. These abrupt changes are apparent when exploring the relative abundaces over time, as there are extreme levels of species turnover at multiple

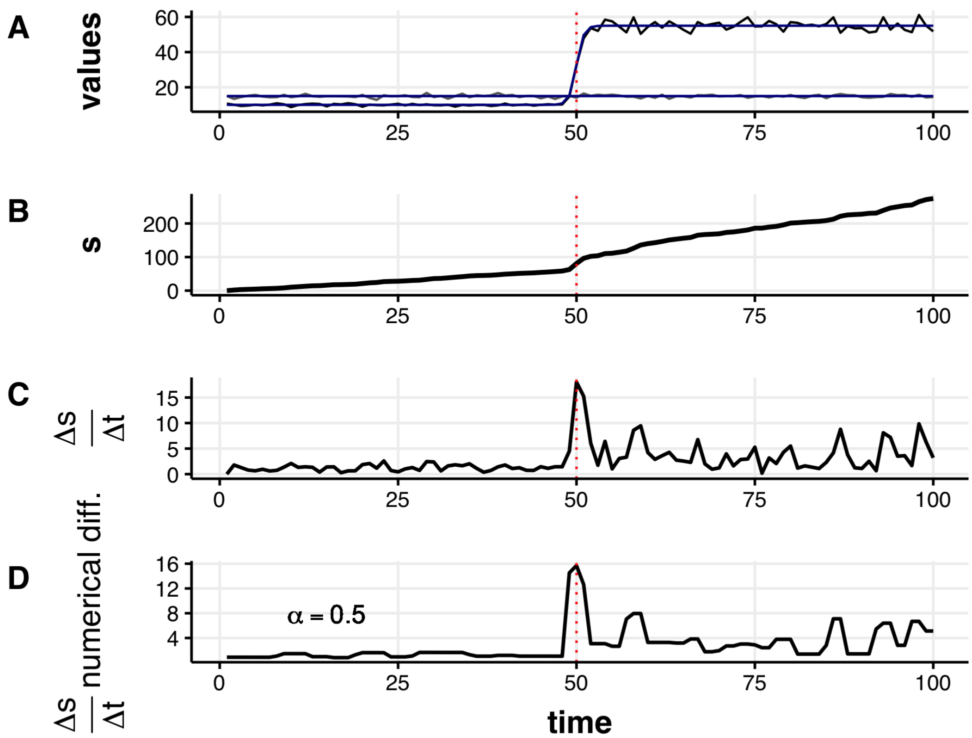


Figure 5.18: The velocity signal is muted when the hyperbolic smoothing parameter, α, is moderate (0.50). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

points in the data (Fig. 5.25). Using Fisher Information and climatological records, Spanbauer et al. (2014) suggest that regime shifts in this system at approximately 1,300 years before present (where present is equal to year 1950). Spanbauer et al. (2014) used different regime detection metrics coupled with regional climatological events to identify regime shifts in the system, suggest that a regime shift occurred at ~1,300 years before present. Using the methods outlined above, I calculated the distance travelled (s) and velocity (v; Fig. 5.29). The results of v and s (Fig. 5.26) on the relative abundance data correspond with both the large shifts in species dynamics (see Fig 5.25, and also with the regime shift identified by Spanbauer et al. (2014). However, two primary results can be made from the metrics v and s that are not obvious nor identified numerically in the results of Spanbauer et al. (2014) (): 1. Two additional large shifts occurred at approximately 2,500, 4,800 and years before 1950

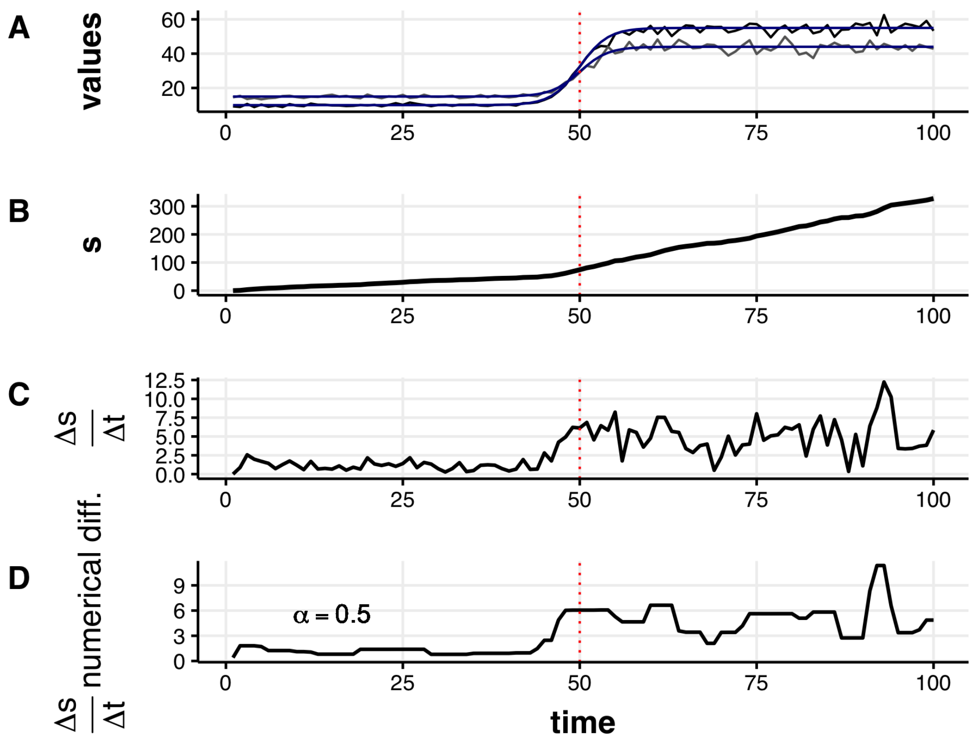


Figure 5.19: The velocity signal is regained under smoooth transition ($\alpha_{tanh} = 0.25$) when both state variables undergo a shift in the mean. True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

1. The periods before the first and after the second large shifts appear oscillatory (Fig. 5.27).

To determine whether removing the noise in the data, I interpolated the each time series using function `stats::approx` to 700 time points. Next, I calculated the distance travelled of the entire system, s. Finally, I obtained the derivative of s by using a regularized differentiation (using function `tvdiff::TVRegDiffR`; parameters were $iter = 2000$, scale = small, $ep = 1x10^{-}6$, and $\alpha = 100$). This method of regularized differentiation is an ideal approach to smoothing s because it assumes the data are non-smooth and incorporates finite differencing. The total variation regularized differentiation is desribed in Chartrand (2011), Price & Burnett (2019), and in the previous first-level section. The smoothed velocity (Fig. 5.29) provides a similar but smoother picture of the velocity of the system trajectory. Comparing the

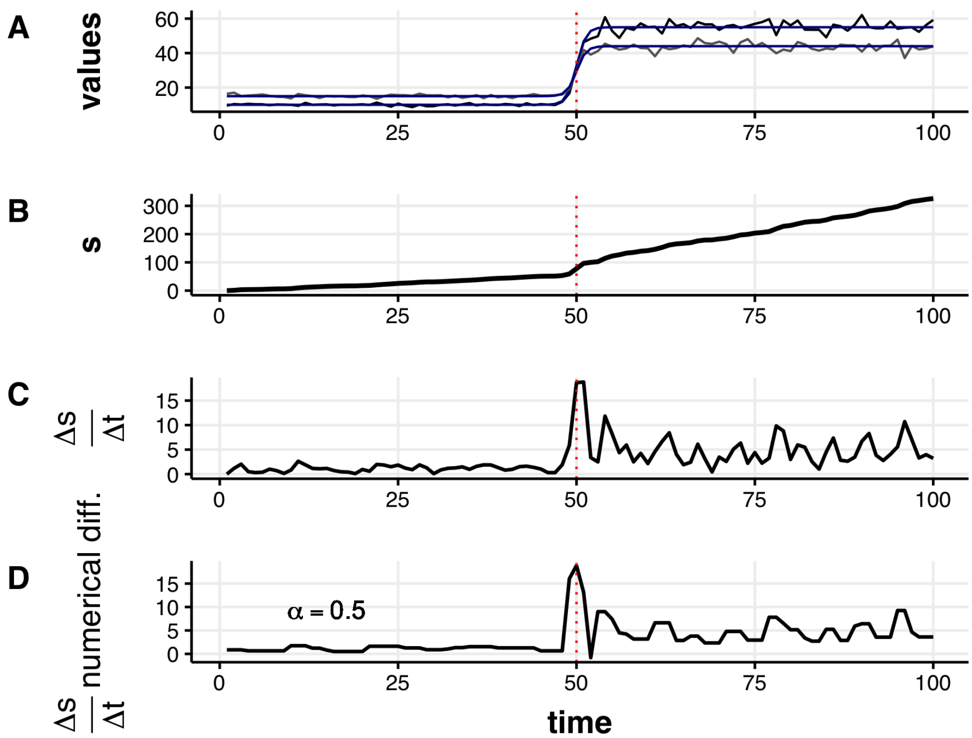


Figure 5.20: The velocity signal is regained under smoooth transition ($\alpha_{tanh} = 0.75$) when both state variables undergo a shift in the mean. True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

smoothed (Fig. 5.29) to the non-smoothed velocity (Fig. 5.26) yields similar inference regarding the location of the regime shifts at 2,200 and 1,300 years before present, however, it more clearly demonstrates potential inter-regime dynamics (e.g., between 7,000 and 4,800 years before present), which were not identified in previous study of this ssytem (Spanbauer et al., 2014).

5.6 Discussion

Here, I described the steps for calculating a novel regime detection metric, system velocity (v). First described in Fath et al. (2003), v is used as a single step for caclulating a more complicated regime detection metric, Fisher Information (see also Chapter 3). System velocity is arguably simple to calculate, as shown in this chapter,

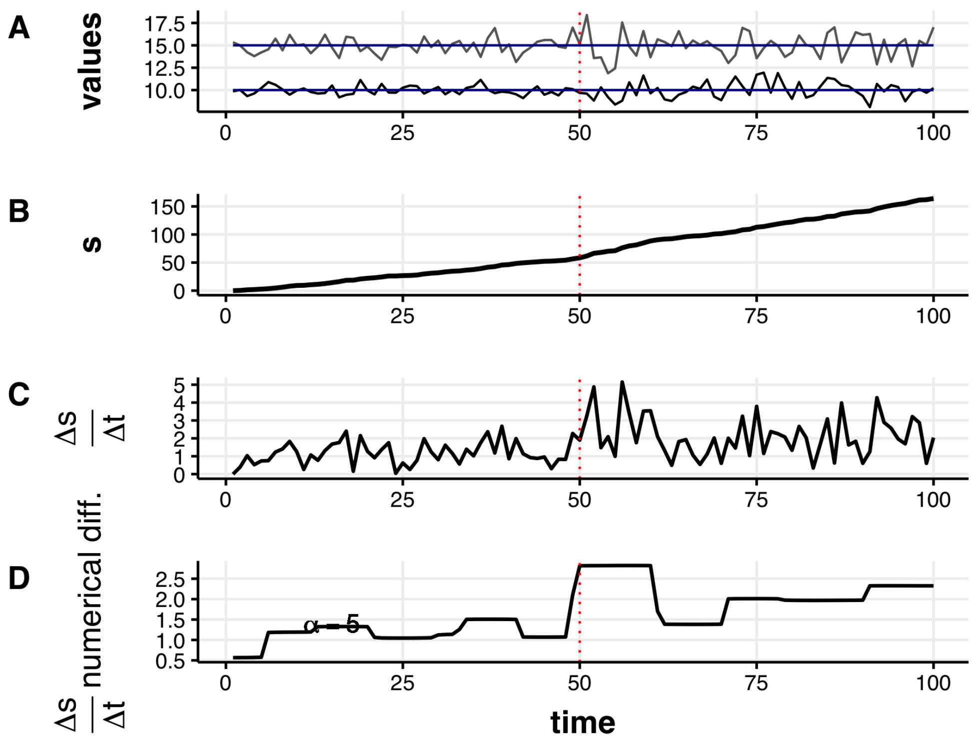


Figure 5.21: The velocity signals a rapid shift in the variance of both state variables under a moderately abrupt transition ($\alpha_{tanh} = 0.75$). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

captures the total change in system variables under a variety of mean and variance conditions. The metric does not, however, perform well as variance increases (Fig. **??**), and smooothing the original data does not reduce the noise surrounding this metric when variance is moderate (Fig. **??**). Variance is a commonly-used indicator of ecological regime shifts (Brock & Carpenter (2006)), however, is difficult to interpret when the number of variables is $\ggg$ a few. System velocity, v, may be useful in situations where the number of state variables is $\ggg$ few, and appears especially useful when the magnitude of change in one or more state variables is high (Figs. 5.8,5.20). For example, this method will likely identify signals of regime shifts where the shift is defined as high species turnover within a community (Fig. 5.24).

This study provides baseline expectations of the velocity of the distance travelled, V, as an indicator of abrupt change in a multivariable system. Although a useful first

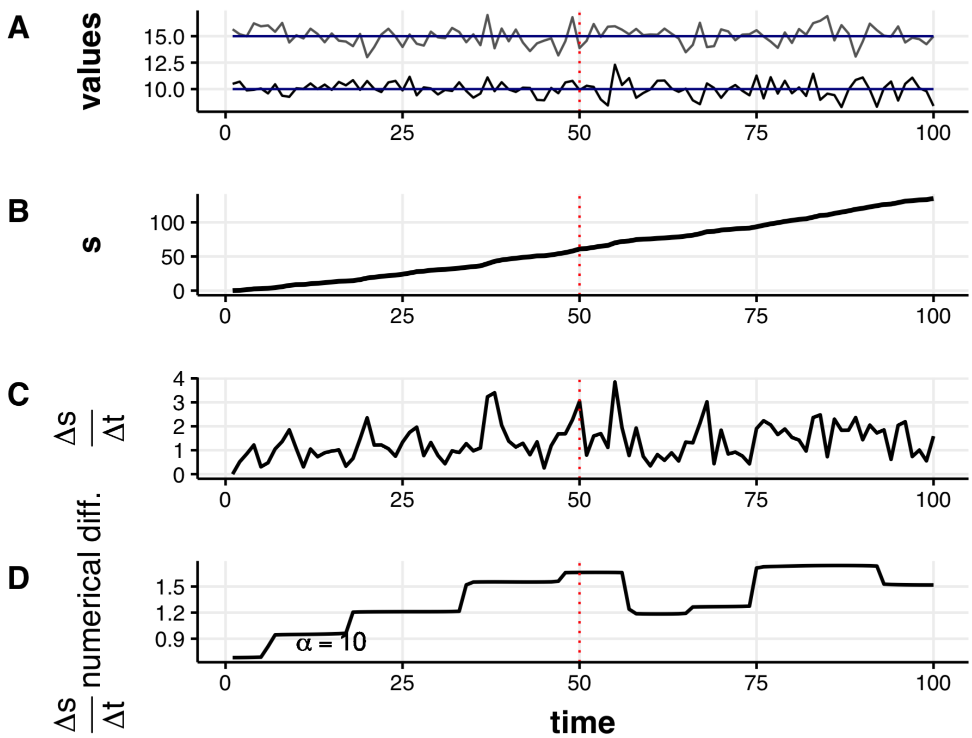


Figure 5.22: The velocity does not signal shifts in the variance of a single variable (x_1) under a moderately abrupt transition ($\alpha_{tanh} = 0.75$). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

step, this metric should next be critiqued in a sensitivity analytical approach, where a statistical measure is used to determine whether V indicates abrupt shifts prior to occurrence (c.f. during or after), particularly with respect to its performance in community-level empirical data. The paleolitic diatom data used in the last section of this chapter is also presented in the documentation for my R Package, **regimeDetectionMeasures** (Appendix **??**). In this case study, the 'distance travelled', s (Eq. (5.1)), clearly exhibits shifts at points where expert opinion and species turnover (in species dominance) agree that a large change occurred. Further, velocity, v (see *dsdt* in package materials) indicates a large shift at only the most predonimnant shift in the time series, perhaps due to the metric's sensitivity to variance (Fig. 5.8.

Further work is required to determine the utility of system velocity as a regime detection metric, however, this chapter demonstrates that the metric may indicate

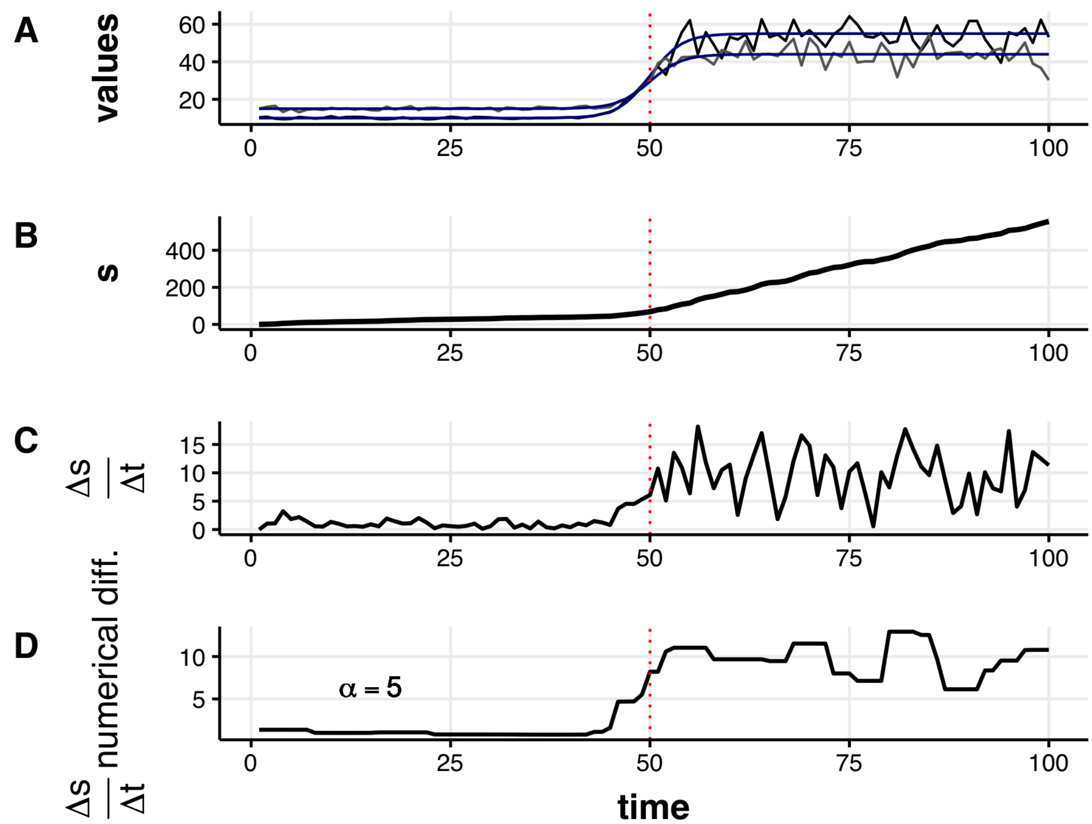


Figure 5.23: The velocity signals a shift when both variables undergo shifts in the mean and variance under a slightly abrupt transition ($\alpha_{tanh} = 0.25$). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

clear shifts in variable means and variability about the means. In addition to examining high-dimensional and noisy data, a study of the performance of v under conditions where few variables exhibit large changes while many variables are relatively constant may also prove useful. Additionally, this metric may be a useful tool for reducing the dimensionality of high dimensional data. Although the metric loses much information, as opposed to some dimension reduction techniques, e.g. Principal Components Analysis PCA, the metric is simple to calculate (even by hand), is computationally inexpensive, and is intuitive, unlike many clustering algorithms (e.g., Non-metric Multidimensional Scaling NMDS). Like system velocity, methods of the latter variety (e.g. NMDS) require post-hoc statistical analyses to confirm the location of clusters

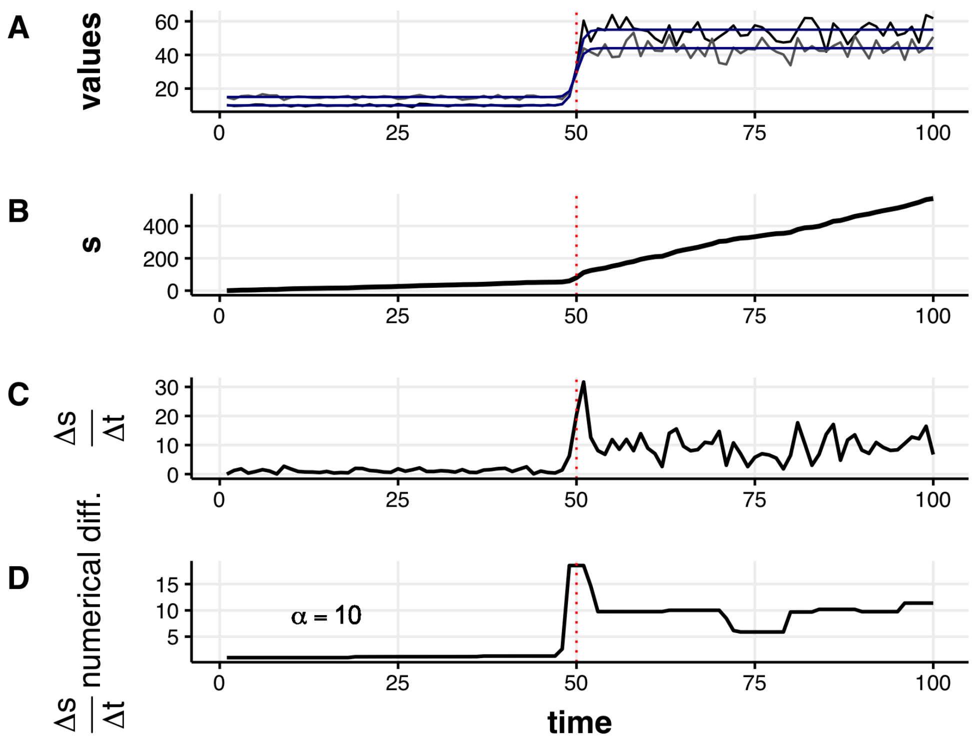


Figure 5.24: The velocity signals a shift when both variables undergo shifts in the mean and variance under a slightly abrupt transition ($\alpha_{tanh} = 1.00$). True and observed values of x_i (panel A), observed distance travelled (s, panel B), observed velocity (C), and the smoothed velocity (D).

(or abrupt change, regime shifts), while methods of the former variety (e.g. PCA) retain loadings but do not necessarily identify the locations of abrupt shifts.

5.7 Supplementary Figures

This section contains additional examples of the behavior of velocity, v when varying the mean and/or variance prior to and/or after the induced abrupt shiftt in the toy system with a discontinuous transition at $t = 50$.

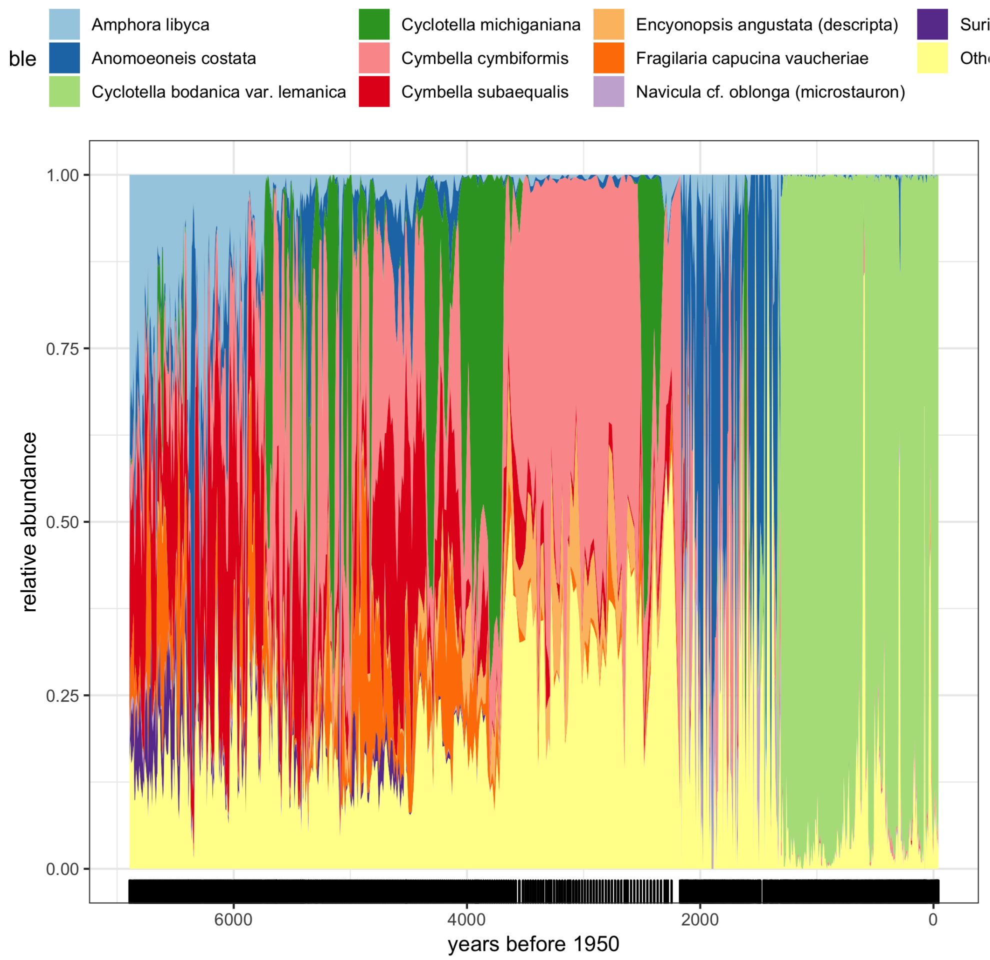


Figure 5.25: Relative abundances of the most common diatom species in the time series. Few species dominate the data over the entire time series, and turnover is apparent at multiple observations.

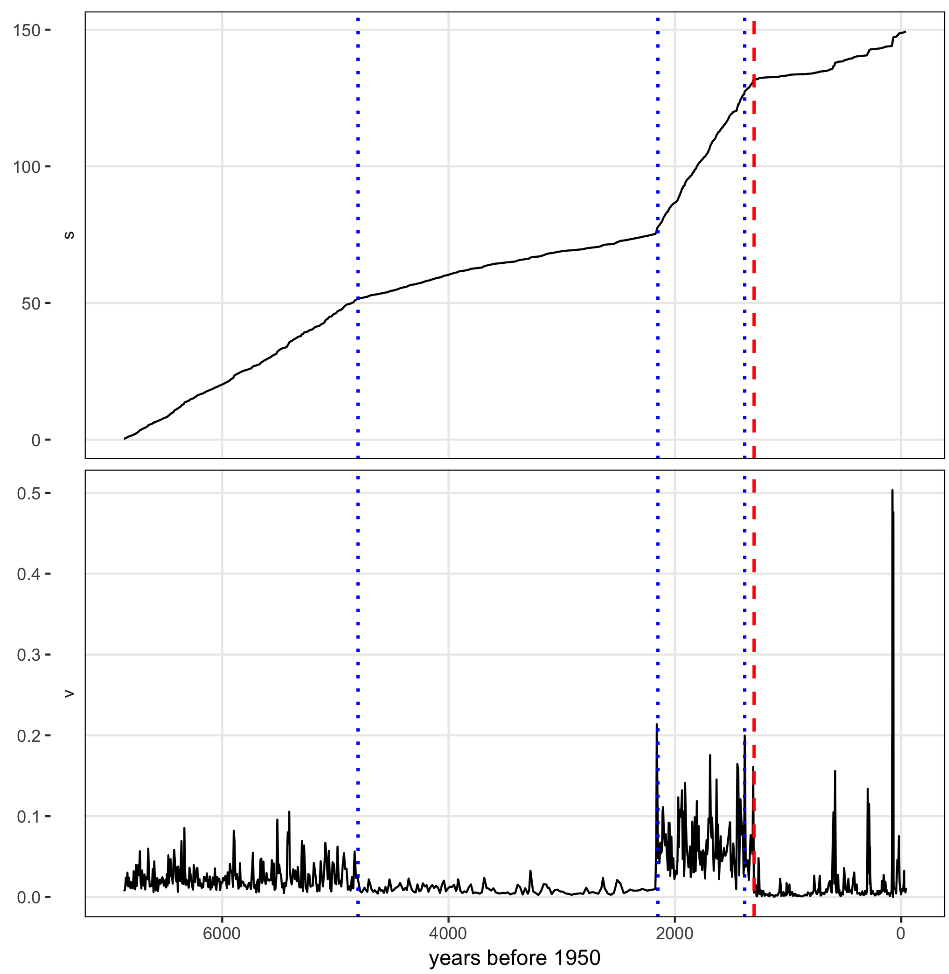


Figure 5.26: Velocity v and distance travelled s of the paleodiatom time series. Dashed line at 1,300 years before 1950 indicates the regime shift identifed in Spanbauer et al. (2014). Dotted lines indicate regime shifts as visually identified on metrics s and v.

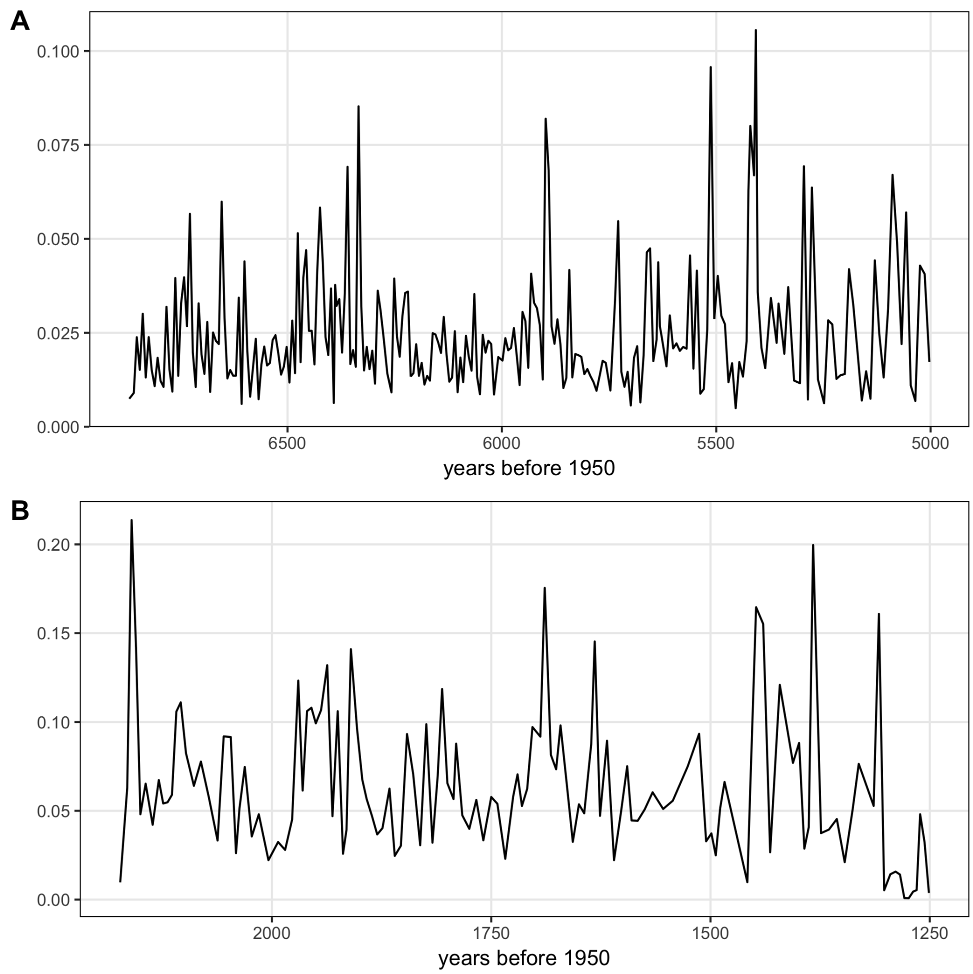


Figure 5.27: Velocity (v) indicates periodic-like conditions in the first (A) and second (B) regimes.

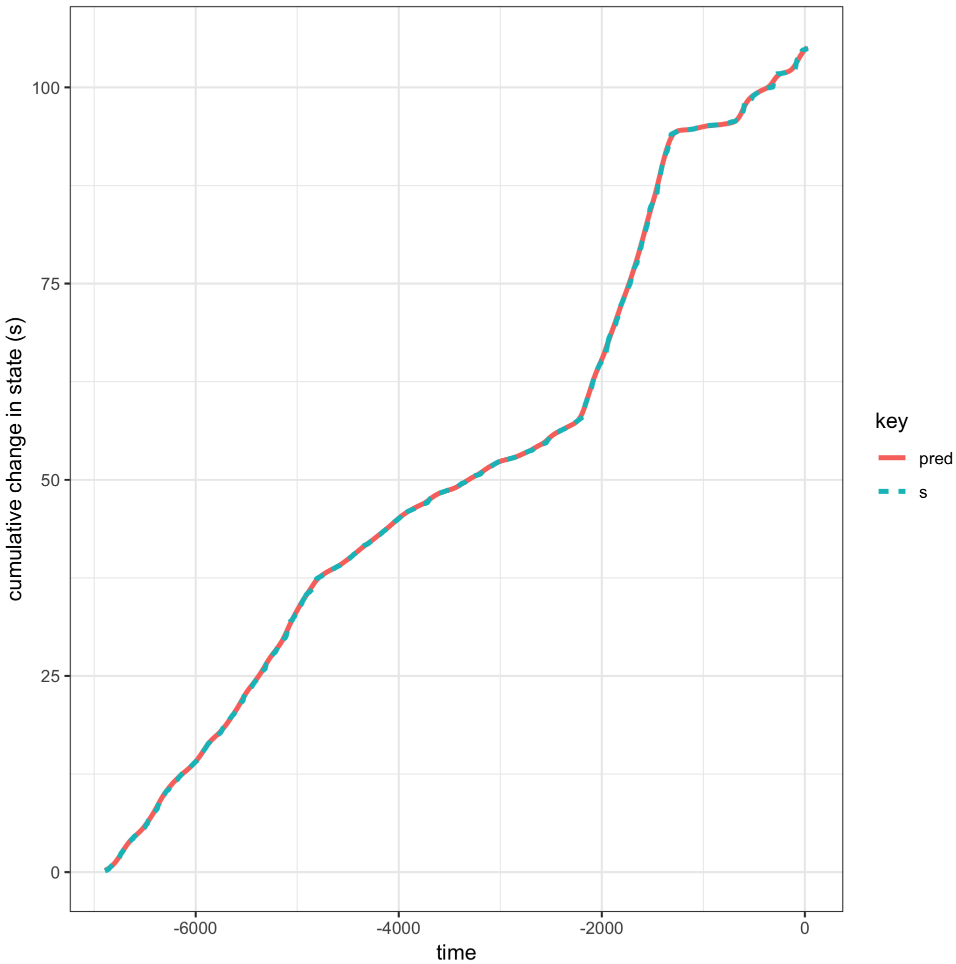


Figure 5.28: The regularized differentiation of s was best fit using $\alpha = 100$. Higher overlap of s and pred indicates a good fit of the regularized differentiated metric to the non-smoothed metric, s.

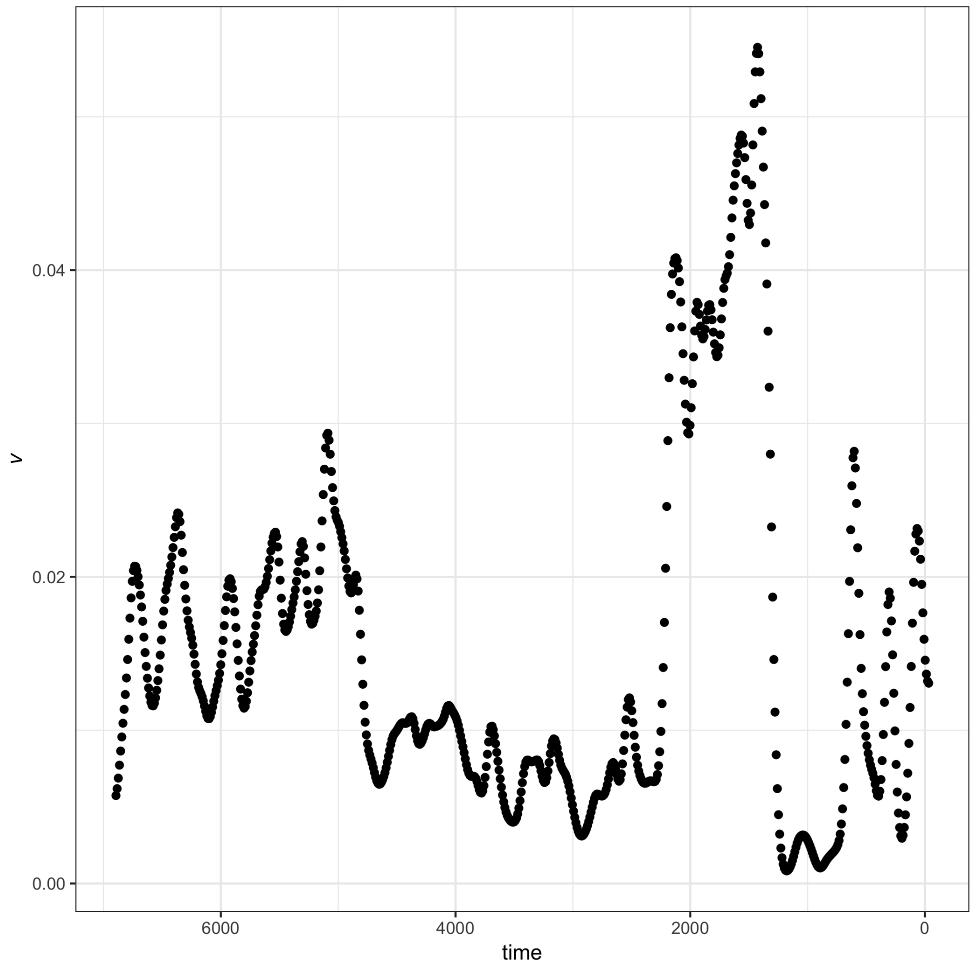


Figure 5.29: The velocity metric (v) signals potential periodicities in the paleo diatom time series data when the distance travelled metric, s, is smoothed using regularized differentiation methods [@price2019tvdiff].

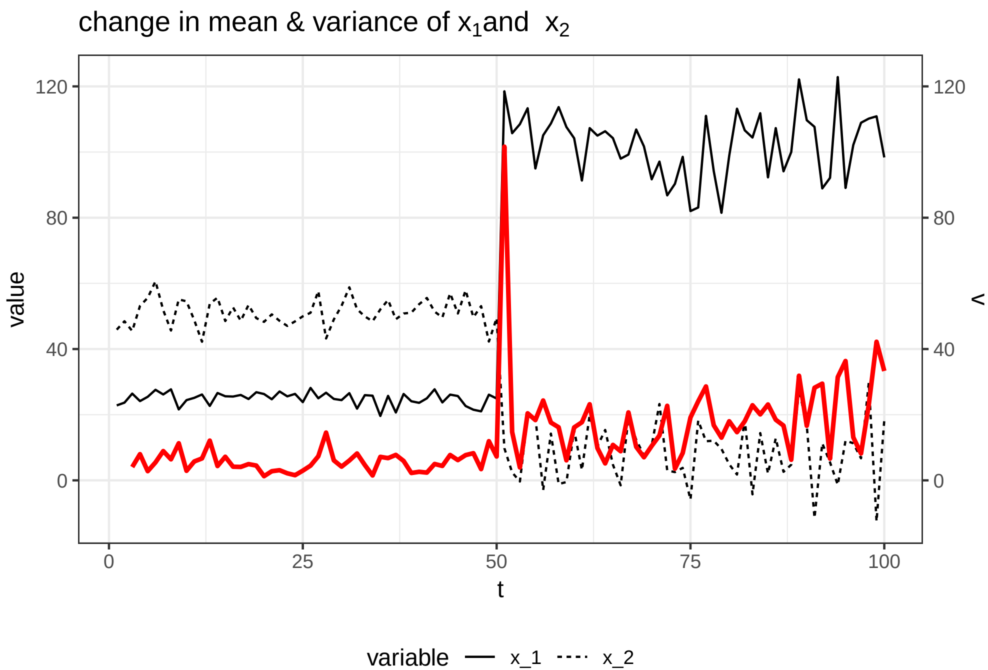


Figure 5.30: System change (s) and velocity (v) of the model system over the time period. Change in means ($\bar{x}_{1_{pre}} = 25$, $\bar{x}_{1_{post}} = 100$, $\bar{x}_{2_{pre}} = 50$, $\bar{x}_{2_{post}} = 10$) and an increase in variance ($\sigma_{1_{pre}} = 2$, $\sigma_{1_{post}} = 10$, $\sigma_{2_{pre}} = 5$, $\sigma_{2_{post}} = 10$).

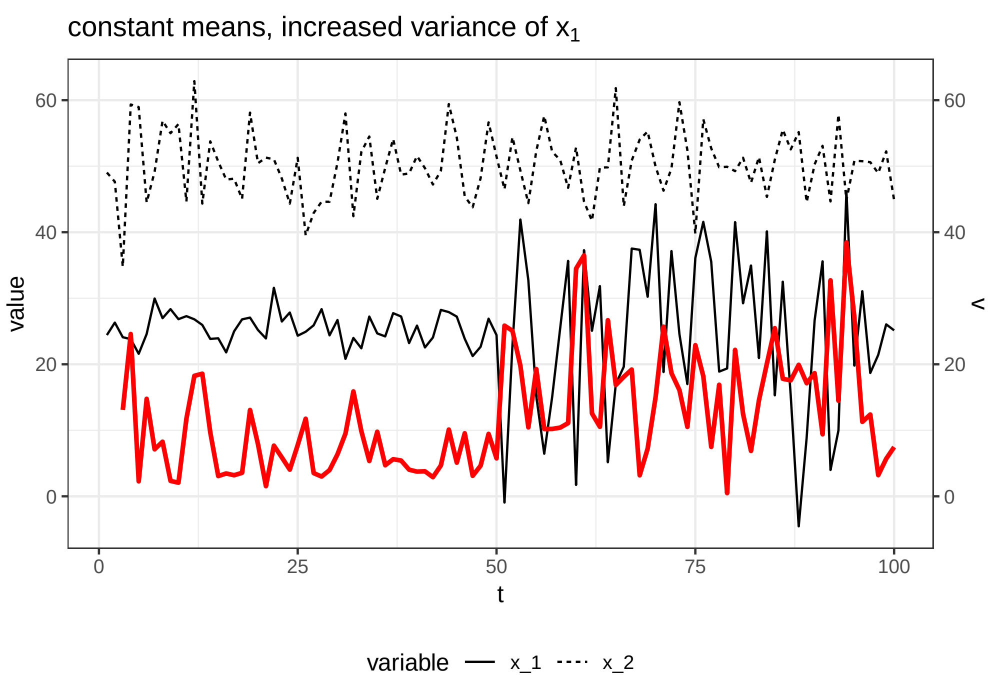


Figure 5.31: System change (s) and velocity (v) of the model system over the time period. Constant means ($\bar{x}_1 = 25$, $\bar{x}_2 = 50$) and sharp change in variance for one state variable $\sigma_{1_{pre}} = 2$, $\sigma_{1_{post}} = 12$, $\sigma_{2_{pre,post}} = 5$

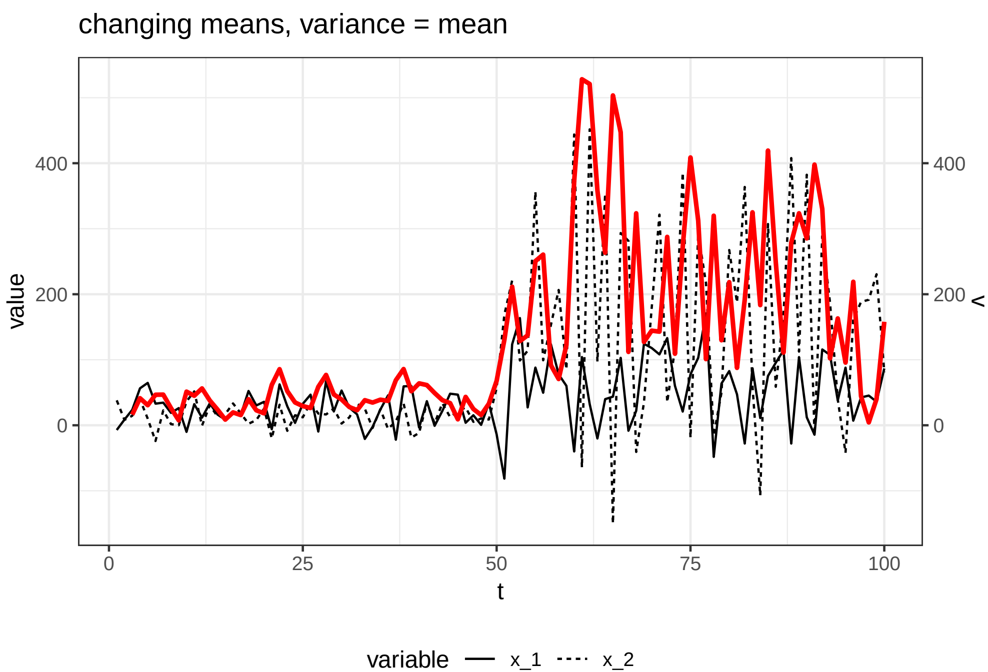


Figure 5.32: System change (s) and velocity (v) of the model system over the time period. Variance equal to mean ($/bar x_i = /sigma_i$), where means ($/bar x_{1_{pre}} = 25$, $/bar x_{1_{post}} = 50$, $/bar x_{2_{pre}} = 15$, $/bar x_{2_{post}} = 150$).

Chapter 6

Robustness of Multivariate Regime Detection Measures to Varying Data Quality and Quantity Using Resampling

6.1 Introduction

Ecological systems have many unpredictable and variably interacting components (Jorgensen et al., 2011). Methods for analyzing these complex systems, e.g. Dynamic Bayesian Networks, network models, and food webs are designed to handle these complexities, yet require data- and knowledge-intensive models. Although ecological data collection and data management techniques are improving (La Sorte et al., 2018), the aforementioned approaches to modeling and understanding complex system are often infeasible in ecosystem research and management (Clements & Ozgul, 2016).

A growing concern with anthropogenic impacts on the environment has increased the demand for mathematical and statistical techniques that capture these dynamics. These often undesirable changes in the structure or functioning of ecological systems are often referred to as *regime shifts*, *regime changes*, *state change*, *abrupt change*, etc. (Andersen et al., 2009) . A yet-unattained goal of ecological research and management is to reach a point where these methods can predict impending regime shifts in real-time and with high confidence. Ideally, ecological regime shift detection methods (hereafter, regime detection measures) would require little knowledge of the intrinsic drivers of the system, and the users of the method would not be required to know if and where a regime shift occurred in the data.

Despite the suite of regime detection measures in the environmental and ecological research literatures, they are not used in ecological management. We can describe the current state of regime detection measures as being either system specific (i.e., the method is not system agnostic) or not. Methods of the latter type are convenient in that they can be applied across various system and data types, but the results of these analyses require some degree of subjective interpretation (Clements & Ozgul, 2018; *c.f.* Batt, Carpenter, Cole, Pace, & Johnson, 2013). Efforts to develop and/or improve regime detection measures that do not require such subjectivity will aid the advance of regime detection measures research and application.

Current efforts to improve regime detection measures may be stunted by the lack of application beyond simple and/or theoretical (toy) systems data. Like most statistical and mathematical approaches, the evolution of many regime detection measures begins with application to theoretical data, followed by application to empirical data. Current applications of regime detection measures to empirical, ecological data are largely limited to data describing populations (Alheit et al., 2005; Anderson & Piatt, 1999; deYoung et al., 2008), climatic, marine, and Paleolithic regime shifts (Kong et al., 2017; Spanbauer et al., 2014; Yang & Wu, 2006), with few applications to terrestrial data (*c.f.* Bahlai, Werf, O'Neal, Hemerik, & Landis, 2015; Sundstrom et al., 2017). Although testing the performance and inference boundaries of theoretical and simple systems is important, they are of little use to ecosystem managers if they are not proven to be easily and reliably applicable to their system. Additionally, regime detection measures should be capable of handling empirical ecological data, which are often sparse, noisy, and irreguarly sampled..

Ecological systems data is expensive to capture, and has large process variation and observation errors. This variability reduces data quality and quantity, limiting the numerical tools for identifing trends and changes in the system (Thrush et al., 2009). Some methods, new and old, proposed as regime detection measures are purported to handle the data limitation and quality issues inherent in ecological data, and minimize subjective decisions for choosing state variables and interpreting results. For example, variable reduction techniques, e.g. principal components analysis (Andersen et al., 2009; Reid et al., 2016; S. Rodionov & Overland, 2005), clustering algorithms (Weijerman, Lindeboom, & Zuur, 2005; Weissmann & Shnerb, 2016), an index of variance (Brock & Carpenter, 2006), and Fisher Information (Cabezas & Fath, 2002; Fath & Cabezas, 2004; Karunanithi et al., 2008) were introduced as methods which collapse the system into a single indicator of ecological regime shifts. Although these methods have been used on empirical ecological systems data, their robustness to empirical data quality and quantity have yet to be examined.

In this Chapter I examine the influence of observation and process errors on the inference obtained from select multivariable regime detection measures. There are three major objectives:

1. Identify the effects of data quality on regime detection measure inference.

2. Identify the effects of data quantity on regime detection measure inference.

3. Explore the relative performance of velocity (described in Chapter 5) to the above mentioned methods under multiple scenarios.

This Chapter provides baseline relative performance estimates of select, multivariable regime detection measures under various scenarios of data quality and quantity. The results from this Chapter inform the practical ecologist of the potential limitations to consider when applying these regime detection measures to their data, and has potential to inform the data collection process. Additionally, the software accompanying this Chapter allows the end user to implement these methods on this diatom system, a toy system, or their own data.

6.2 Data and Methodology

6.2.1 Study system and data

I used paleodiatom time series from a freshwater system in North America (Foy Lake, present day Montana) that apparently underwent rapid shifts in algal community dynamics at multiple points in time. This data comes from a single soil core sample, from which the relative abundances of 109 diatom species were identified at 768 observations (time points) over $\approx 7,000$ years (Figure 6.1. Althouh the soil core was sampled at regular distances, the soil accumulation process is not necessarily linear over time, resulting in irregularly-sampled observations (i.e., time elapsed between sampling points differs varies; see Figure 6.2). The data were published in Spanbauer et al. (2014) and can be downloaded at the publisher's website.

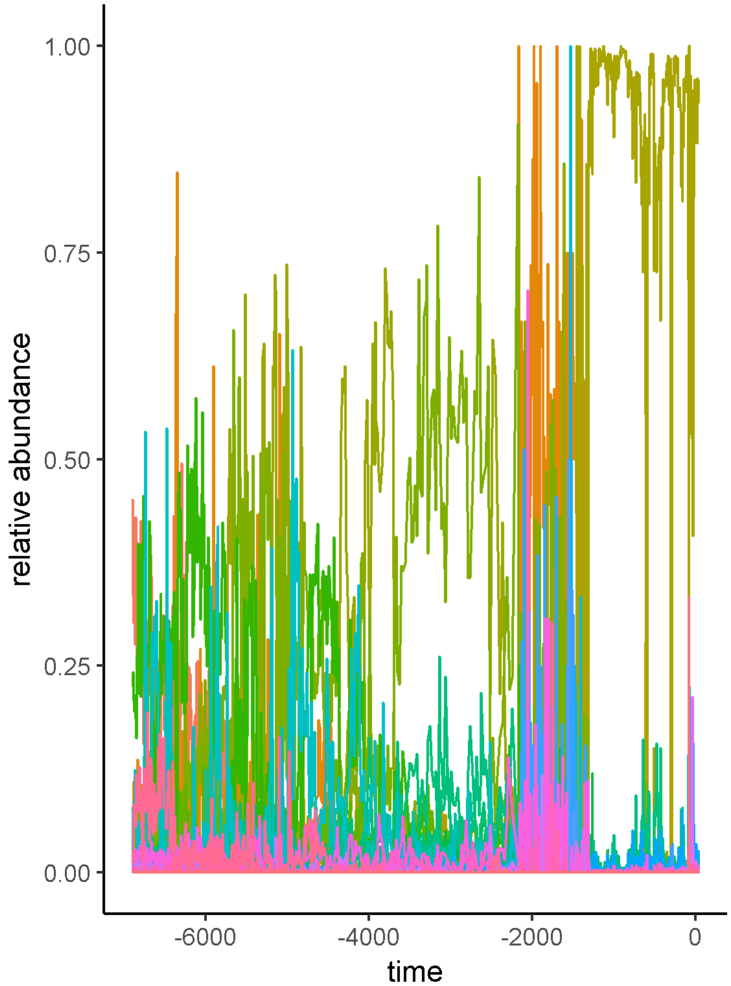


Figure 6.1: Relative abundances of the diatom species in Foy Lake over the time period.

6.2.2 Regime detection measures

Fewer model-free regime detection metrics exist than do model-based metrics (Chapter 2) and of these, only a few are suggested for multivariable data. Here, I compare the results for three regime detection metrics that are model-free and can handle multivariable data: velocity (Chapter 5), the Variance Index (Brock & Carpenter, 2006) and Fisher Information (Fath et al., 2003). I chose the Variance Index, as this is one of the more widely applied multivariate, model-free regime detection measures, and has been shown to, in some empirical data, identify regime shifts *post hoc*. I

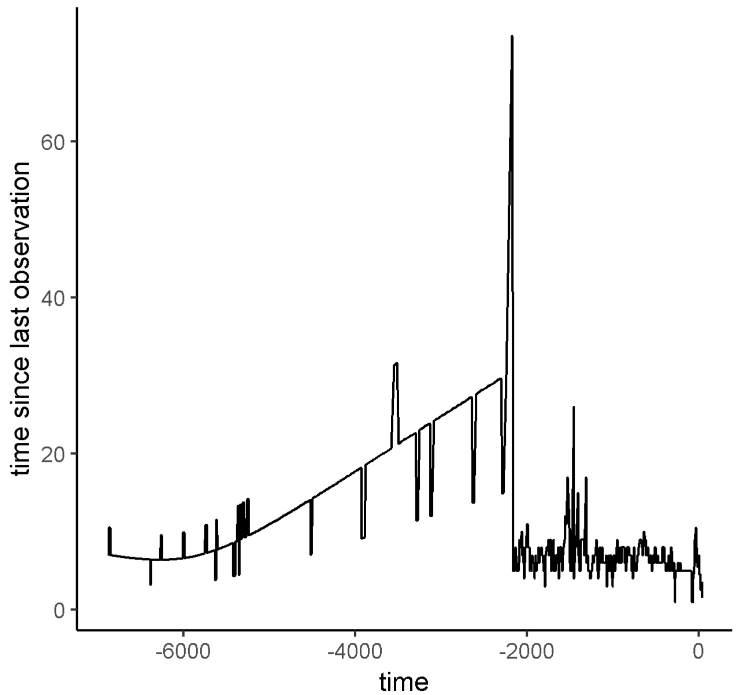


Figure 6.2: The amount of time elapsed between observations.

introduced the velocity in Chapter 5 as a new, potential regime detection metric. As this is the first time it has been used for such a purpose, including it in this approach allows us to further identify potential flaws with the method, but also to gain some baseline estimates of its relative performance. In Chapter 3 I presented the Fisher Information metric as it is used in detecting ecological regime shfits, and discuss the situations under which it may or may not be a good metric.

Velocity (v)

In Chapter 5, I describe a new method, **velocity**, v, as a potential dimension reduction and regime detection method. First introduced by Fath et al. (2003) as one of multiple steps in calculating their variant of Fisher Information, velocity calculates the cumulative sum of the square root of the sum of the squared change in all state variables over a period of time (Eq. (6.1)). Steps for calculating this metric are described in detail in Chapters 3 and 5.

$$\Delta s_i = \sqrt{\sum_{j=1}^{n}(x_{i,j} - x_{i-1,j})^2} s_k = \sum_{i=2}^{k} \Delta s_i 2 \leq k \leq nv = \frac{\Delta s}{\Delta t} \tag{6.1}$$

Variance Index (VI)

The Variance Index was introduced by Brock & Carpenter (2006), and is simply defined as the maximum eigenvalue of the covariance matrix of the system over some period (window) of time. The Variance Index (also called Variance Indicator) was originally applied to a modelled system (Brock & Carpenter, 2006), and has since been applied to empirical data (Spanbauer et al., 2014; Sundstrom et al., 2017). Although rising variance has been shown to manifest prior to abrupt shifts in some empirical systems data (Brock & Carpenter, 2006; Nes & Scheffer, 2005), the Variance Index, which is intended for multivariate data, appears most useful when the system exhibits a discontinuous (non-linear) shift (Brock & Carpenter, 2006).

Fisher Information (FI)

Fisher Information (I) is essentially the area under the curve of the acceleration to the fourth degree (s''^4) divided by the squared velocity (s'^2; also referred to as v in Chapter 5) of the distance travelled by the system, s over some period of time (T), and is given in Eq. (6.2):

$$I = \frac{1}{T}\int_0^T dt \left[\frac{s''^2}{s'^4}\right]^2 \tag{6.2}$$

I refer the reader to Chapter 3 for a complete description and to Cabezas & Fath (2002) for a complete derivation of Fisher Information.

Using moving window analysis to calculate Fisher Information and Variance Index

Unlike *velocity*, the Variance Index and Fisher Information are calculated using moving window analysis. That is, over the entire time series, T^*, these metrics are calculated within multiple windows of time, T. In this approach, all state variables, x_i, are used to inform the calculations (of Variance Index and Fisher Information) over a time interval, T, where T is the length in [time] units of the time interval and satisfies the following condition: $2 \leq T < (T^* - 1)$. If $T = T^* - 1$, then only a single value of the metric will be calculated for entire time series, which does not allow for any estimate of change.

When using these metrics in the context of identifying abrupt changes in ecological systems data across $T*$, it is ideal the value of T meets the following conditions: $3 < T \ll T^* - 1$. The length of a time window dictates the number of calculations one can obtain over T^*, such that the number of potential metric calulations increases as $\frac{T}{T^*}$ decreases. Previous applications of moving window analyses to calculate Fisher Information found that at least eight observations (time points) should be used [citation].

An additional parameter is required when conducting moving window analyses: the number of time points by which the window advances. In order to maximize the data, I advance the window at a rate of one time unit. However, it is important to note that because these data are not sampled annually and the because the window always advances by a single time unit, the number of observations included in each calculation will not be the same. If fewer than 5 observations are in a window, I did not calculate metrics, advancing the window forward. I assigned the calcuated values of Fisher Information and Variance Index within each moving window to the **end** (the last time unit) of the moving window. In temporal analyses, assigning the value[which value] to any other point in time (e.g., the beginning or the middle) muddles the interpretation of the metric over T^*. Also note that this method has the potential to result in calculating a metric for all integers between $0.20T^*$ and T^*.

6.2.3 Simulating data quality and quantity issues using resampling techniques

Using a resampling approach I calculated the regime detection measures over different scenarios simulating data quality and data quantity issues common to ecological data anlaysis. The scenarios are categorized as *observations* and *species*. The observations scenario simulates a loss of temporal observations (decreasing the number of times the system was observed), and the species scenario simulates a loss of information about the system by removing some proportion of the species. The loss of temporal observations and the loss of species were examined at three proportions: $\mathbf{P} = [0.25, 0.50, 0.75, 1.00]$, where $\mathbf{P}$ is the proportion of species and time points **retained** for analysis. For example, when $\mathbf{P} = 0.25$, a random selection of 25% of the species are retained for analysis in the species scenario. I resampled the data over $10,000$ iterations (N_{samp}) for each scenario and $\mathbf{P}$ combination. Note that because when $\mathbf{P} = 1.00$, all data are retained. Therefore, no resampling was conducted at this level because only a single metric (e.g. Velocity) value is possible.

6.2.4 Comparing regime detection measures

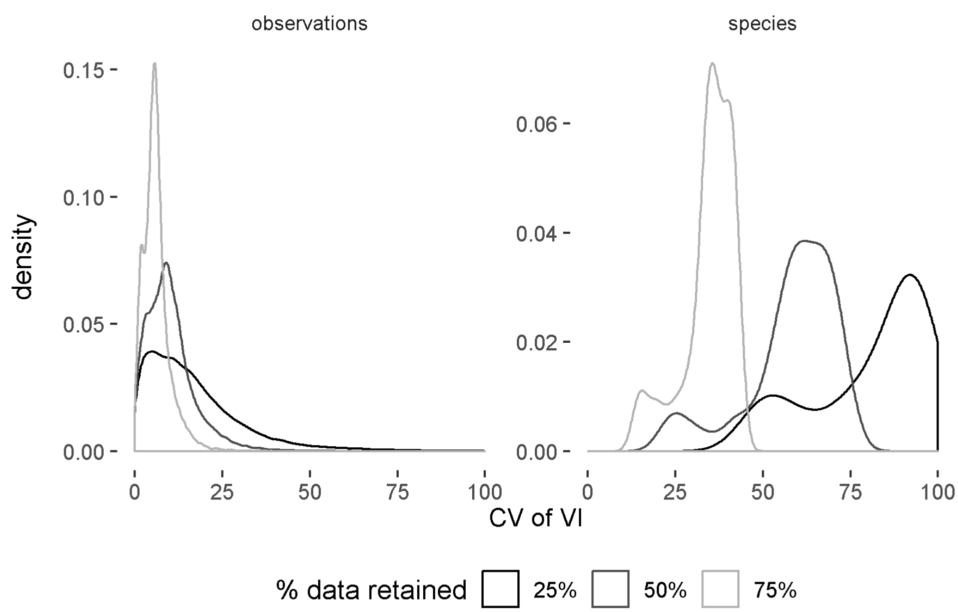
Interpretation of the regime detection measures used in this analysis are currently limited to visual inspection. Therefore, I limit inference in this study largely to the impact of data loss on the variability with a regime detection measure (i.e. how robust is the measure to data loss). It is important to not only identify the influence of data quality and quantity on the performance of individual regime detection metrics, but also to somehow relate these qualities. I visually inspect the relatve performance of these metrics by comparing the coefficient of variation of the resampled samples for the results of resampling method (**M**; species, observations) and sampling percentage (**P**; 25%, 50%, 75%) combination for each metric (FI, VI, v). The coefficient of variation measures provides a relative measure of the variability in the estimated metric across resampled samples as $100\frac{\sigma}{\mu}$, where σ is the standard deviation and μ is the mean value.

I observed the distributions of the CV [get rid of the error to mean ratio, confuses the issue] to identify potential flaws in the metrics should data quality or quantity (**M**, **P**) decrease. First, within a value of **P** a low error to mean ratio (CV) indicates that the metric value is similar across the resampled samples ($N_{samp} = 10,000$). The efficacy of the metric should be questioned as CV$\rightarrow 1$, and perahps even abandoned as CV$\gg 1$. Next, we can examine how the distribution of CV changes within **M** and across **P**. As we increase **P**, we are increasing the volume of data we are feeding the metric. Intuitively, we can assume that as we add more data (volume), we are supplying the metric with more *information*, theoretically increasing the signal-to-noise ratio. Following this logic, we should expect the distribution of CV to generally decrease in mean CV value and also become less variable (less dispersion around the mean CV). A visual examination of the distribution of CV across **P** and within **M** was suffcient to achieve inference regarding the quality of these metrics upon data loss and lessened quality.

6.3 Results

6.3.1 Velocity of the distance travelled (v)

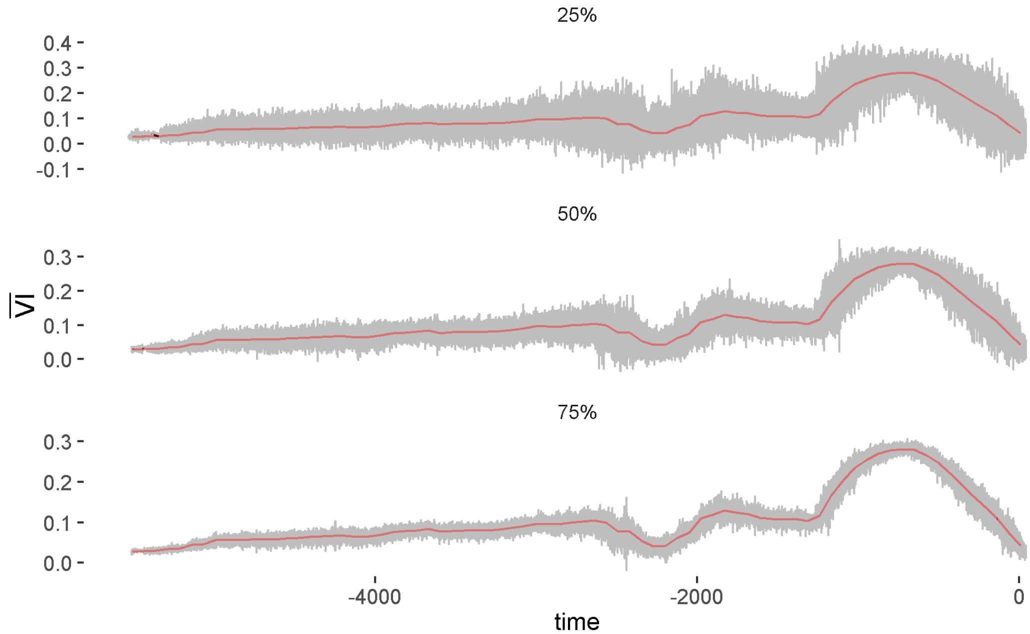
The velocity of the distance travelled, $\frac{ds}{dt}$ or v, exhibited dispersion across the values of **P**, however, yielded consistent results (i.e., high overlap in the densities of the CV across values of **P** and across methodologies; see Fig. **??**). Further, it should be noted that because v is calculated using first differences, it will be sensitive to large changes in the state variables. By examining the density plot of the CV of the dsitance travelled, s, we notice that this measure is highly *insensitive* to data loss (Fig. **??**), suggesting that a finite differencing appraoch (e.g., using total variation regularized differentiation; see Chapter) which can yield a much smoother derivatie than the approach used here, may decrease the sensitivity of v to data loss. This hypothesis is further supported when examining the effect of species (Fig. **??**) and temporal observation loss (Fig. **??**) on the velocity metric. These conditions are representative of the other $\mathbf{P} - \mathbf{M}$ combinations.



\begin{figure}

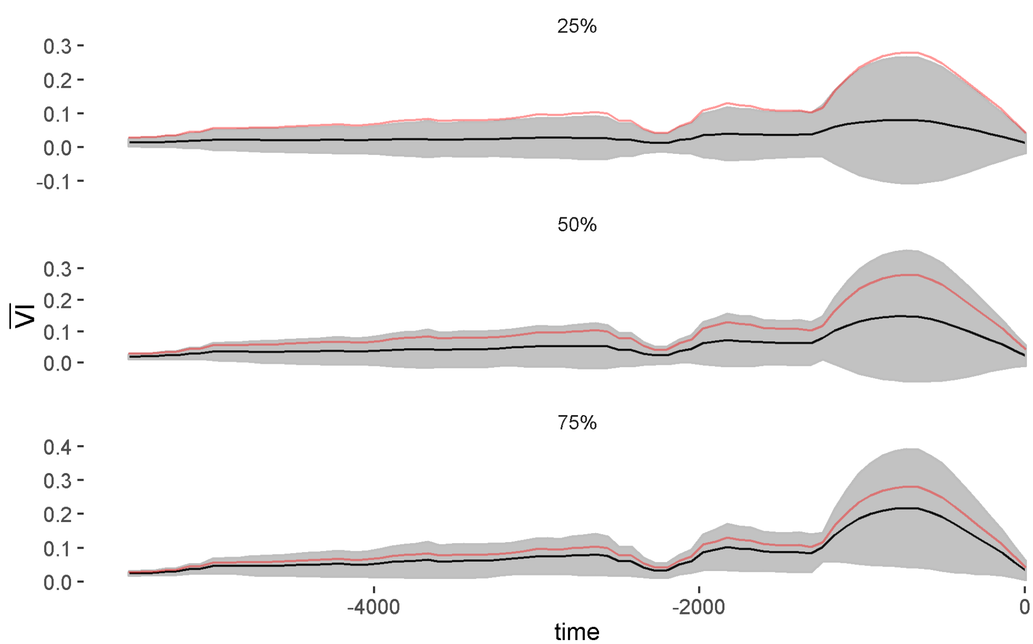
\caption{Density plot of the coefficient of variation (CV) as a percentage (%) of the

Variance Index resampled values over 10,000 iterations. Densities are drawn based on all values of CV but values >100% are not printed.} \end{figure}



\begin{figure}

\caption{Mean Variance Index (VI) and associated 95% confidence intervals over 10,000 iterations using the observations resampling method. Red line indicates the value of VI when **M** and **P** = 100%.} \end{figure}



\begin{figure}

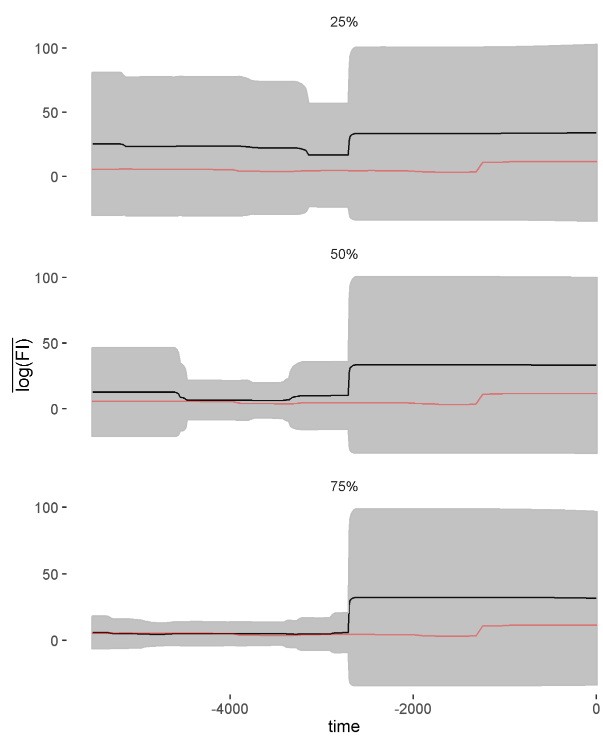
\caption{Mean Variance Index (VI) and associated 95% confidence intervals over 10,000 iterations using the species resampling method. Red line indicates the value of VI when **M** and **P** = 100%.} \end{figure}

6.3.2 Variance Index

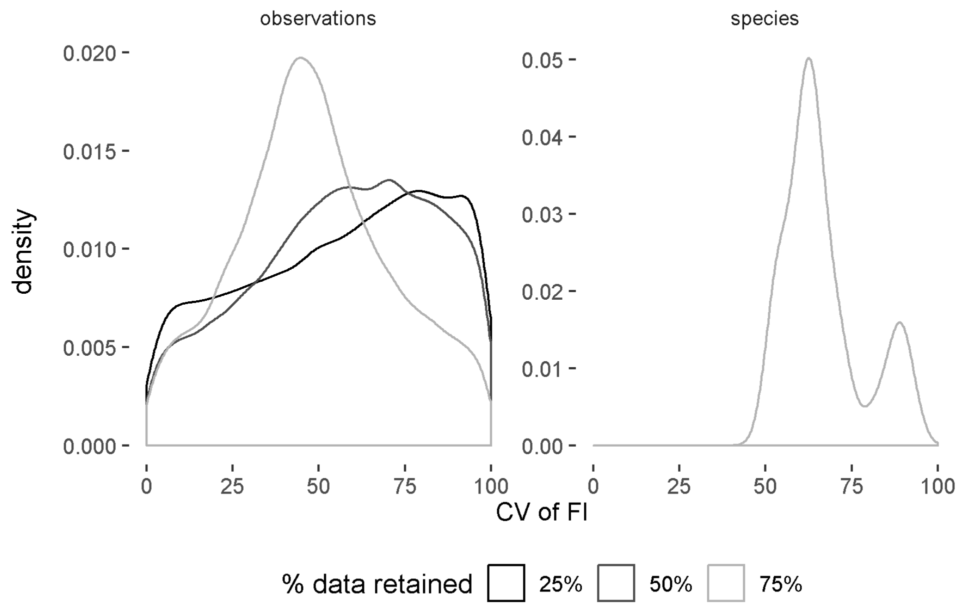
The Variance Index (VI) performed best under the the observations resampling method, exhibiting low values for and low dispersion in the CV density (Fig. **??**) across iterations. However, the VI appears sensitive to high losses of species information, where the density of the CV still exhibits low dispersion but with higher overall mean values (Fig. **??**). Surprisingly, the Variance Index was insensitive to temporal observation loss (Fig. 6.3.1), exhibiting a similar amount of noise across various degrees of data loss (**P**). Although the signal was dampened under the species method, the signals for the shifts in community composition were not lost across levels of **P** (Fig. 6.3.1). This is likey due to the high probably that the dominant species were rarely always excluded from the resampled observations.

6.3.3 Fisher Information is highly sensitive to information loss

The Fisher Information method did not yield conclusive results regarding the abrupt shifts in the paleodiatom community composition. Further, this method appears highly sensitive to varying quality and quantitaties of data (Figs. 6.3.3, 6.3.3). Although the Fisher Information identifies the shift in community composition at $\sim 1,300$ years before present, it fails to identify shifts outside this period. Further, it is difficult to visually analyze any value of the Fisher Information on the original scale as the values range from ≈ 0 to $10^{1}5$ (Fig. 6.3.3). In addition to failing to identify the shifts in community composition, the standard deviation of Fisher Information far exceeded the mean value of Fisher Information under all $\mathbf{M}-\mathbf{P}$ scenarios (Fig. 6.3.3). When I resampled the data using 25% and 50% of the species the ratio of mean Fisher Information to standard deviation (CV) of Fisher Information is always $\gg 1$ (i.e, not pictured in Fig. 6.3.3). The high variation in FI values across resampled iterations coupled with the high dispersion within each $\mathbf{M}-\mathbf{P}$ combination (Fig. 6.3.3) suggests Fisher Information will not produce similar trends when we lose or distort the data collected. This is also sugggested by the high confidence intervals surrounding each $\mathbf{M}-\mathbf{P}$ combination (Fig. 6.3.3).



\begin{figure} \caption{Mean Fisher Information (FI; note the scale) and associated 95% confidence intervals over 10,000 iterations using the species resampling method. Red line indicates the value of FI when $\mathbf{M}$ and $\mathbf{P}$ = 100%. A very small value was added to the mean FI prior to log transformation.} \end{figure}



\begin{figure}

\caption{Density plot of the coefficient of variation (CV) as a percentage (%) of the Fisher Information resampled samples (10,000 iterations). Densities are drawn based on all values of CV, but values >100% are not printed.} \end{figure}

6.4 Detrending the Data Prior to Calculations

If and how to manipulate the original data prior to calculating various regime detection methods is an important consideration, and a line of research that has not yet been fully explored. Although most of the multivariate methods identified in the literature review do not require data conforms to a specific distribution, how th results of these methods can vary as we change the quality and characteristics of the original data (Michener & Jones, 2012). In fact, since many of the methods for regime shift detection are specifically looking for changes in variance structure and autocorrelation, standardizing variances is not counterintuitive. Some studies detrend the original time series prior to data aggregation and calculation of regime detection metrics. I did not detrend the original data for two reasons. First, the authors of the original paper

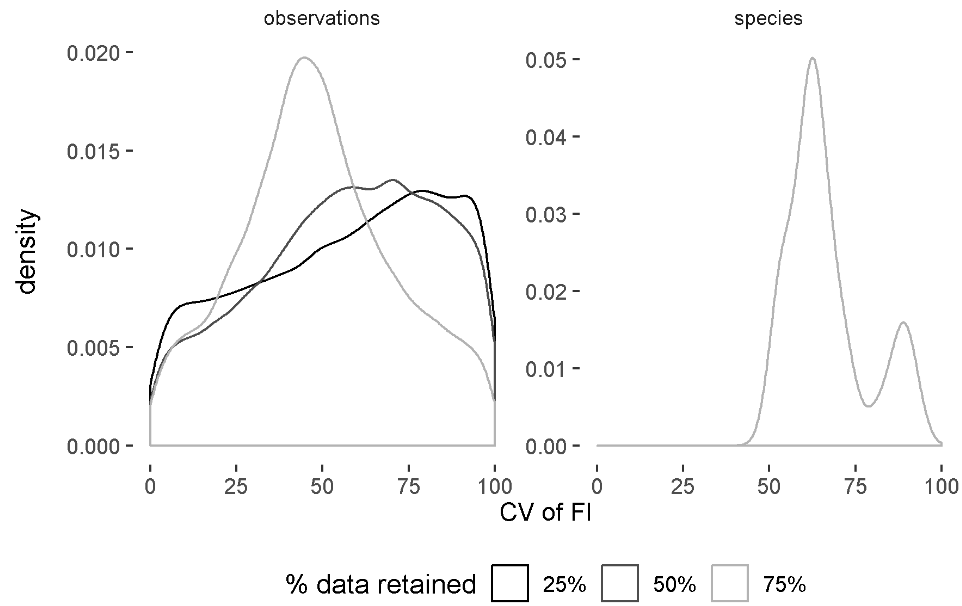


Figure 6.3: Local regression (loess) smoothing of a dominant species in the paleodiatom community, extitAnomoeoneis costata varies with the span parameter, making it difficult to justify smoothing the data prior to calculating various regime detection metrics.

analysing this dataset (Spanbauer et al., 2014) did not detrend species time series. Like Spanbauer et al. (2014) I only scaled the original data, rather than detrending. Second, detrending a time series requires yet another subjective decision by the data analyst. For example, a "spanning" parameter must be chosen when detrending (smoothing) non-linear time series using local regression (Loess) regression (see Fig. 6.3). Other smoothing methods are being explored for both detrending (e.g., PcR; Beck et al., 2018) and regime shift identification (e.g., generalized additive modelling; Beck et al., 2018). Finally, this data exhibits rapid and drastic shifts in community composition *and* contains a disproprtionate amount of dominant versus non-dominant species. Consequently, most species contain more zero than non-zero observations, which makes loess smoothing difficult. Although this chapter concerns impacts of data quality and quantity based on hypothetical data collection and analytical decisions, adding yet another parameter necessitates another layer of compartive analysis. Future work studying the impact of detrending, data scaling, outlier removal, and other related decisions would be of value in understanding the efficacy of these and other regime detection measures in real-world situations.

6.5 Conclusion

In this chapter I provide additional evidence for the sensitivity of select regime detection measures to information (data) quality and quantity loss. The loss of data quantity was simulated by randomly sampling subsets of both the species and the temporal observations, and the reduction in data quality manifests as a function of removing whole species from the community profile. Previous studies of the robustness of univariate regime detection metrics have found similar results, suggesting the measures fail in numerous real-world ecological conditions (Andersen et al., 2009; Contamin & Ellison, 2009). This chapter also highlights the relative insensitivity of the new velocity metric (see Chapters 3, 5) to data and information quality and quantity (e.g., Fig. **??**) loss.

6.6 Ackowledgements

This study was conceptualized at the International Institute for Applied Systems Analysis (IIASA) as part of the Young Scholars Summer Program in 2018. I thank my IIASA program supervisors, Drs. Brian Fath and Elena Rovenskaya, for their advisement of this project and IIASA scientists Drs. Matthias Jonas, Chai Molina, Piotr Zebrowski for additional support.

Chapter 7

Discontinuity chapter under construction

7.1 Introduction

Body size influences the frequency and intensity of inter- and intraspecific competition for resources, territory, and mates, thereby dictating the spatial and temporal scales at which a species of a distinct body size operates (Allen et al., 2006; Peters & Wassenberg, 1983; Silva & Downing, 1995). The scaling structure of terrestrial communities are have been found to have 'lumpy' distributions — that is, they are not well-described using paramteric statistical descriptions. Thes scaling structures as manifested in body mass distirbutions of communities are considered reflective of the discontinuous and heterogeneous nature of resource use. Specifically, Holling (1992) suggests that the body mass distribution of a community or group of species reflects the discontinuous nature of environmental structures and processes. Quantitative analyses of animal body sizes (Allen et al., 2006; Nash et al., 2014) and other similar distributions has revealed the prevalence of discontinuous size distributions in faunal body mass

(Skillen & Maurer, 2008), plant biomass (Spanbauer et al., 2016), city population sizes (Garmestani, Allen, & Bessey, 2005), and species<U+2019> home range size (Restrepo & Arango, 2008).

Species operating at similar spatial and temporal scales are those close in body size, and are identified in discontinuity analyses as belonging to the same body mass aggregations or <U+2018>lumps<U+2019> (Allen and others 1999). The interactions among species in a single aggregation presumably experience a higher frequency and intensity of interspecific interactions with each other as opposed to those in different aggregations (Peterson and others 1998).

barichievy2018method - method taht caleb used.

Multiple global change drivers are exerting influence in a south-to-north pattern within the Great Plains. For instance, in the Great Plains, climate change is shifting native and agricultural plant phenologies (Richardson et al., 2013) and geographic centers of species distributions (Hovick et al., 2016). Woody plant encroachment is causing regime shifts from historically grassland regimes to woodland or shrubland regimes (Engle et al., 2008); whole ecoregions in the southern Great Plains have shifted to woodlands in the past century, and northern ecoregions are increasingly on the brink of wholesale regime shifts (Twidwell et al., 2013). Interacting with climate change and woody plant encroachment, fire frequency and size have also increased by >400% in the Great Plains, especially in the southern portions that have transitioned to woodlands (Donovan et al., 2017). Energy development such as oil and gas extraction reduced net primary productivity by approximately 4.5 Tg between 2000-2015, with much of the development focused on the southern Great Plains (Allred et al., 2015). Although the rate of agricultural land conversion had greatly slowed by the 1950s (Brown et al., 2005), the northern plains lost much of its remaining grassland after commodity prices surged at the beginning of the 21st century (Drummond et al., 2012). Urbanization and population growth in the Great Plains has continually increased in and around already populated areas (Brown et al., 2005), with the greatest growth occurring in the southern portions of the Great Plains. In light of this, we selected a belt transect on the ecotone of the Great Plains and Eastern Temperate Forests extending from the Gulf of Mexico to the edge of the boreal forest in Canada. Specifically, the belt transect extended south-north from 28 - 49 degrees latitude (approximately 2300 km) and east-west from 93 - 97 degrees longitude (approximately 350 km). Statistical Analysis Identifying discontinuities For each route falling within the belt transect, we identified discontinuities in avian community body masses by rank-ordering the log-transformed body masses of each species observed at each route for each year. We obtained mean body mass estimates for all species in the analysis from the CRC Handbook of Avian Body Masses (Dunning Jr, 2007). We then used the <U+201C>discontinuity detector<U+201D> method (Barichievy et al., 2018) on the log-ranked body masses, which is based on the Gap Rarity Index for detecting discontinuities in continuous data (C. Stow, Allen, & Garmestani, 2007). For taxa with determinant growth, mean body mass has been shown to reliably differentiate size aggregations and is strongly allometric to the scales at which functions are carried out by organisms (Nash et al., 2014; Sundstrom & Allen, 2014). Because the discontinuity detector method is known to overestimate discontinuities in observations with low species richness, we removed any routes with $<$ 40 species observed within a given year (Table 3.1). We used a power table (Lipsey, 1990) to account for sample size (the number of species observed at each BBS route in a given year) and average variance in body masses (Dunning Jr, 2007) to adjust the critical d-value (the value based on Monte Carlo simulations that identifies significant discontinuities) where N varied (Allen, Forys, & Holling, 1999) (Table 3.2).

7.2 Methods

I used discontinuity analysis to identify the cross-scale structure of the avian communities across space-time (Barichievy et al., 2018).

7.2.1 Spatial sampling scheme

I sampled BBS routes from a North-South-running transect in Central North America running from the Southern Great Plains ecoregion near the Gulf of Mexico ($\sim 28°$ latitude) to boreal Canadaian forest ($\sim 49°$ latitude), running $\sim 5°$ west from ($\sim 93°$ longitude) to ($\sim 97°$ longitude). This region of the world is experience widespread range expansion of a native invasive woody plant, Eastern Redcedar [*Juniperus `virginiana*;@twidwell2016plant`;Donovan et al. (2018)]. This widespread and steadfast transition of native and replanted short and tallgrass prairie habitat to semi-wooded areas should manifest in the avian community composition across large spatial scales, potentially manifesting as changes in cross-scale strucutre of avian body masses. Given the body of evidence suggesting avian community composition shifts upon woody encroachment, we can make two predictions. First, an area undergoing woody encroachment should exhibit a shift in community composition, given the extreme sensitivity of grassland bird species to woody cover. As such, we should next expect a shift in the body mass aggregation structure of the avian community. However, this compositional shift need not manifest in the discontinuous strucutre of the body mass distribution (e.g., shifting the number, size, and/or location of body mass aggregations within a distribution). Rather, at a minimum, we should expect species turnover and potentially shifts in the species which dominate (comprise the center of) or are peripherally associated with (comprise the edges of) each body mass aggregation.

7.2.2

Given the extent and resolution of our data, we are able to only test the biotic interaction and textural discontinuity hypotheses (Allen et al., 2006). “If scale<U+2010>dependent resource variability is introduced into the model, then a single mode can separate into multiple modes (Marquet et al. 1995), indicating an interaction between the distribution of resources in the landscape and body mass aggregations.” from allen2006patterns

Chapter 8

Conclusions

$$\begin{aligned} Data &= Information \\ &= Signal \\ &= Process + Noise \end{aligned} \tag{8.1}$$

Climate change is expected to induce an increase in both the intensity and frequency of rapid ecological change or disturbance, impacting social systems, potentially to the detriment of human communities most vulnerable. Identifying and forecasting these changes is critical for community and ecological planning, management, and disaster mitigation. Because ecological and social systems are tightly coupled, it is commonplace to use ecological indicators to identify change and potential changes that may impact these systems. Many papers introducing or discussing regime detection measures suggest the ecologist uses multiple lines of evidence, ranging from historical observations to ecological modelling results, for identifying an ecological regime shift (Lindegren et al., 2012). Although valid, comparing results of multiple methods or lines of evidence within a single system has yielded inconsistent results, and inconsistent results can result in either improper conclusions, or in what I am calling **method mining**. That is, a dataset is analyzed using until a sufficient number of methods yield affirmative results.

8.1 Method mining regime detection methods

Many regime detection measures have yet to be properly and statistically (or numerically) scrutinized. However, it should be noted that, in part due to both (i) the popularity and (ii) the sheer number of 'new' methods a handful of authors[1]. Ecological indicators (a.k.a. indices, metrics) have been suggested as 'early-warning indicators' of ecological regime shifts or abrupt change (Chapters 1 and 2) and are methods of measurement designed to provide inference about one or more unobserved or latent processes, are inherently biased. Regardless of the state of the theory supporting *regime shifts* in ecology, ecological indicators and the methods for calculating them should be heavily scrutinized prior to being used in an ecological management or policy-making setting. Rather, new methods (indices, metrics, etc.) are being introduced into the literature at a rate exceeding that at which they are scrutinized (Chapter 2). This dissertation demonstrates that, while potentially useful, regime detection metrics are inconsistent, not generalizable, and are currently not validated using probabilities or other statistical measurements of certainty.

8.2 Ecological data are noisy

Regime detection metrics appear more reliable when the signal-to-noise ratio is high (Ch. 2, Ch. 5, Taranu et al., 2018). Ecological systems are noisy, and the observational data we are collecting at large scales (e.g., the North American Breeding Bird survey),

[1] S.R. Carpenter is one example of an author who has relative infamy in the field and has, as primary author or otherwise, introduced a relatively large number of new methods (e.g., rising variance, the variance index, Fourier transform, online dynamic linear modelling, TVARSS, to name a few)

is noisy. Using methods incapable of identifying meaningful signals in noisy data appears futile, yet, methods for doing so are increasingly introduced in the scientific literature (Ch. 2).

8.3 Data collection and munging biases and limits findings

Regime detection measures and other ecological indicators can signal (see (8.1)) various changes in the data, however, understanding what processes are embedded in the signals (i.e., removing the noise) requires expert judgement. And because a consequence of data collection and data analysis limits the extent to which we can identify and infer processes and change within an ecological system, **I suggest the practical ecologist scrutinizes her data prior to identifying and conducting analyses**, including those that are purely exploratory. By collecting and analysing data, the ecologist has defined the bounaries of the system *a priori*ˆ+(+ Beisner, Haydon, & Cuddington, 2003 states this eloquently as, "The number and choice of variables selected to characterize the community will be determined by what we wish to learn from the model"). The influence of state variable selection is ignored by some metrics (e.g. Fisher Information Eason et al., 2014b and v Chapter 5), in that the resulting measure is composite and carries no information regarding the influence of state variables on the metric result.

The actual limitations to the system should be, theoretically, known as a result of bounding the system. Inference beyond this system is extrapolation, and should be treated as speculation, especially when not accompanied by a measure of uncertainty around one's predictions.

8.4 Common Limitations of Regime Detection Measures

Limitations of the findings in this dissertation and of the regime detection methods used herein are largely influenced by the **data collection**, **data munging** processes. Although the below mentioned points may seem logical to many, these assumptions are overlooked by many composite indicators, including regime detection measures.

1. Signals in the indicators are restricted to the ecological processes captured by the input data. Extrapolation occurs when processes manifest at scales different than the data collected [resolution; Chapter 4]

2. Normalization and weighting techniques often impact results (whether ecological or numerical) (Appendices 8.7 and 8.5)

3. Data aggregation techniques often impact results (Chapter 6)

4. Some indices fail to generalize across systems or taxa (see Chapters 1 and 2)

8.5 Specific synthesis of chapter results

pendix {-}"

Appendix A: R package regimeDetectionMeasures

This appendix contains example analysis associated with the R Package, `regimeDetectionMeasures`. Development source code for this package is available on GitHub as a compressed file at `https://github.com/TrashBirdEcology/regimeDetectionMeasures/archive/master.zip` or at `https://github.com/TrashBirdEcology/regimeDetectionMeasures`.

8.6 Measures/metrics calculated

This package will calculate a various regime detection metrics that have been used to 'detect ecological regime shifts'. A 'new' metric, **distance travelled** is also calculated (Burnett and others, *in prep*). **Composite measures:** 1. Distance travelled -see also package `distanceTravelled`.

1. Fisher Information 1. Variance Index

Single-variable measures: 1. Skewness (mean and mode versions)

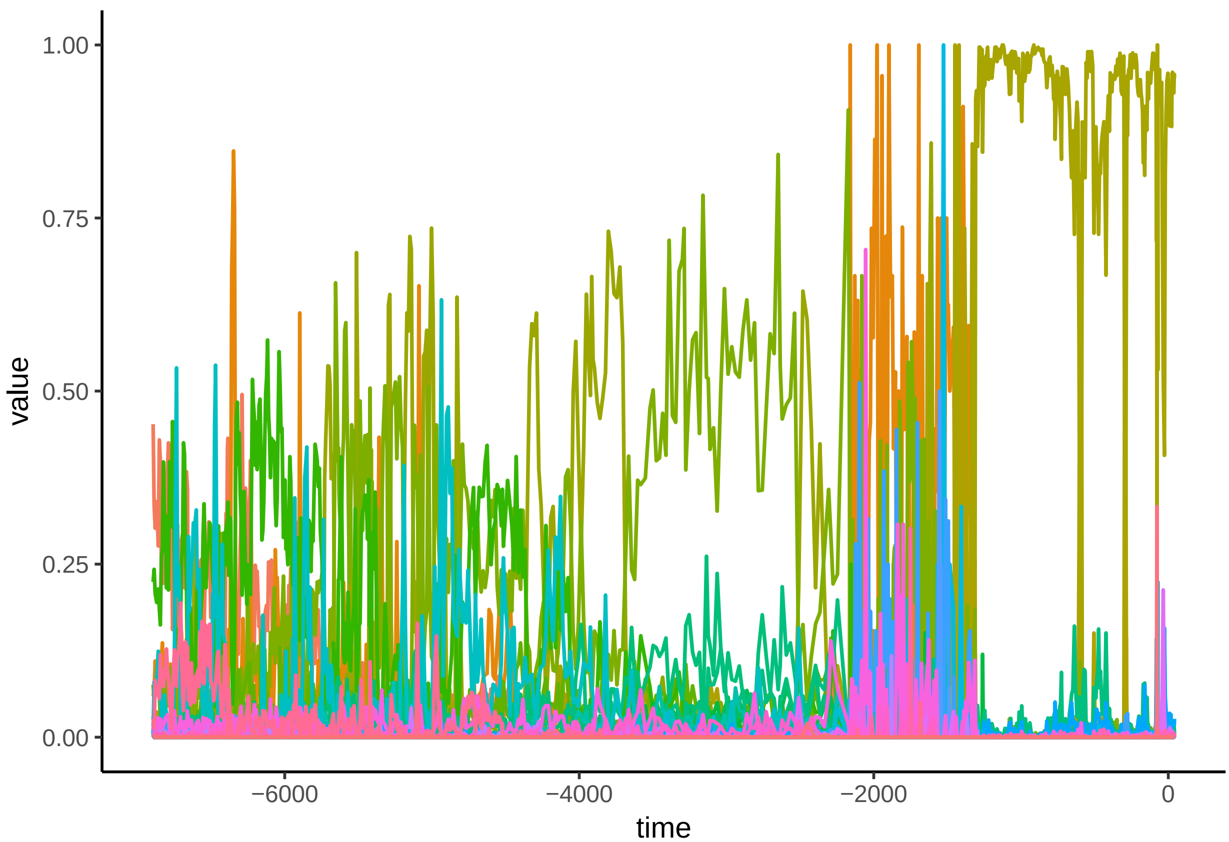
1. Kurtosis

1. Variance

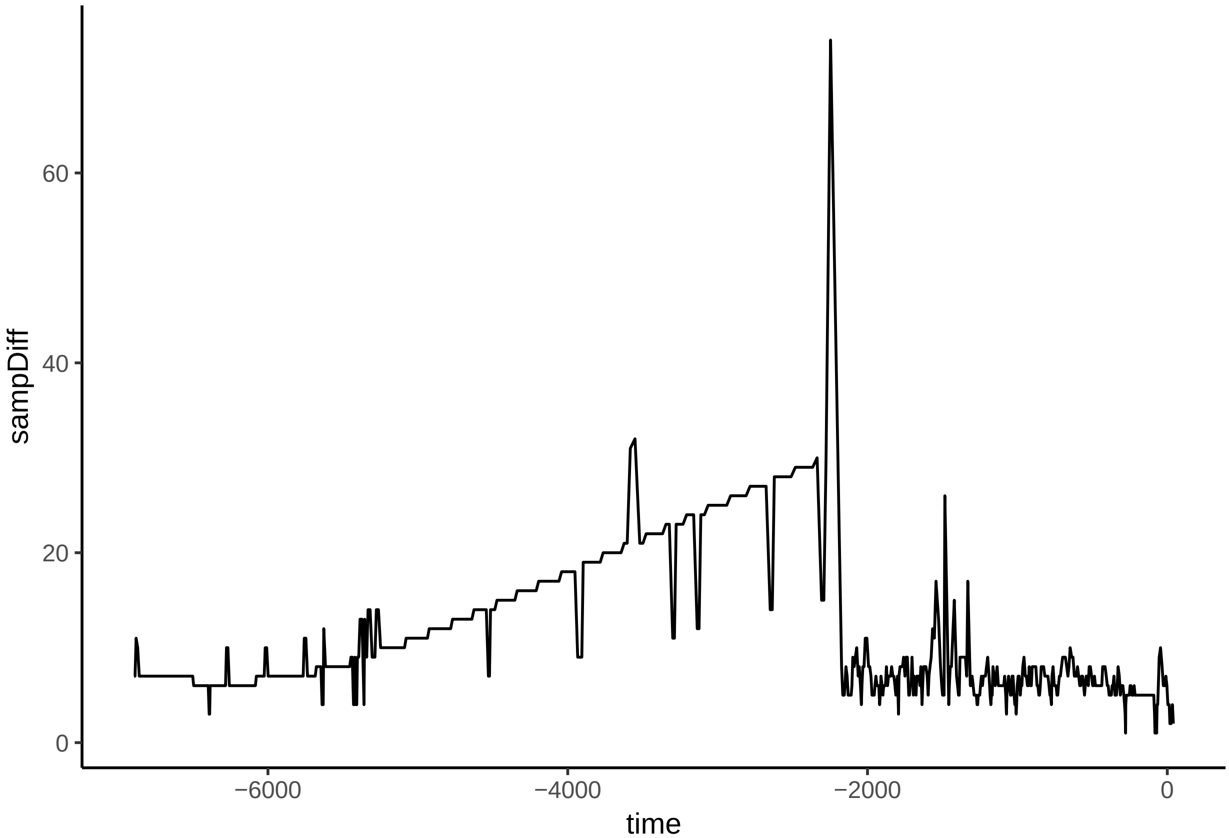
1. Mean

1. Mode
1. Coefficient of variation, CV
1. Autocorrelation lag-1 (using `stats::acf`)

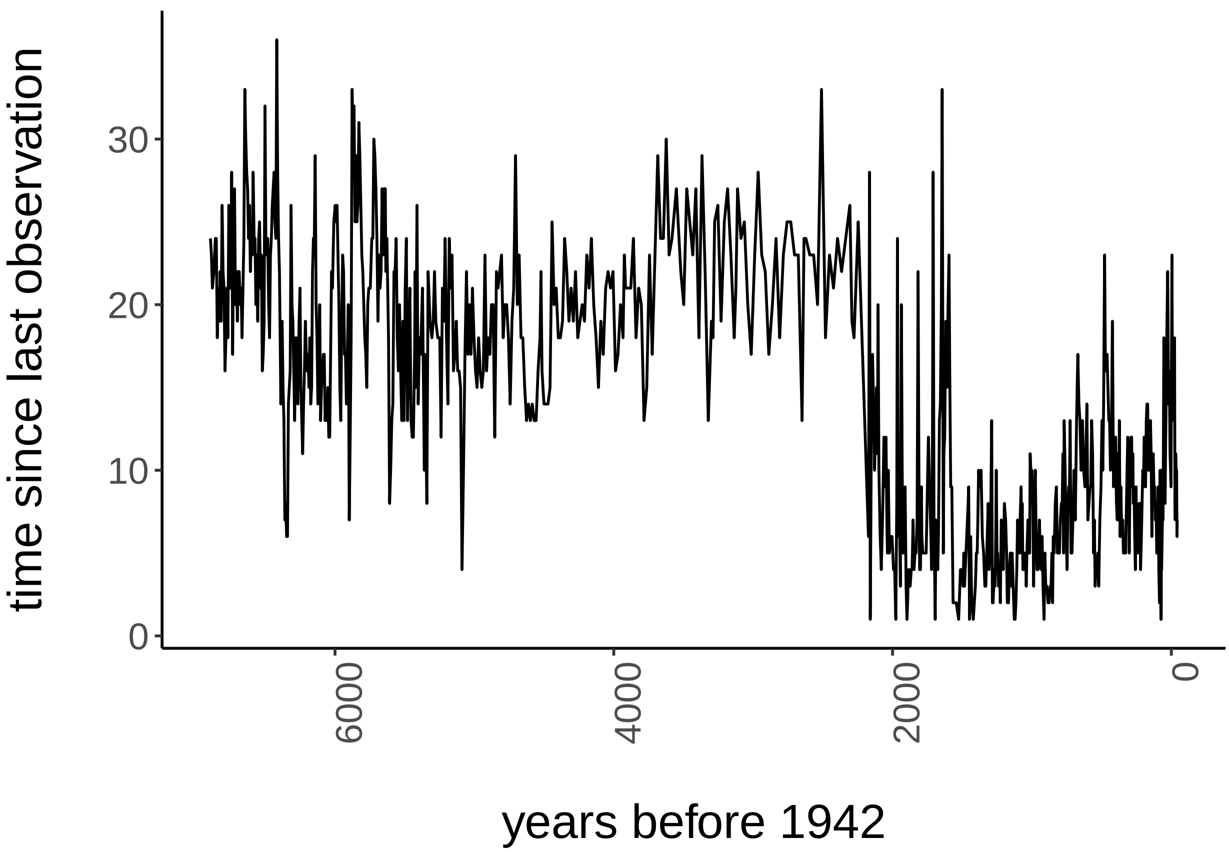
8.7 Example analysis




```
Warning: Removed 1 rows containing missing values (geom_path).
```


sampDiff
60
40
20
0
−6000
−4000
−2000
0
time

2485


time since last observation
30
20
10
0
6000
4000
2000
0
years before 1942

2486

```
# Calculate FI, VI, and early warning signals ----------------------------------

# Uses a moving window analysis to calculate FI and Vi within each window

results <-

    rdm_window_analysis(

        origData,

        winMove = 0.25,

        overrideSiteErr = F,

        min.window.dat = 2,

        fi.equation = "7.12",

        to.calc = c('EWS', 'VI')

    )


# Reults will return all results in a single data frame.

head(results)
```

```
        metricValue cellID_min cellID_max winStart   winStop metricType

1 0.157575450026098          1          1 -1746.63  -12.6975         VI

2 0.186533784869325          1          1 -1820.16  -86.2275         VI

3 0.211304354488511          1          1 -1893.69 -159.7575         VI

4 0.235451240210219          1          1 -1967.22 -233.2875         VI

5  0.25306928580876          1          1 -2040.75 -306.8175         VI

6 0.260835104190261          1          1 -2114.28 -380.3475         VI

  variable cellID

1       NA     NA

2       NA     NA

3       NA     NA

4       NA     NA
```

```
        NA      NA
        NA      NA
```

Appendix B: R package bbsRDM

This appendix contains a vignette associated with the R Package, `bbsRDM`. Development source code for this package is available on GitHub as a compressed file, `https://github.com/TrashBirdEcology/bbsRDM/archive/master.zip` or at `https://github.com/TrashBirdEcology/rRDM`.

This vignette runs through the capabilites of the bbsRDM package, which relies on the package `trashbirdecology::regimeDetectionMeasures`. Although this package can be used to calculate and visualize BBS data using time series, the example at hand runs presents an application to spatial transects.

.1 Load packages & create local directories

There are a lot of dependencies to load.

Create directories to locally store BBS data and results.

```
#> Warning in dir.create(resultsDir): './chapterFiles/appendix_bbsRDM/
#> myResults' already exists
#> Warning in dir.create(resultsDirEWS): './chapterFiles/appendix_bbsRDM/
```

```
#> myResults/ews' already exists
#> Warning in dir.create(resultsDirDIST): './chapterFiles/appendix_bbsRDM/
#> myResults/distances' already exists
```

.2 Download the BBS data and save to file locally

If necessary, download all the state data. This takes 10-15 minutes, so only run if you have not recently downloaded the BBS data.

```
#> [1] "NOT DOWNLOADING BBS DATA. If you wish to download the BBS data, please remove file
#> [1] "NOT DOWNLOADING BBS DATA. If you wish to download the BBS data, please remove file
```

.3 Create the sampling grid

Next we build the sampling grid to force route information onto a regular grid.

Now we load in the BBS data from the feathers we created and align with the sampling grid. This requires a bit of memory, proceed with caution.

.4 Next, we can subset the BBS data by species and/or functional traits (OPTIONAL but recommended)

Although subsetting the speices is optional, I recommend removing waterfowl, wading birds, and shorebirds from analyses, especially as the spatial extent of the analysis increases.

.5 Subset species according to AOU species codes (i.e. by family, genera, etc..)

For this example we will remove shorebirds, wading birds, and waterfowl (i.e., AOU species' codes 0000:2880). *See `R/subsetByAOU.R` source code or documentation for options (see: `subset.by`)

.6 Subset species by trait, body mass, taxonomically, etc…

.7 Calculate regime detection metrics across space or time

.7.1 First, define the parameters required to calcualte the metrics.

.8 Define the years we want to analyze

For this (spatial) example, we will analyze only every fifth year

.9 Conduct analysis!

This section will loop through `years.use` and `dir.use`, running each BBS route (temooral analysis) or spatial transect by year (spatial analysis) at a time. Results are saved in directories created in @ref(#createDirs)

.10 Import and munge the results to prepare for visualization

First, import and combine the results as created in @ref(#calcMetrics) This chunk will import the EWS results and the distance results separately, combining each into their own data frames.

```
#> [1] "I am importing 474 files. Does this sound right?!"
#> [1] "I am importing 475 files. Does this sound right?!"
```

Next, get the results to align with our sampling grid for visualizing results across space.

.11 Visualize results: one regime detection metric at a time

First, specify plotting parameters.

We can visualize either the distance results (`distResults`) or the early-warning signal results (`ewsResults`).

Define the results we want to visualize:

Plot individual transects **Note: please specify dirID.ind as desired spatial transect number, and dirInd as direction (E-W or N-S)**

References

Abadi, F., Gimenez, O., Arlettaz, R., & Schaub, M. (2010). An assessment of integrated population models: Bias, accuracy, and violation of the assumption of independence. *Ecology*, *91*(1), 7–14.

Alheit, J., M??llmann, C., Dutz, J., Kornilovs, G., Loewe, P., Mohrholz, V., & Wasmund, N. (2005). Synchronous ecological regime shifts in the central Baltic and the North Sea in the late 1980s. *ICES Journal of Marine Science: Journal Du Conseil*, *62*(7), 1205–1215. doi: 10.1016/j.icesjms.2005.04.024

Allen, C. R., Garmestani, A., Havlicek, T., Marquet, P. A., Peterson, G., Restrepo, C., ... Weeks, B. (2006). Patterns in body mass distributions: Sifting among alternative hypotheses. *Ecology Letters*, *9*(5), 630–643.

Andersen, T., Carstensen, J., Hernandez-Garcia, E., & Duarte, C. M. (2009). Ecological thresholds and regime shifts: Approaches to identification. *Trends in Ecology & Evolution*, *24*(1), 49–57.

Anderson, P. J., & Piatt, J. F. (1999). Community reorganization in the Gulf of Alaska following ocean climate regime shift. *Marine Ecology Progress Series*, 117–123.

Angel, E. (2000). *Interactive computer graphics : A top-down approach with opengl.* Boston, MA: Addison Wesley Longman.

Angel, E. (2001a). *Batch-file computer graphics : A bottom-up approach with quicktime.* Boston, MA: Wesley Addison Longman.

Angel, E. (2001b). *Test second book by angel.* Boston, MA: Wesley Addison Longman.

Aria, M., & Cuccurullo, C. (2017). Bibliometrix: An r-tool for comprehensive science mapping analysis. *Journal of Informetrics*, *11*(4), 959–975. Retrieved from `https://doi.org/10.1016/j.joi.2017.08.007`

Bahlai, C. A., Werf, W. vander, O'Neal, M., Hemerik, L., & Landis, D. A. (2015). Shifts in dynamic regime of an invasive lady beetle are linked to the invasion and insecticidal management of its prey. *Ecological Applications*, *25*(7), 1807–1818.

Barichievy, C., Angeler, D. G., Eason, T., Garmestani, A. S., Nash, K. L., Stow, C. A., ... Allen, C. R. (2018). A method to detect discontinuities in census data. *Ecology and Evolution*, *8*(19), 9614–9623.

Batt, R. D., Carpenter, S. R., Cole, J. J., Pace, M. L., & Johnson, R. A. (2013). Changes in ecosystem resilience detected in automated measures of ecosystem metabolism during a whole-lake manipulation. *Proceedings of the National Academy of Sciences*, *110*(43), 17398–17403.

Beaugrand, G. (2004). The north sea regime shift: Evidence, causes, mechanisms and consequences. *Progress in Oceanography*, *60*(2-4), 245–262.

Beck, K. K., Fletcher, M.-S., Gadd, P. S., Heijnis, H., Saunders, K. M., Simpson, G. L., & Zawadzki, A. (2018). Variance and rate-of-change as early warning signals for a critical transition in an aquatic ecosystem state: A test case from Tasmania, Australia. *Journal of Geophysical Research: Biogeosciences.*

Beisner, B. E., Haydon, D. T., & Cuddington, K. (2003). Alternative stable states in ecology. *Frontiers in Ecology and the Environment*, *1*(7), 376–382.

Benedetti-Cecchi, L., Tamburello, L., Maggi, E., & Bulleri, F. (2015). Experimental perturbations modify the performance of early warning indicators of regime shift. *Current Biology*, *25*(14), 1867–1872.

Bennett, E. M., Peterson, G. D., & Gordon, L. J. (2009). Understanding relationships among multiple ecosystem services. *Ecology Letters*, *12*(12), 1394–1404.

Bestelmeyer, B. T., Ellison, A. M., Fraser, W. R., Gorman, K. B., Holbrook, S. J., Laney, C. M., ... others. (2011). Analysis of abrupt transitions in ecological systems. *Ecosphere*, *2*(12), 1–26.

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

Boettiger, C., Ross, N., & Hastings, A. (2013). Early warning signals: The charted and uncharted territories. *Theoretical Ecology*, *6*(3), 255–264.

Brock, W. A., & Carpenter, S. R. (2010). Interacting regime shifts in ecosystems: Implication for early warnings. *Ecological Monographs*, *80*(3), 353–367.

Brock, W., & Carpenter, S. (2006). Variance as a Leading Indicator of Regime Shift in Ecosystem Services. *Ecology and Society*, *11*(2). doi: 10.5751/ES-01777-110209

Burthe, S. J., Henrys, P. A., Mackay, E. B., Spears, B. M., Campbell, R., Carvalho, L., ... others. (2016). Do early warning indicators consistently predict nonlinear change in long-term ecological data? *Journal of Applied Ecology*, *53*(3), 666–676.

Butitta, V. L., Carpenter, S. R., Loken, L. C., Pace, M. L., & Stanley, E. H. (2017). Spatial early warning signals in a lake manipulation. *Ecosphere*, *8*(10), n/a–n/a.

doi: 10.1002/ecs2.1941

Byrski, J., & Byrski, W. (2016). A double window state observer for detection and isolation of abrupt changes in parameters. *International Journal of Applied Mathematics and Computer Science*, *26*(3), 585–602.

Cabezas, H., & Fath, B. D. (2002). Towards a theory of sustainable systems. *Fluid Phase Equilibria*, *194*(Supplement C), 3–14. doi: 10.1016/S0378-3812(01)00677-X

Carpenter, S., & Brock, W. (2011). Early warnings of unknown nonlinear shifts: A nonparametric approach. *Ecology*, *92*(12), 2196–2201.

Carpenter, S. R., & Brock, W. A. (2006). Rising variance: A leading indicator of ecological transition. *Ecology Letters*, *9*(3), 311–318.

Carpenter, S. R., Brock, W. A., Cole, J. J., Kitchell, J. F., & Pace, M. L. (2008). Leading indicators of trophic cascades. *Ecology Letters*, *11*(2), 128–138. doi: 10.1111/j.1461-0248.2007.01131.x

Carpenter, S. R., Cole, J. J., Pace, M. L., Batt, R., Brock, W., Cline, T., … others. (2011). Early warnings of regime shifts: A whole-ecosystem experiment. *Science*, *332*(6033), 1079–1082.

Chartrand, R. (2011). Numerical differentiation of noisy, nonsmooth data. *ISRN Applied Mathematics*, *2011*.

Clements, C. F., & Ozgul, A. (2016). Including trait-based early warning signals helps predict population collapse. *Nature Communications*, *7*. doi: 10.1038/ncomms10984

Clements, C. F., & Ozgul, A. (2018). Indicators of transitions in biological systems. *Ecology Letters*, *21*(6), 905–919.

Cobo, M. J., López-Herrera, A. G., Herrera-Viedma, E., & Herrera, F. (2011). An approach for detecting, quantifying, and visualizing the evolution of a research field: A practical application to the fuzzy sets theory field. *Journal of Informetrics*, *5*(1), 146–166.

Contamin, R., & Ellison, A. M. (2009). Indicators of regime shifts in ecological systems: What do we need to know and when do we need to know it. *Ecological Applications*, *19*(3), 799–816. doi: 10.1890/08-0109.1

Dakos, V., Carpenter, S. R., Brock, W. A., Ellison, A. M., Guttal, V., Ives, A. R., ... Nes, E. H. van. (2012). Methods for detecting early warnings of critical transitions in time series illustrated using simulated ecological data. *PloS One*, *7*(7), e41010.

Dakos, V., Carpenter, S. R., Nes, E. H. van, & Scheffer, M. (2015a). Resilience indicators: Prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1659), 20130263. doi: 10.1098/rstb.2013.0263

Dakos, V., Carpenter, S. R., Nes, E. H. van, & Scheffer, M. (2015b). Resilience indicators: Prospects and limitations for early warnings of regime shifts. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *370*(1659), 20130263.

Dakos, V., Van Nes, E. H., D'Odorico, P., & Scheffer, M. (2012). Robustness of variance and autocorrelation as indicators of critical slowing down. *Ecology*, *93*(2), 264–271.

Davis, E. P., & Karim, D. (2008). Comparing early warning systems for banking crises. *Journal of Financial Stability*, *4*(2), 89–120.

DeAngelis, D. L., & Yurek, S. (2017). Spatially explicit modeling in ecology: A review. *Ecosystems*, *20*(2), 284–300.

deYoung, B., Barange, M., Beaugrand, G., Harris, R., Perry, R. I., Scheffer, M., & Werner, F. (2008). Regime shifts in marine ecosystems: Detection, prediction and management. *Trends in Ecology & Evolution*, *23*(7), 402–409. doi: 10.1016/j.tree.2008.03.008

Donovan, V. M., Burnett, J. L., Bielski, C. H., Birgé, H. E., Bevans, R., Twidwell, D., & Allen, C. R. (2018). Social–ecological landscape patterns predict woody encroachment from native tree plantings in a temperate grassland. *Ecology and Evolution*, *8*(19), 9624–9632.

Ducré-Robitaille, J.-F., Vincent, L. A., & Boulet, G. (2003). Comparison of techniques for detection of discontinuities in temperature series. *International Journal of Climatology: A Journal of the Royal Meteorological Society*, *23*(9), 1087–1101.

Dutta, P. S., Sharma, Y., & Abbott, K. C. (2018). Robustness of early warning signals for catastrophic and non-catastrophic transitions. *Oikos*, *127*(9), 1251–1263.

Eason, T., & Cabezas, H. (2012). Evaluating the sustainability of a regional system using Fisher information in the San Luis Basin, Colorado. *Journal of Environmental Management*, *94*(1), 41–49.

Eason, T., Garmestani, A. S., & Cabezas, H. (2014a). Managing for resilience: Early detection of regime shifts in complex systems. *Clean Technologies and Environmental Policy*, *16*(4), 773–783.

Eason, T., Garmestani, A. S., & Cabezas, H. (2014b). Managing for resilience: Early detection of regime shifts in complex systems. *Clean Technologies and Environmental Policy*, *16*(4), 773–783.

Fath, B. D., & Cabezas, H. (2004). Exergy and Fisher Information as ecological indices. *Ecological Modelling*, *174*(1), 25–35. doi: 10.1016/j.ecolmodel.2003.12.045

Fath, B. D., Cabezas, H., & Pawlowski, C. W. (2003). Regime changes in ecological systems: An information theory approach. *Journal of Theoretical Biology*, *222*(4), 517–530. doi: 10.1016/S0022-5193(03)00067-5

Filatova, T., Polhill, J. G., & Ewijk, S. van. (2016). Regime shifts in coupled socio-environmental systems: Review of modelling challenges and approaches. *Environmental Modelling & Software*, *75*, 333–347.

Fisher, R. A. (1922). On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, *222*, 309–368. Retrieved from `http://www.jstor.org/stable/91208`

Folke, C., Carpenter, S., Walker, B., Scheffer, M., Elmqvist, T., Gunderson, L., & Holling, C. S. (2004). Regime shifts, resilience, and biodiversity in ecosystem management. *Annu. Rev. Ecol. Evol. Syst.*, *35*, 557–581.

Frieden, B. R. (1990). Fisher information, disorder, and the equilibrium distributions of physics. *Physical Review A*, *41*(8), 4265–4276. doi: 10.1103/PhysRevA.41.4265

Frieden, B. R., & Gatenby, R. (Eds.). (2007). *Exploratory Data Analysis Using Fisher Information*. Retrieved from `http://www.springer.com/us/book/9781846285066`

Garmestani, A. S., Allen, C. R., & Bessey, K. M. (2005). Time-series analysis of clusters in city size distributions. *Urban Studies*, *42*(9), 1507–1515.

Graham, R., & Tél, T. (1984). Existence of a potential for dissipative dynamical systems. *Physical Review Letters*, *52*(1), 9.

Groffman, P. M., Baron, J. S., Blett, T., Gold, A. J., Goodman, I., Gunderson, L. H.,

... Wiens, J. (2006). Ecological Thresholds: The Key to Successful Environmental Management or an Important Concept with No Practical Application? *Ecosystems*, *9*(1), 1–13. doi: 10.1007/s10021-003-0142-z

Guttal, V., & Jayaprakash, C. (2009). Spatial variance and spatial skewness: Leading indicators of regime shifts in spatial ecological systems. *Theoretical Ecology*, *2*(1), 3–12.

Guttal, V., Jayaprakash, C., & Tabbaa, O. P. (2013). Robustness of early warning signals of regime shifts in time-delayed ecological models. *Theoretical Ecology*, *6*(3), 271–283.

Hastings, A., & Wysham, D. B. (2010a). Regime shifts in ecological systems can occur with no warning. *Ecology Letters*, *13*(4), 464–472.

Hastings, A., & Wysham, D. B. (2010b). Regime shifts in ecological systems can occur with no warning. *Ecology Letters*, *13*(4), 464–472. doi: 10.1111/j.1461-0248.2010.01439.x

Hawkins, S. J., Bohn, K., & Doncaster, C. P. (2015). Ecosystems: The rocky road to regime-shift indicators. *Current Biology*, *25*(15), R666–R669.

Hefley, T. J., Tyre, A. J., & Blankenship, E. E. (2013). Statistical indicators and state–space population models predict extinction in a population of bobwhite quail. *Theoretical Ecology*, *6*(3), 319–331.

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

Holling, C. S. (1992). Cross-scale morphology, geometry, and dynamics of ecosystems. *Ecological Monographs*, *62*(4), 447–502.

Hughes, T. P. (1994). Catastrophes, phase shifts, and large-scale degradation of a caribbean coral reef. *Science*, *265*(5178), 1547–1551.

Hughes, T. P., Carpenter, S., Rockström, J., Scheffer, M., & Walker, B. (2013). Multiscale regime shifts and planetary boundaries. *Trends in Ecology & Evolution*, *28*(7), 389–395.

Jorgensen, S. E., Fath, B., Bastianoni, S., Marques, J. C., Muller, F., Nielsen, S. N., ... Ulanowicz, R. E. (2011). *A new ecology: Systems perspective*. Elsevier.

Jorgensen, S. E., & Svirezhev, Y. M. (2004). *Towards a Thermodynamic Theory for Ecological Systems*. Elsevier.

Karunanithi, A. T., Cabezas, H., Frieden, B. R., & Pawlowski, C. W. (2008). Detection and assessment of ecosystem regime shifts from fisher information. *Ecology and Society*, *13*(1).

Kefi, S., Guttal, V., Brock, W. A., Carpenter, S. R., Ellison, A. M., Livina, V. N., ... Dakos, V. (2014). Early warning signals of ecological transitions: Methods for spatial patterns. *PloS One*, *9*(3), e92097.

Kong, X., He, Q., Yang, B., He, W., Xu, F., Janssen, A. B., ... others. (2017). Hydrological regulation drives regime shifts: Evidence from paleolimnology and ecosystem modeling of a large shallow chinese lake. *Global Change Biology*, *23*(2), 737–754.

La Sorte, F. A., Lepczyk, C. A., Burnett, J. L., Hurlbert, A. H., Tingley, M. W., & Zuckerberg, B. (2018). Opportunities and challenges for big data ornithology. *The Condor: Ornithological Applications*, *120*(2), 414–426.

Lenton, T. M. (2011). Early warning of climate tipping points. *Nature Climate Change*, *1*(4), 201.

Lindegren, M., Dakos, V., Gröger, J. P., Gårdmark, A., Kornilovs, G., Otto, S. A.,

& Möllmann, C. (2012). Early detection of ecosystem regime shifts: A multiple method evaluation for management application. *PLoS One*, *7*(7), e38410.

Litzow, M. A., & Hunsicker, M. E. (2016). Early warning signals, nonlinearity, and signs of hysteresis in real ecosystems. *Ecosphere*, *7*(12), n/a–n/a. doi: 10.1002/ecs2.1614

Liu, J., Dietz, T., Carpenter, S. R., Alberti, M., Folke, C., Moran, E., ... others. (2007). Complexity of coupled human and natural systems. *Science*, *317*(5844), 1513–1516.

Liu, S., Yamada, M., Collier, N., & Sugiyama, M. (2013). Change-point detection in time-series data by relative density-ratio estimation. *Neural Networks*, *43*, 72–83.

Mac Nally, R., Albano, C., & Fleishman, E. (2014). A scrutiny of the evidence for pressure-induced state shifts in estuarine and nearshore ecosystems. *Austral Ecology*, *39*(8), 898–906.

Mantua, N. (2004). Methods for detecting regime shifts in large marine ecosystems: A review with approaches applied to North Pacific data. *Progress in Oceanography*, *60*(2), 165–182. doi: 10.1016/j.pocean.2004.02.016

Mayer, A. L., Pawlowski, C., Fath, B. D., & Cabezas, H. (2007). Applications of fisher information to the management of sustainable environmental systems. In *Exploratory data analysis using fisher information* (pp. 217–244). Springer.

Mayer, A. L., Pawlowski, C. W., & Cabezas, H. (2006). Fisher information and dynamic regime changes in ecological systems. *Ecological Modelling*, *195*(1-2), 72–82.

Michener, W. K., & Jones, M. B. (2012). Ecoinformatics: Supporting ecology as a data-intensive science. *Trends in Ecology & Evolution*, *27*(2), 85–93.

Möllmann, C., Folke, C., Edwards, M., & Conversi, A. (2015). *Marine regime shifts around the globe: Theory, drivers and impacts.* The Royal Society.

Mumby, P. J., Steneck, R. S., & Hastings, A. (2013). Evidence for and against the existence of alternate attractors on coral reefs. *Oikos*, *122*(4), 481–491.

Nash, K. L., Allen, C. R., Barichievy, C., Nyström, M., Sundstrom, S., & Graham, N. A. (2014). Habitat structure and body size distributions: Cross-ecosystem comparison for taxa with determinate and indeterminate growth. *Oikos*, *123*(8), 971–983.

Nes, E. H. van, & Scheffer, M. (2005). Implications of spatial heterogeneity for catastrophic regime shifts in ecosystems. *Ecology*, *86*(7), 1797–1807.

Nicholls, K. H. (2011). Detection of regime shifts in multi-species communities: The Bay of Quinte phytoplankton example. *Methods in Ecology and Evolution*, *2*(4), 416–426. doi: 10.1111/j.2041-210X.2011.00093.x

Nicholls, K., Hoyle, J., Johannsson, O., & Dermott, R. (2011). A biological regime shift in the bay of quinte ecosystem (lake ontario) associated with the establishment of invasive dreissenid mussels. *Journal of Great Lakes Research*, *37*(2), 310–317.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annu. Rev. Ecol. Evol. Syst.*, *37*, 637–669.

Perretti, C. T., & Munch, S. B. (2012). Regime shift indicators fail under noise levels commonly observed in ecological systems. *Ecological Applications*, *22*(6), 1772–1779.

Perretti, C. T., Munch, S. B., & Sugihara, G. (2013). Model-free forecasting outperforms the correct mechanistic model for simulated and experimental data.

Proceedings of the National Academy of Sciences, *110*(13), 5253–5257.

Peters, R. H., & Wassenberg, K. (1983). The effect of body size on animal abundance. *Oecologia*, *60*(1), 89–96.

Petersen, J. K., Hansen, J. W., Laursen, M. B., Clausen, P., Carstensen, J., & Conley, D. J. (2008). Regime shift in a coastal marine ecosystem. *Ecological Applications*, *18*(2), 497–510.

Price, Nathaniel B., & Burnett, J. L. (2019). *Tvdiff: An r package for numerical differentiation of noisy, nonsmooth data.*

Ratajczak, Z., Carpenter, S. R., Ives, A. R., Kucharik, C. J., Ramiadantsoa, T., Stegner, M. A., ... Turner, M. G. (2018). Abrupt change in ecological systems: Inference and diagnosis. *Trends in Ecology & Evolution.*

Reid, P. C., Hari, R. E., Beaugrand, G., Livingstone, D. M., Marty, C., Straile, D., ... Zhu, Z. (2016). Global impacts of the 1980s regime shift. *Global Change Biology*, *22*(2), 682–703. doi: 10.1111/gcb.13106

Restrepo, C., & Arango, N. (2008). Discontinuities in the geographical range size of north american birds and butterflies. *Discontinuities in Ecosystems and Other Complex Systems. Columbia University Press, New York, New York, USA*, 101–135.

Roberts, C. P., Twidwell, D., Burnett, J. L., Donovan, V. M., Wonkka, C. L., Bielski, C. L., ... others. (2018). Early warnings for state transitions. *Rangeland Ecology & Management*, *71*(6), 659–670.

Rockström, J., Steffen, W. L., Noone, K., Persson, Å., Chapin III, F. S., Lambin, E., ... others. (2009). Planetary boundaries: Exploring the safe operating space for humanity. *Ecology and Society.*

Rodionov, S. N. (2005). A brief overview of the regime shift detection methods. *Large-Scale Disturbances (Regime Shifts) and Recovery in Aquatic Ecosystems: Challenges for Management Toward Sustainability*, 17–24. Retrieved from `http://www.beringclimate.noaa.gov/regimes/rodionov_overview.pdf`

Rodionov, S., & Overland, J. E. (2005). Application of a sequential regime shift detection method to the Bering Sea ecosystem. *ICES Journal of Marine Science*, *62*(3), 328–332. doi: 10.1016/j.icesjms.2005.01.013

Roy Frieden, B. (1998). *Physics from Fisher information*. Cambridge: Cambridge University Press.

Sagarin, R., & Pauchard, A. (2012). *Observation and ecology: Broadening the scope of science to understand a complex world*. Island Press.

Salehpour, S., Gustafsson, T., & Johansson, A. (2011). An on-line method for estimation of piecewise constant parameters in linear regression models. *IFAC Proceedings Volumes*, *44*(1), 3171–3176.

Scheffer, M. (2009). *Critical transitions in nature and society*. Princeton University Press.

Scheffer, M., Carpenter, S., Foley, J. A., Folke, C., & Walker, B. (2001). Catastrophic shifts in ecosystems. *Nature*, *413*(6856), 591.

Scheffer, M., & Carpenter, S. R. (2003). Catastrophic regime shifts in ecosystems: Linking theory to observation. *Trends in Ecology & Evolution*, *18*(12), 648–656.

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

Seddon, A. W., Froyd, C. A., Witkowski, A., & Willis, K. J. (2014). A quantitative framework for analysis of regime shifts in a galápagos coastal lagoon. *Ecology*, *95*(11), 3046–3055.

Silva, M., & Downing, J. A. (1995). The allometric scaling of density and body mass: A nonlinear relationship for terrestrial mammals. *The American Naturalist*, *145*(5), 704–727.

Skillen, J. J., & Maurer, B. A. (2008). The ecological significance of discontinuities in body-mass distributions. *Discontinuities in Ecosystems and Other Complex Systems*, 193–218.

Smith, V. H., & Schindler, D. W. (2009). Eutrophication science: Where do we go from here? *Trends in Ecology & Evolution*, *24*(4), 201–207.

Sommer, S., Benthem, K. J. van, Fontaneto, D., & Ozgul, A. (2017). Are generic early-warning signals reliable indicators of population collapse in rotifers? *Hydrobiologia*, *796*(1), 111–120.

Spanbauer, T. L., Allen, C. R., Angeler, D. G., Eason, T., Fritz, S. C., Garmestani, A. S., … Stone, J. R. (2014). Prolonged instability prior to a regime shift. *PLoS One*, *9*(10), e108936.

Spanbauer, T. L., Allen, C. R., Angeler, D. G., Eason, T., Fritz, S. C., Garmestani, A. S., … Sundstrom, S. M. (2016). Body size distributions signal a regime shift in a lake ecosystem. *Proceedings of the Royal Society B: Biological Sciences*, *283*(1833), 20160249.

Steel, E. A., Kennedy, M. C., Cunningham, P. G., & Stanovick, J. S. (2013). Applied statistics in ecology: Common pitfalls and simple solutions. *Ecosphere*, *4*(9), 1–13.

Sugihara, G., May, R., Ye, H., Hsieh, C.-h., Deyle, E., Fogarty, M., & Munch, S. (2012).

Detecting causality in complex ecosystems. *Science*, *338*(6106), 496–500.

Sundstrom, S. M., Eason, T., Nelson, R. J., Angeler, D. G., Barichievy, C., Garmestani, A. S., ... Allen, C. R. (2017). Detecting spatial regimes in ecosystems. *Ecology Letters*, *20*(1), 19–32. doi: 10.1111/ele.12709

Takens, F. (1981). Detecting strange attractors in turbulence. In *Dynamical systems and turbulence, warwick 1980* (pp. 366–381). Springer.

Taranu, Z. E., Carpenter, S. R., Frossard, V., Jenny, J.-P., Thomas, Z., Vermaire, J. C., & Perga, M.-E. (2018). Can we detect ecosystem critical transitions and signals of changing resilience from paleo-ecological records? *Ecosphere*, *9*(10).

Taylor, L. (1961). Aggregation, variance and the mean. *Nature*, *189*(4766), 732.

Thrush, S. F., Hewitt, J. E., Dayton, P. K., Coco, G., Lohrer, A. M., Norkko, A., ... Chiantore, M. (2009). Forecasting the limits of resilience: Integrating empirical research with theory. *Proceedings of the Royal Society B: Biological Sciences*, *276*(1671), 3209–3217.

Vasilakopoulos, P., Raitsos, D. E., Tzanatos, E., & Maravelias, C. D. (2017). Resilience and regime shifts in a marine biodiversity hotspot. *Scientific Reports*, *7*(1), 13647.

Walker, B., Holling, C. S., Carpenter, S., & Kinzig, A. (2004). Resilience, adaptability and transformability in social–ecological systems. *Ecology and Society*, *9*(2).

Walther, G.-R., Post, E., Convey, P., Menzel, A., Parmesan, C., Beebee, T. J., ... Bairlein, F. (2002). Ecological responses to recent climate change. *Nature*, *416*(6879), 389.

Weijerman, M., Lindeboom, H., & Zuur, A. F. (2005). Regime shifts in marine ecosystems of the north sea and wadden sea. *Marine Ecology Progress Series*, *298*, 21–39.

Weissmann, H., & Shnerb, N. M. (2016). Predicting catastrophic shifts. *Journal of Theoretical Biology*, *397*, 128–134.

Wolkovich, E. M., Cook, B. I., McLauchlan, K. K., & Davies, T. J. (2014). Temporal ecology in the Anthropocene. *Ecology Letters*, *17*(11), 1365–1379. doi: 10.1111/ele.12353

Yang, Q., & Wu, X. (2006). 10 challenging problems in data mining research. *International Journal of Information Technology & Decision Making*, *05*(04), 597–604. doi: 10.1142/S0219622006002258

Ye, H., Beamish, R. J., Glaser, S. M., Grant, S. C., Hsieh, C.-h., Richards, L. J., … Sugihara, G. (2015). Equation-free mechanistic ecosystem forecasting using empirical dynamic modeling. *Proceedings of the National Academy of Sciences*, *112*(13), E1569–E1576.

Yin, D., Leroux, S. J., & He, F. (2017). Methods and models for identifying thresholds of habitat loss. *Ecography*, *40*(1), 131–143.

Zhou, T., & Shumway, R. (2008). One-step approximations for detecting regime changes in the state space model with application to the influenza data. *Computational Statistics & Data Analysis*, *52*(5), 2277–2291.