Decoupling the Calculation of Fisher Information

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

Introduction

Changes in the feedback(s) governing ecosystem processes can trigger unexpected and sometimes undesirable responses in environmental conditions [@scheffer_catastrophic_2001; @walther_ecological_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_ecological_2009]. Climate change and globalization are triggering novel and unexpected changes in ecosystems [@hughes_catastrophes_1994;@scheffer_catastrophic_2001 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_towards_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 [@groffman_ecological_2006; @deyoung_regime_2008; @carpenter2011early; @sagarin_observation_2012; @wolkovich_temporal_2014; @jorgensen_towards_2004].

Numerous quantitative approaches have been proposed as regime shift detection methods [@clements_including_2016 ; @rodionov_application_2005 @andersen_ecological_2009; @mantua_methods_2004], but few are consistently applied to terrestrial ecological data [@deyoung_regime_2008]. I broadly classify these methods as either model-based or model-free [@boettiger_quantifying_2012; @hastings_regime_2010; @dakos_methods_2012. Model-based methods use mathematical (mechanistic) representations of the system [@hefley2013statistical], which often carrying strict assumptions that are easily violated by dynamic systems such as ecosystems [@abadi2010assessment]. Further, model misspecification may yield spurious results [@perretti_model-free_2013]. Model-free [or metric-based, per @dakos_methods_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 [@mantua_methods_2004; @rodionov_application_2005; @andersen_ecological_2009] and composite measures of multiple variables, notably principal components analysis [@petersen2008regime;@mollmann2015marine], clustering algorithms [@beaugrand2004north], and variance index [@brock_variance_2006].

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_towards_2002; @karunanithi_detection_2008]. As a multivariate, model-free method, Fisher Information integrates the information present in the entire data of a system and distills this

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

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 [@burthe2016early]. 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.

The Sustainable Regimes Hypothesis

Fisher Information [hereafter, FI; @fisher_mathematical_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 [@mayer_applications_2007;@karunanithi_detection_2008; @eason_managing_2014;@eason_evaluating_2012]. 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_physics_1998;@frieden_fisher_1998], and (iii) a measure representing the dynamic order/organization of a system [@cabezas_towards_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_towards_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 $t??? \inf$ 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 $t??? \inf$ FI approaches zero (no information). Per the Sustainable Regimes Hypothesis [@cabezas_towards_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.

###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_prolonged_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.

###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_regime_2003; @mayer_applications_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.

Methods

###Predator-Prey Model System Our model system is a two-species predator-prey model [Eq. \@ref(eq:predprey); @fath_regime_2003; @mayer_applications_2007; @frieden_exploratory_2007], hereafter referred to as the "model system":

$$dx_1 = g_1 x_1 (1 - \frac{x_1}{k}) - \frac{l_{12} x_1 x_2}{1 + \beta x_1} dx_2 = \frac{g_{21} x_1 x_2}{1 + \beta x_1} - m_2 x_2)(\#eq : predprey) \tag{1}$$

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_regime_2003; @mayer_applications_2007; @frieden_exploratory_2007]. The initial conditions (predator and prey abundances,) for the model system were not provided in the original references [@fath_regime_2003; @mayer_applications_2007]. I used package `deSolve` in Program R (version 3.3.2) to solve the model system (Eq. Eq. @ref(eq:predprey)), 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.

###Inducing a Regime Shift @mayer_applications_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. @ref(eq:mayerCase) where k_1 is the initial carrying capacity, k_2 is the final carrying capacity, $t_s hift$ 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. @ref(eq:mayerCase)) 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. @ref(eq:mayerCase)). I assume $k_1 = 800$ and $k_2 = 625$, values corresponding to the range of carrying capacities explored by @mayer_applications_2007. I simulated a time series of 600 time units, introducing a regime change after 200 time units, and $\alpha = 0.05$.

$$k(t) = k_1 - 0.5(k_1 - k_2)(\tanh(\alpha(t - t^*)) + 1) k'(t) = 0.5\alpha(k_1 - k_2)(\tanh(\alpha(t - t^*))^2 + 1)(\#eq : mayerCase) \tag{2}$$

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_towards_2002 and the binning" method (first appearing in @karunanithi_detection_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_regime_2003;@mayer_applications_2007] and I refer the reader to [@cabezas_towards_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. @ref(eq:s)

$$\frac{ds}{dt} = \sqrt{(}\frac{dx_1}{dt}^2 + \frac{dx_2}{dt}^2)(\#eq:s) \tag{3}$$

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_towards_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_applications_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_towards_2002]. Equation @ref(eq:fiAdapted) is a general form of FI and Equation @ref(eq:fiDerivs) is the form used in the derivative-based method for FI [see eq. 7.3b and 7.12 in @mayer_applications_2007]. To better understand the FI calculation, note that Eq.@ref(eq:fiDerivs) 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. @ref(eq:fiDerivs) 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. @ref(eq:fiDerivs)).

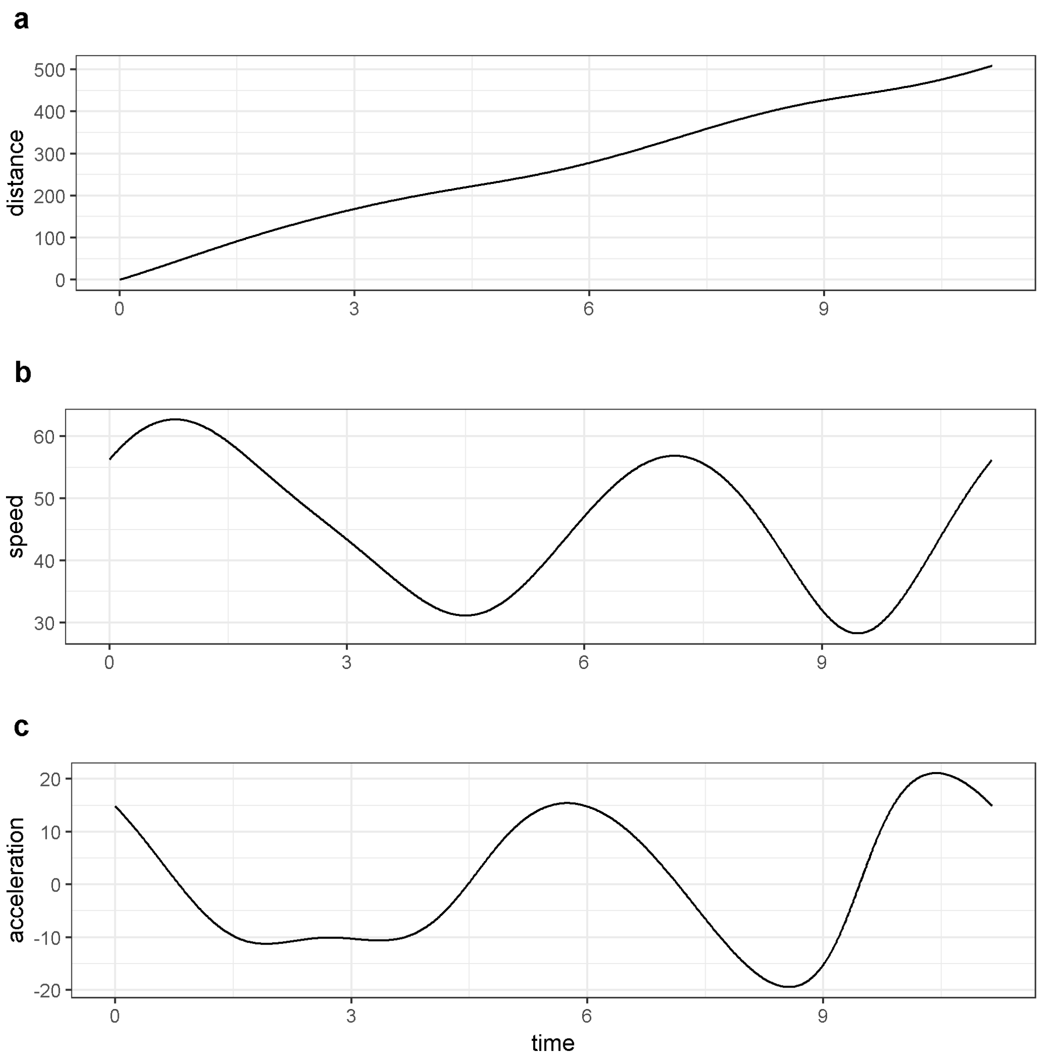

Figure 1: From top to bottom, distance traveled in phase space, speed tangential to system trajectory, acceleration tangential to system trajectory.

$$I = \int \frac{ds}{p(s)} \left[\frac{dp(s)}{ds}\right]^2 (\#eq:fiAdapted) \tag{4}$$

$$I = \frac{1}{T} \int_0^T dt \left[\frac{s''^2}{s'^4}\right]^2 (\#eq:fiDerivs) \tag{5}$$

##Results Distance travelled (s), speed ($\frac{ds}{dt}$), and acceleration ($\frac{d^2s}{dt^2}$) capture the dynamics of the model system (Eq. @ref(eq:predprey); Fig. @ref(fig:distSpeedAccel)). I simulated a regime shift in the carrying capacity of this model system, at approximately $t = 200$ (Fig. @ref(fig:kByTime)). 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. @ref(fig:kTrajectories)). 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. @ref(fig:sOverTime)) 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. @ref(fig:dsdtOverTime)).

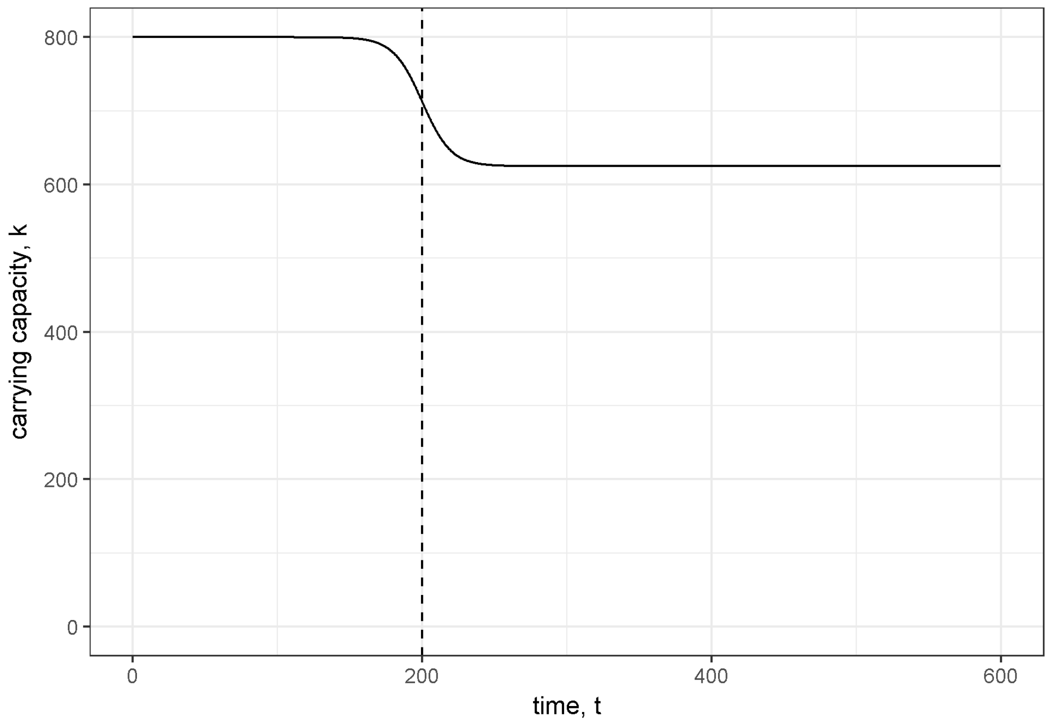

Figure 2: Carrying capacity over time with a regime shift occuring around time 200.

I calculated FI for the distribution of s over a series of non-overlapping time windows. According to @mayer_applications_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. @ref(fig:fiOverTime)).

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_applications_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.

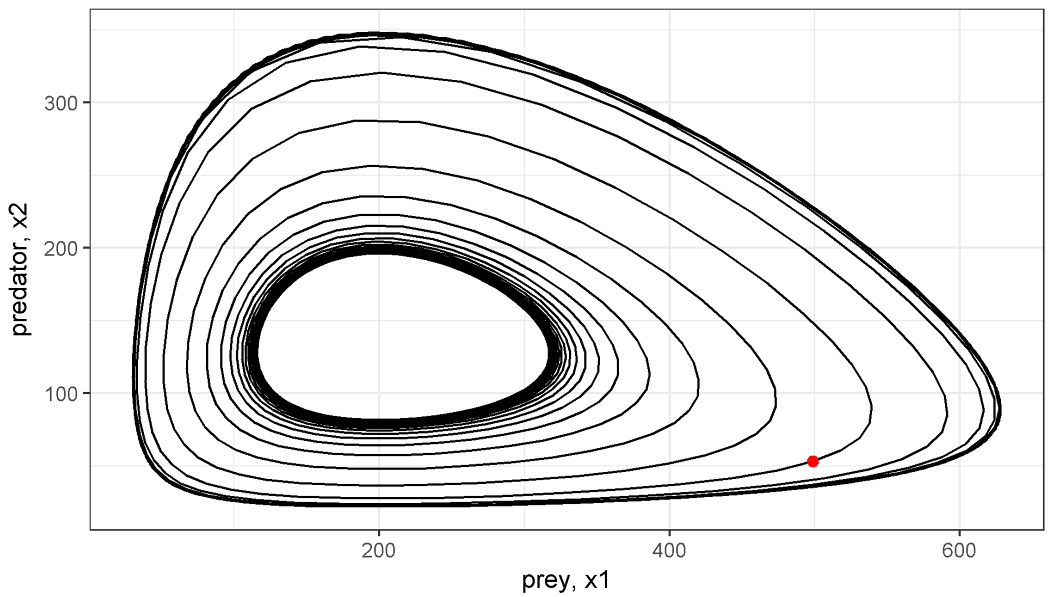

Figure 3: Phase space plot of system trajectories for different values of k

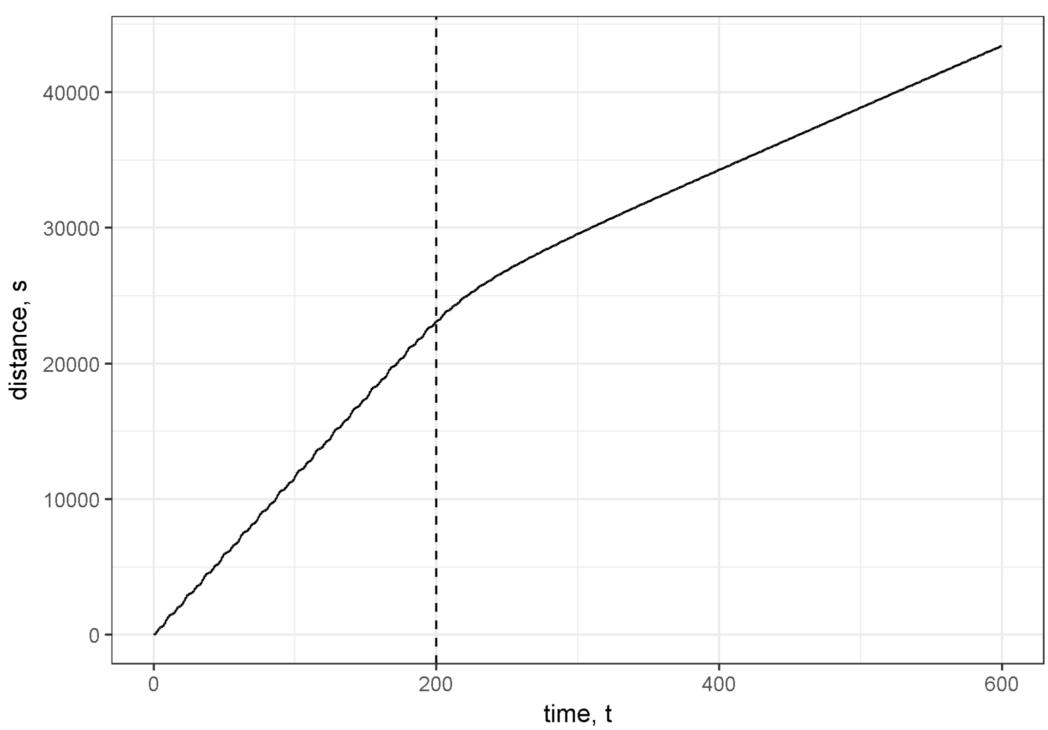

Figure 4: Distance travelled in phase space over time. Dashed vertical line at *t*=200 indicates location of regime shift.

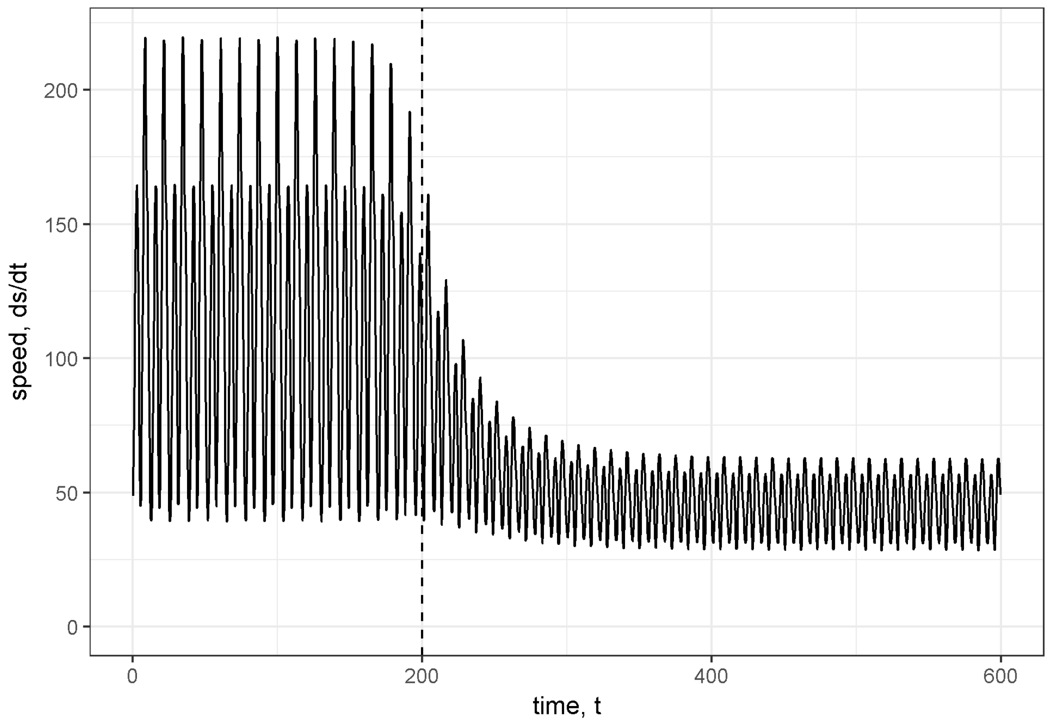

Figure 5: Speed of the system (rate of change) in phase space. Dashed vertical line at time 200 indicates location of regime shift.

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_detection_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 [@mantua_methods_2004; @fath_regime_2003], not unlike most regime detection methods [Ch. @ref(rdmReview)]. 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_towards_2002; @karunanithi_detection_2008; @spanbauer_prolonged_2014]. Although the aforementioned and numerous other works interpret FI in this context [e.g., @eason_managing_2014;@eason_evaluating_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.

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).

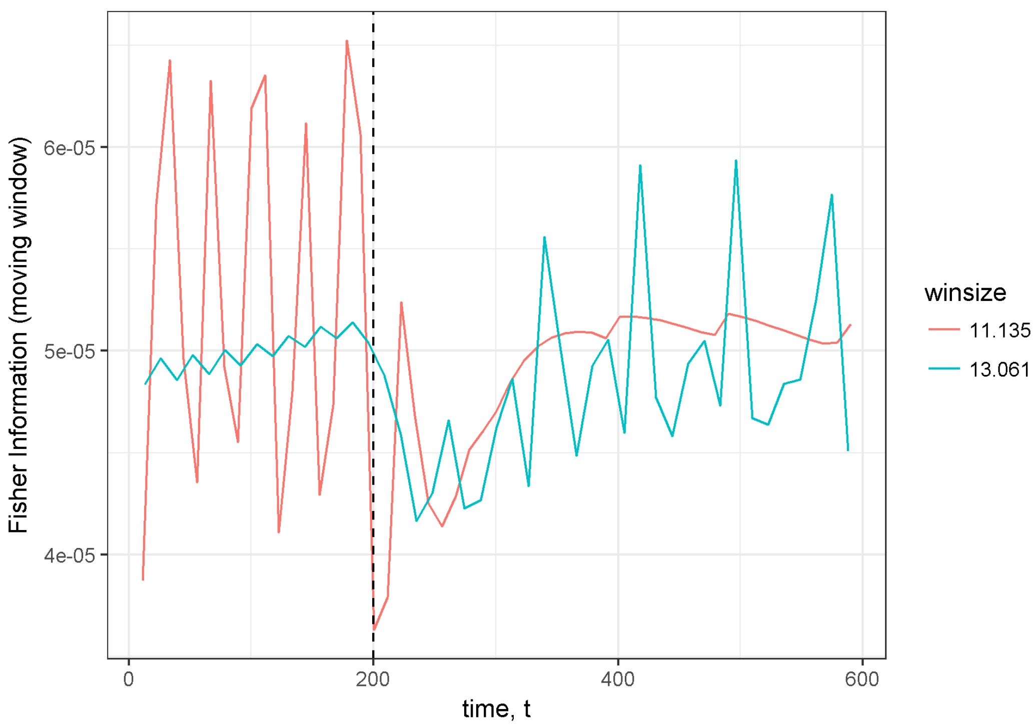

Figure 6: Fisher Information calculated for non-overlapping time windows. Two different window sizes were used as indicated by color. Dashed vertical line at time 200 indicates approximate location of regime shift.