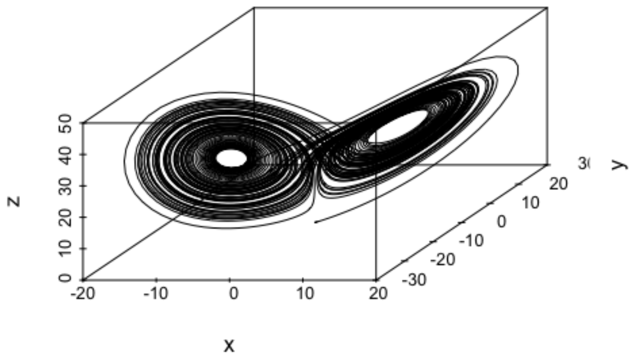


Figure 1: An example solution of the Lorenz ('butterfly') represented in its 3-dimensional 3-dimensional phase-space. Phase plots are typically used to visualize stable areas within a system's trajectory but reconstruction requires the difference models to be known and parameterized.

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

Introduction

When and how systems exhibit abrupt and undesireable change is a hallmark of modern regime shift ecology and can extremely useful when the change has potential to impact society. Quantitatively detecting and forecasting these changes, however, has yet to be accomplished for noisy, ecological systems data (Chapter @ref(rdmReview)). Although statistical

If the system in question cannot be described by a system of equations (i.e. first-difference equations).

Modelling the location, direction, and rate of change of ecosystems can be a powerful tool for understanding dynamical systems.

Visualizing the displacement of systems in space-time has been used to demonstrate theoretical dynamical systems.

For example, the classic 'butterfly' (Lorez) attractor (Fig. @ref(fig:lorenz3D)) is an informative visual of the state-switching behavior of a chaotic system. Although the behavior in phase space are used often in dynamical systems theory and systems ecology to make inference regarding system behavior and dynamics, they have yet to be used outside theoretical studies as a tool ecological data analysis [@takens1981detecting]

c.f. @sugihara2012detecting for an example of phase-space reconstruction using Taken's theorem of ecological time series]. Although methods for reconstructing attractors in ecological data have been explored , reconstruction methods they do not explicitly incorporate the dynamics of whole-systems.

Taken's embedding theorem [@sugihara2012detecting; @ye2015equation]

```
knitr::include_graphics(here::here("/chapterFiles/velocity/figsCalledInDiss/lorenz3D.png"))
```

Here I propose a method which simply describes the rate of change behavior of system dynamics in phase space: **velocity**, V. Reconstructing the phase spaces of noisy ecological data is difficult, and in systems that are data-limited, not possible.

Phase space reconstruction

lekscha2018phase - a new method for PS recon..

Multivariate methods for ecosystem trajectory

Rate of change

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** ($\mathbf{S}$) or **velocity** ($\mathbf{V}$), where ($\mathbf{S}$) is the adirectional magnitude (i.e. it is a scalar) of the displacement of an object over unit time and $\mathbf{V}$ describes both the direction and magnitude (i.e. it is a vector) of the object's movement in spacetime. $\mathbf{S}$ is a scalar taking values of ≥ 0 and $\mathbf{V}$ 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 $\mathbf{S}$ is constant, $\mathbf{V}$ changes in a sunusoidal fashion, where $\mathbf{V}$ 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 $\mathbf{S}$ is useful when estimating other scalar quantities (e.g., $\frac{miles}{gallon}$), given a starting and/or final position in space, $\mathbf{S}$ is not informative of its the path travelled.

ROC is used to simplify

lischeid_monitoring_2016- phase space to reconstruct niche assembly hypotheses - theoretical paper

howROC has been used in ecological systems analysis

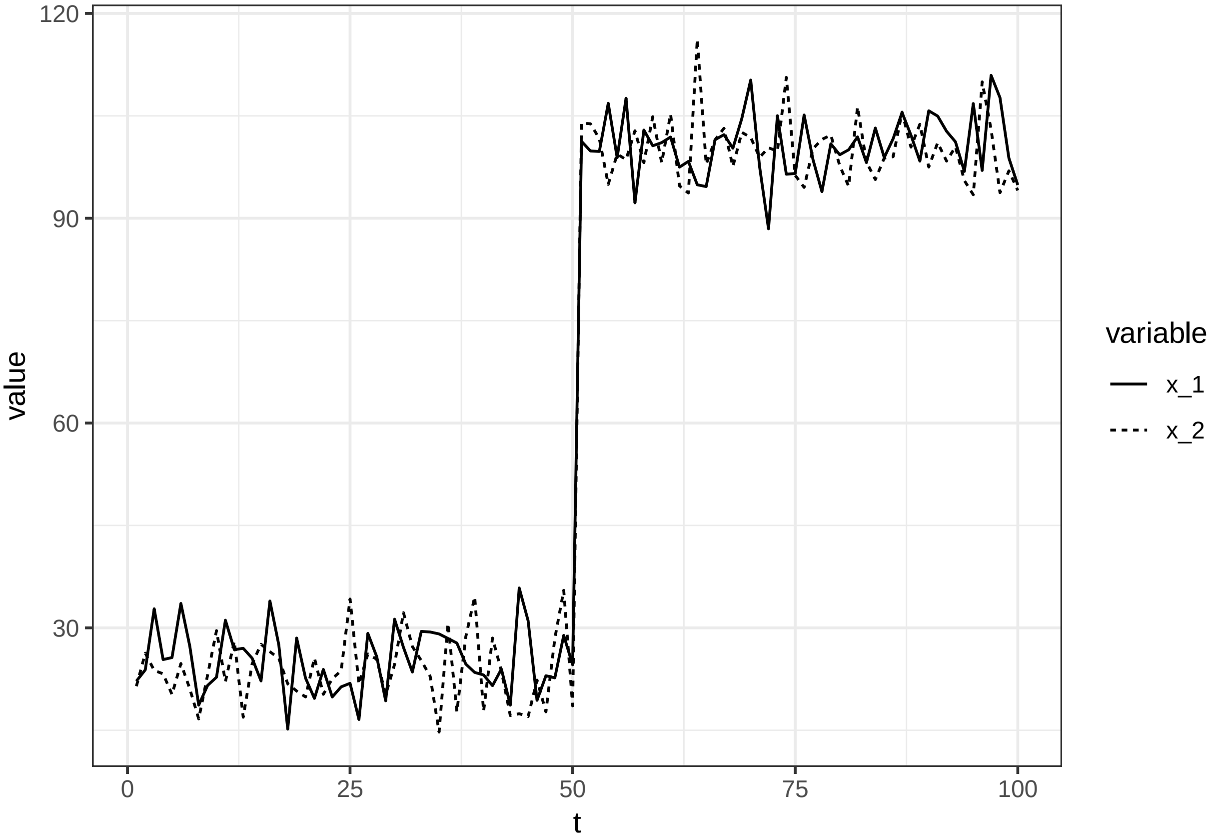
- modern: gams
- op;ld: ### what it may indicate.. The speed at which a system undergoes changes ### some papers tilman1996biodiversity- ROC was found to be high in plant community biomass prior to a drought – they used the ratio of the commuity biomasses in epiermental polota (he also compared to CV) heimann2008terrestrial-

Aims

In this Chapter I describe the steps for calculating a 'new' metric, **system velocity** (v), as a method for reducing the dimensionality and identifying abrupt shifts in high dimensional data. 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. @ref(fiGuide)], first introduced in @fath_regime_2003. Below, I describe the steps for calculating system velocity, simply defined as the cumulative sum of the squared change in all state variables over a period of time.

Data and Methods

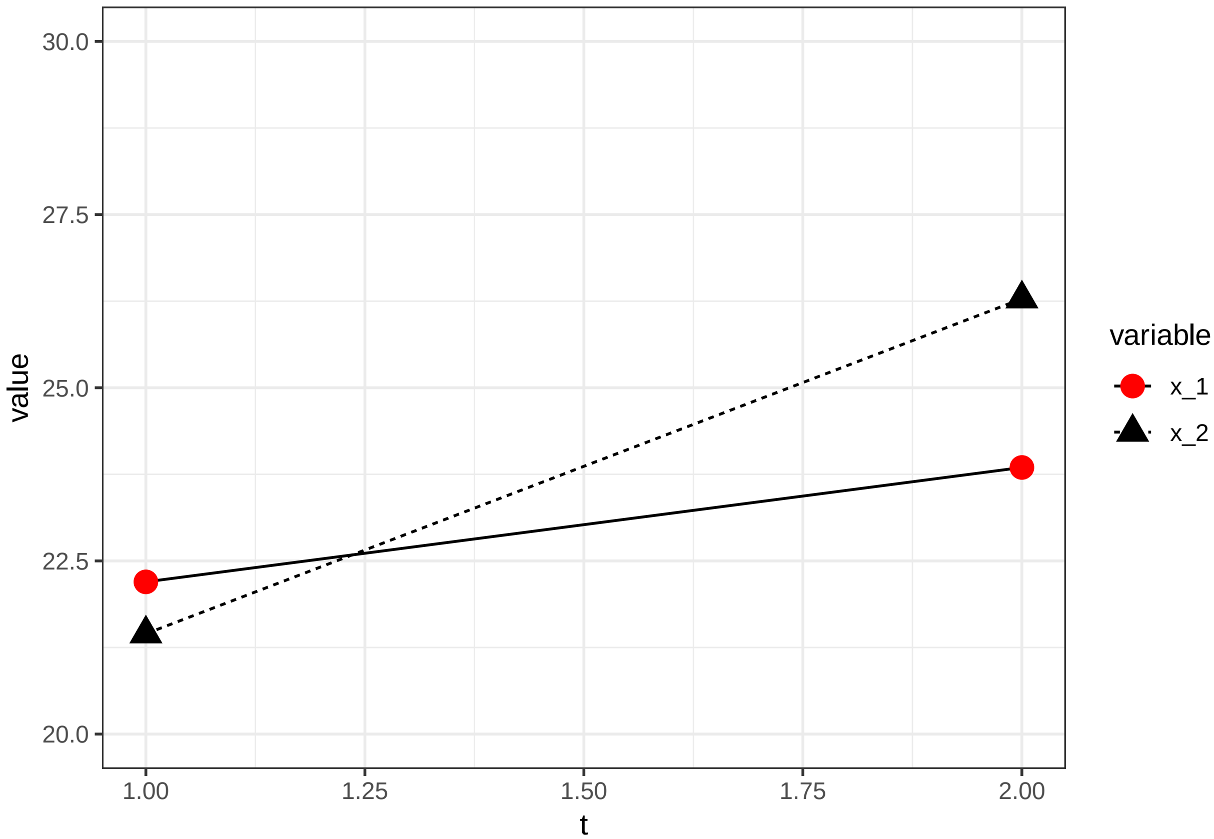
Theoretical system example: two-species time series



Consider a system (Fig. @ref(fig:sysEx)) 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 a simple, two-variable system to demonstrate the calculation of each step below. The system comprises variables x_1 and x_2, with observations occurring at each time point $t = 1, 2, 3, ...100$.

Steps for calculating system velocity, v

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. @ref(fig:sysEx2)). 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 occurs at $t = 50$, where a shift in either or both $\bar{x}_i$ or σ_i.



Step 1: Calculate Δx_i

The first step in calculating v is to obtain the change in values for each state variables, x_1 and x_2 between two consecutive time points (e.g., from $t = 1$ to $t = 2$:

$$\begin{aligned} \Delta x_1 = x_{1_{t=2}} - x_{1_{t=1}} \\ \Delta x_2 = x_{2_{t=2}} - x_{1_{t=1}} \end{aligned} \; (\#eq:diffX) \tag{1}$$

Step 2: Calculate $\sqrt(\sum_i^N \Delta x_1^2)$

After calculating the differences for each state variable, we will next calculate the total change in the system over the time elapsed, following Pythagora's theorem,

$$X_1^2 + X_2^2 = s^2 (\#eq:pythagorean) \tag{2}$$

where s represents the total change in the system, and X_1 and X_2 represent the changes in all state variables $(x_{i_{t=2}} - x_{i_{t=1}})$. We achieve this by first squaring the differences obtained in Eq. @ref(eq:diffX):

$$\begin{aligned} (x_{1_{t=2}} - x_{1_{t=1}})^2 \\ (x_{2_{t=2}} - x_{2_{t=1}})^2 \end{aligned} \; (\#eq:diffXsq) \tag{3}$$

Step 3: Use Pythagorean theorem to isolate s

Next, we isolate s in Eq. @ref(eq:pythagorean), capturing the total change in all state variables into a single measure by taking the 2nd root of the squared sums of all x:

$$\begin{aligned} \sum_{i=1}^N \Delta x_i = \sum_{i=1}^N (x_{t_{i+1}} - x_{t_i})^2 \\ = \Delta s \; (\#eq:diffXsq2) \\ = \sqrt([x_{1_{t=2}} - x_{1_{t=1}}]^2 + [x_{2_{t=2}} - x_{2_{t=1}}]^2) \end{aligned} \tag{4}$$

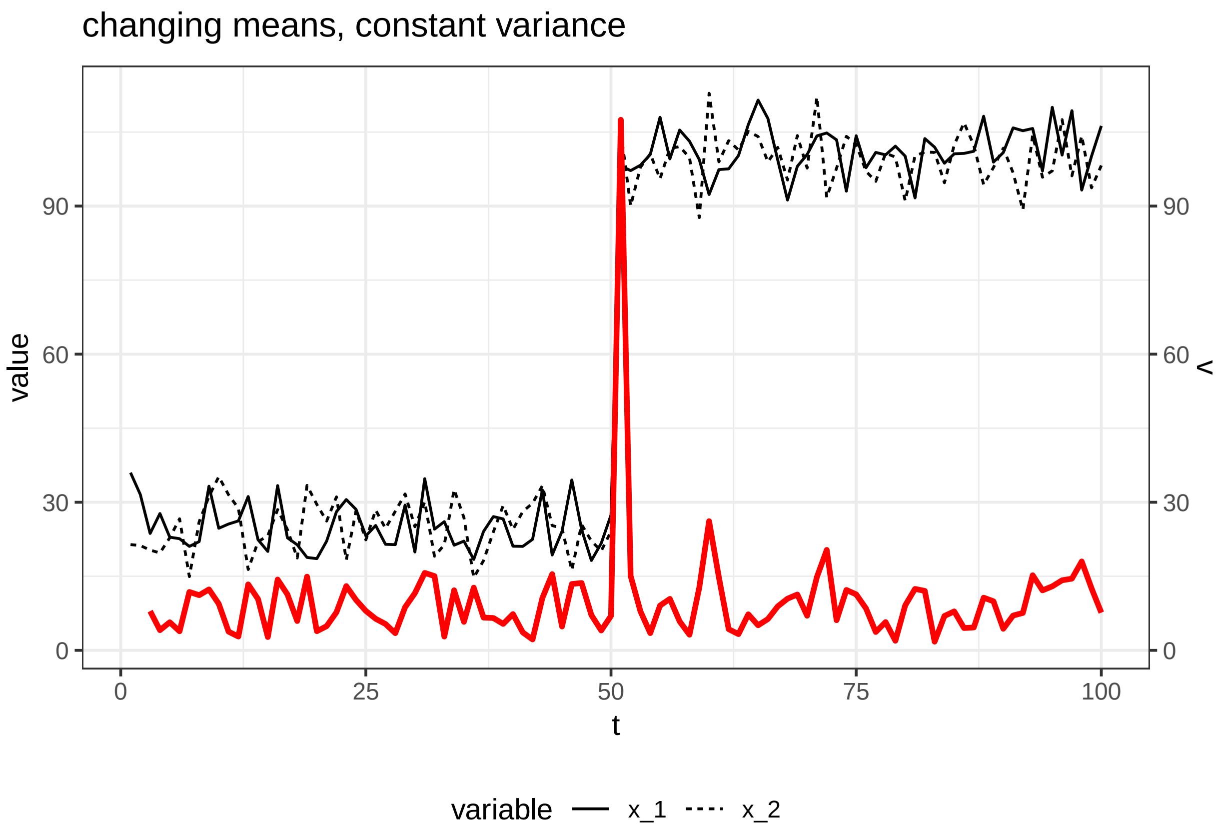
We now have a single measure, Δs (Eq. @ref(eq:diffXsq2)), for each pair of time points in our N-dimensional system. It is obvious that Δs will always be a positive value, since we took the 2nd root of a squared value. Although discussed in a later section, it is important to note that this value is not unitless–that is, our example system takes on the units of our state variables, x_1 and x_2. Because we are interested in identifying abrupt changes in the entire system, we calculate the cumulative sum of Δs at every time point, such that:

$$s = \sum_{t=1}^{T} \Delta s (\#eq:s) \tag{5}$$

Step 4: 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. @ref(eq:s)), and then divide by the total time elapsed between consecutive sampling points:

$$v = \frac{s_{t+1} - s_t}{\Delta t} (\#eq:velocity) \tag{6}$$



The steps for calculating velocity [Eq. @ref(eq:velocity)] are demonstrated using the first five time points of our toy system (Fig. @ref(fig:sysEx)) in Table @ref(tab:distTab).

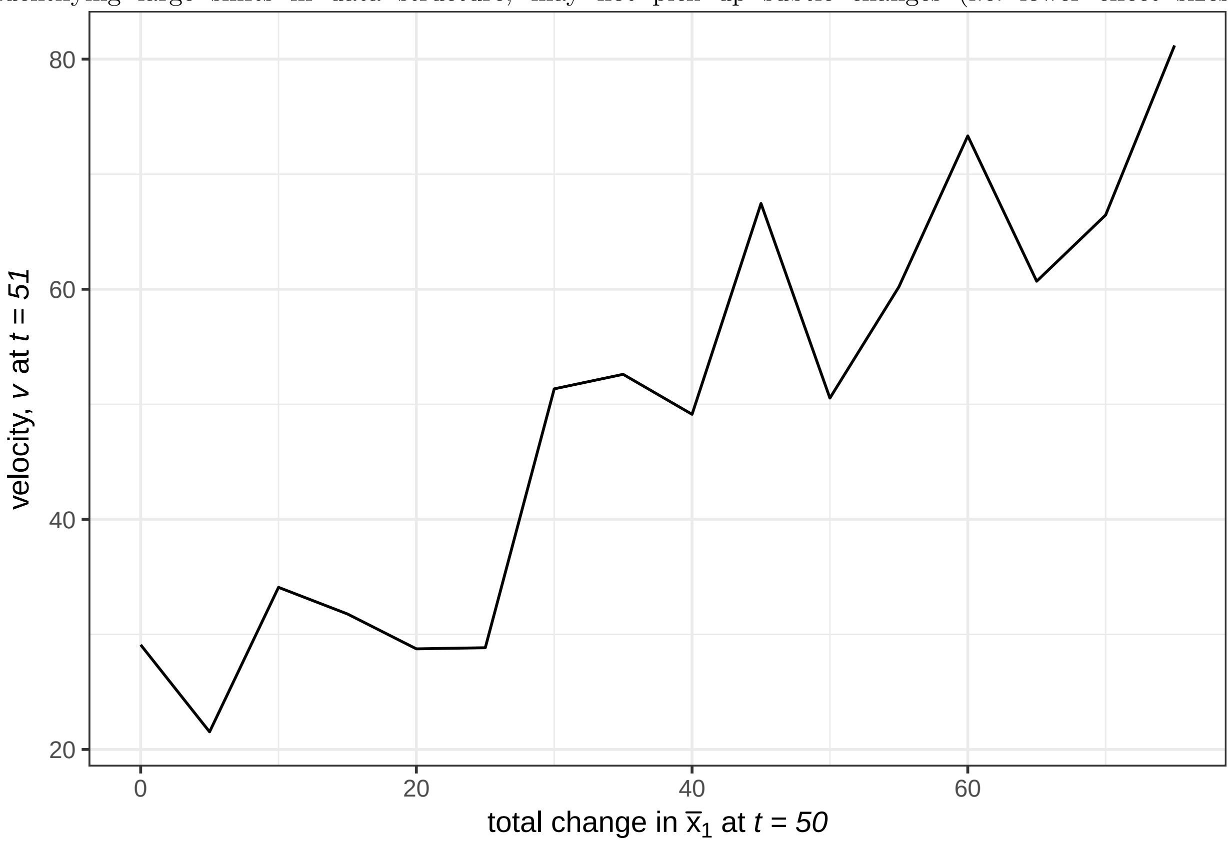
Velocity v performance under varying mean and variance in the toy system

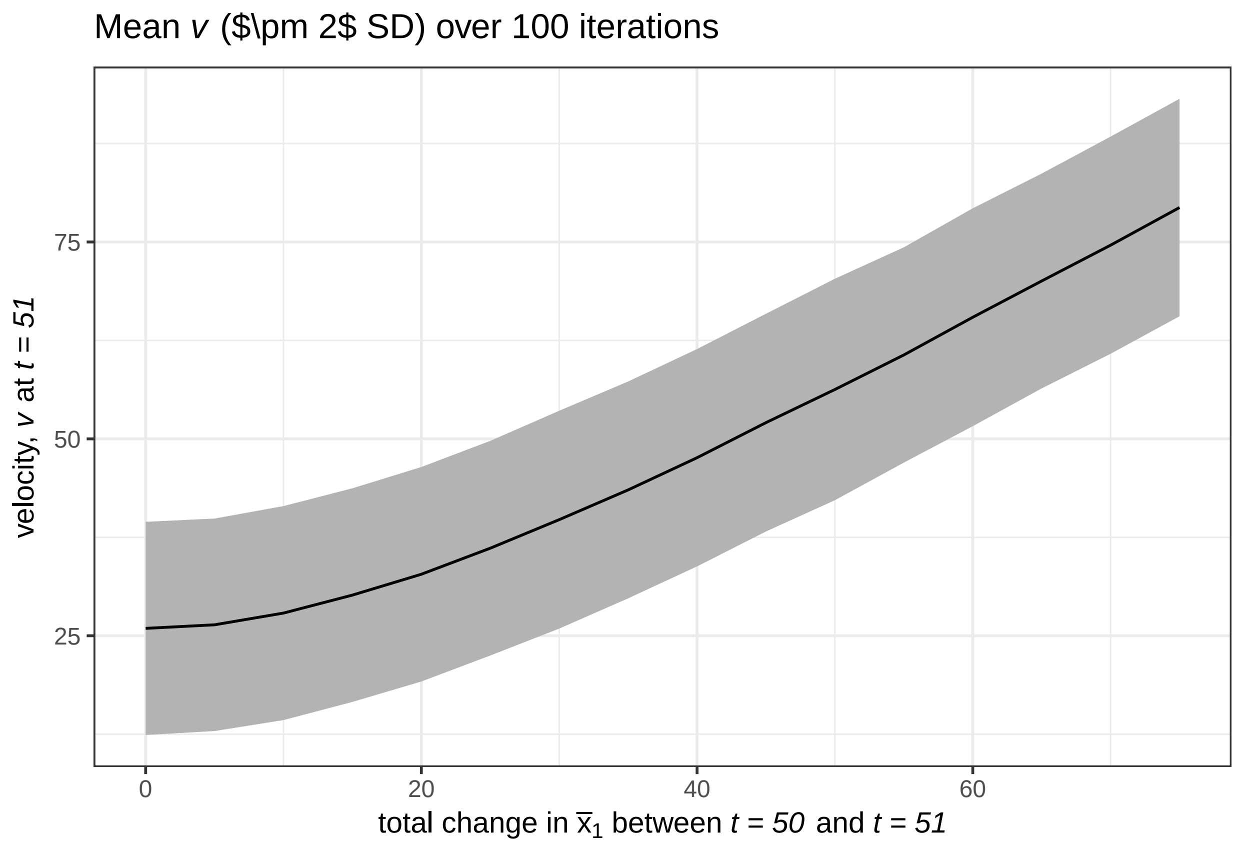
I simulated 10,000 random draws of the toy system, which experiences a rapid shift at $t = 50$, while varying two each of the following system paramters at the regime shift: $\bar{x}_1$, increased the mean value of x_1 σ_1, change in variance of x_1 Simulations consisted of 10,000 random samples drawn 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 mentione 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.

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. @ref(fig:simVplot1),@ref(simVplot2)). As expected, the magnitude of v increased linearly as the total difference between $\bar{x}_{1_{pre}}$ and $\bar{x}_{1_{post}}$ increased (@ref(fig:simVplot2)). This is not surprising, as s increases as the total change in abundance across the entire sytem increases (Eq. @ref(eq:s)), therefore, the potential maximum of v also increases. This may indicate that v, while capable of identifying large shifts in data structure, may not pick up subtle changes (i.e. lower effect sizes).





Varying post-shift variance

In the previous example, variance was constant before and after the shift at $t = 50$. To determine whether the signal emitted by v at the regime shift is lost with increasing variance, I varied the variance parameter, σ_1 in the set $W = \{1, 2, 3, ...25\}$. The variance for both state variables prior to the regime shift, σ_1 and σ_2, was 5, with the change occurring in σ_{1post}. Sytem velocity v appears senstive to increases in the variance at the point of the regime shift (Figs. @ref(fig:simVarplot), @ref(fig:simVarplot2)). This extreme sensitivity of v to σ_{post} (Fig. @ref(fig:simVarplot2)) is unsurprising, given the fact that, without smoothing the derivatives, the tangential speed of a 'noisy' variable will always be noisy itself (see Figs. @ref(velocitySysEx1), @ref(velocitySysEx2), @ref(velocitySysEx3), @ref(velocitySysEx4)).

Smoothing the data prior to calculating v

To ameliorate the influence of noise (e.g. Fig. @ref(simVarPlot)) on the regime shift signal in v, I used linear approximation techniques in attempt to smooth the velocity (derivatives). I used the function *stats::approx* to interpolate values of x_1 and x_2 to regularly-spaced time points in the set $t = \{1 : 100\}$, and then calculated v as described in the steps above (Eqs. @ref(eq:diffX):@ref(eq:velocity)). Increasing the number of points (t) at which the original state variables were smoothed did not influence the amount of noise surrounding the signal of the regime shift (at $t = 50$) in system velocity, v (Fig. @ref(fig:smoothV)).

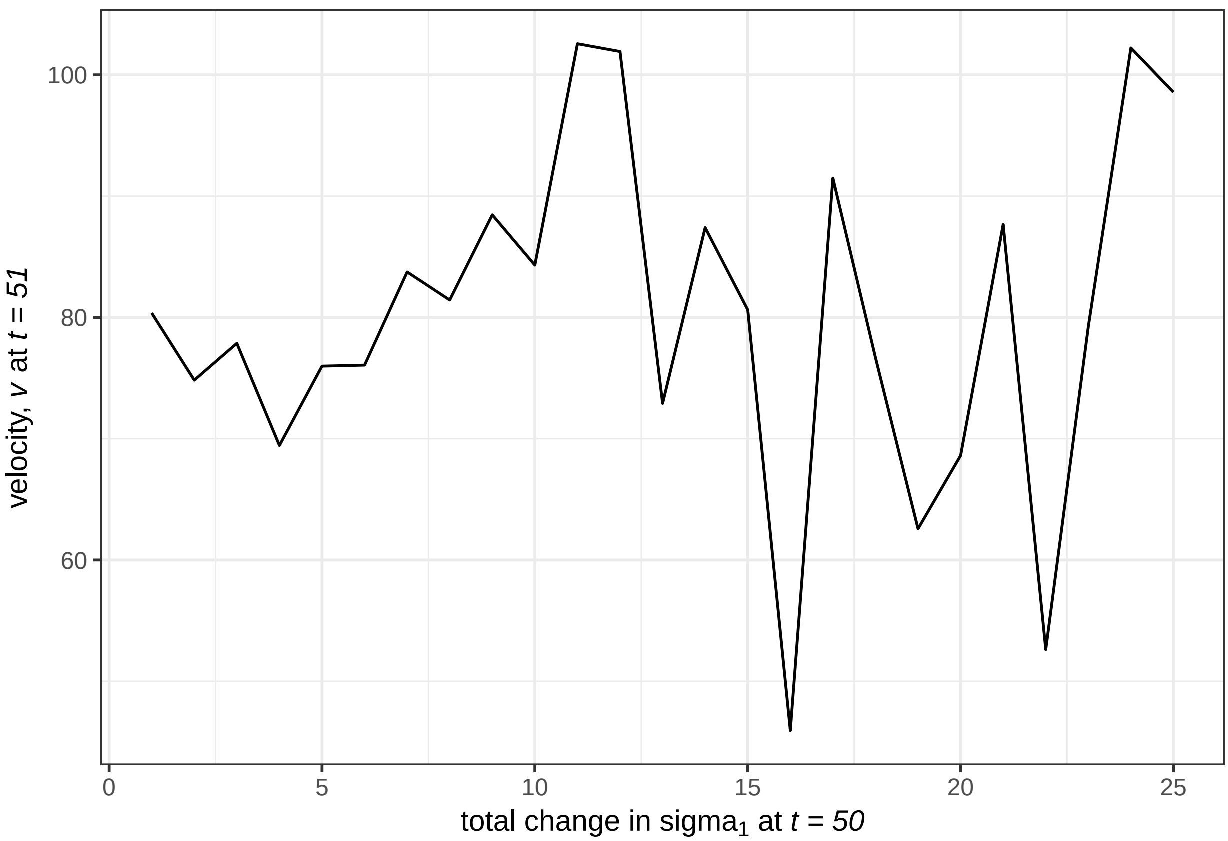


Figure 2: High variance of velocity (v) in a single iteration ($N_{iter} = 1$, seed = 123) of simulations as we increase σ_1 at $t = 50$.

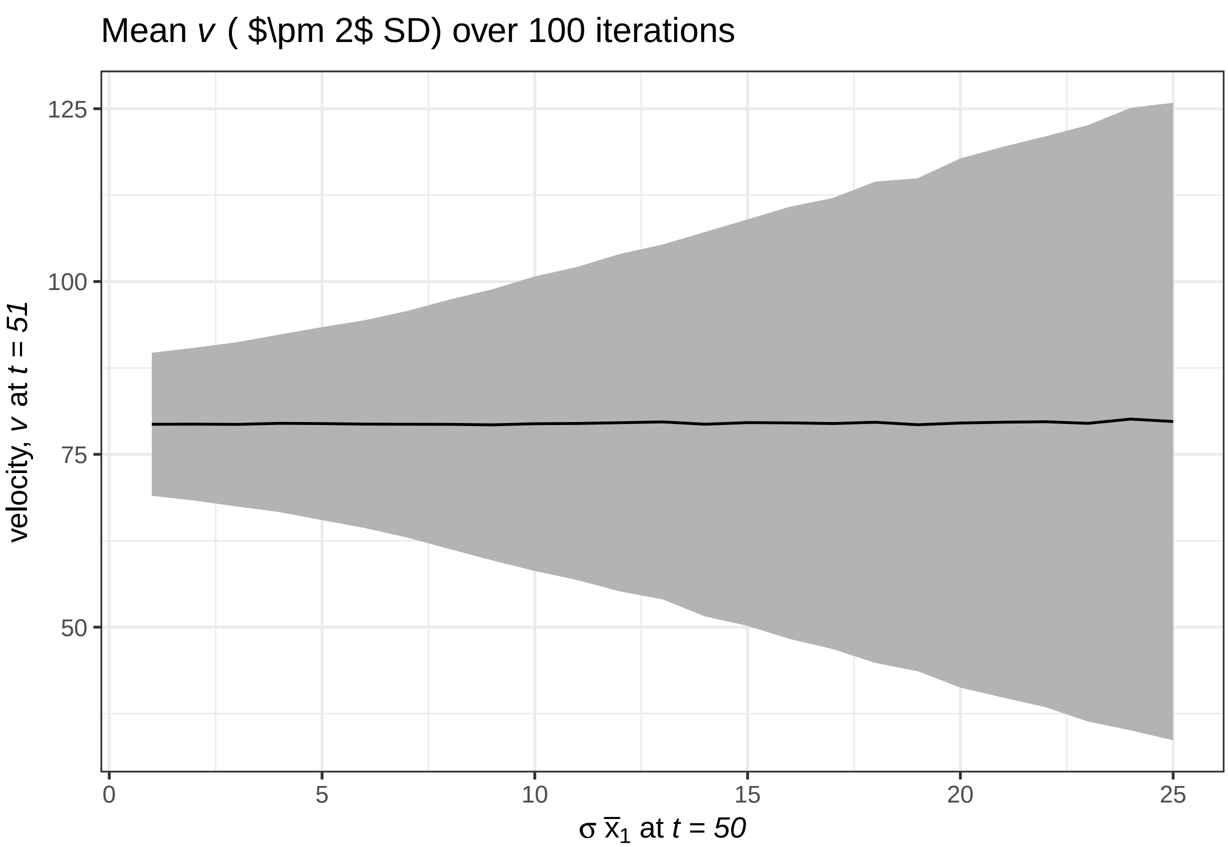
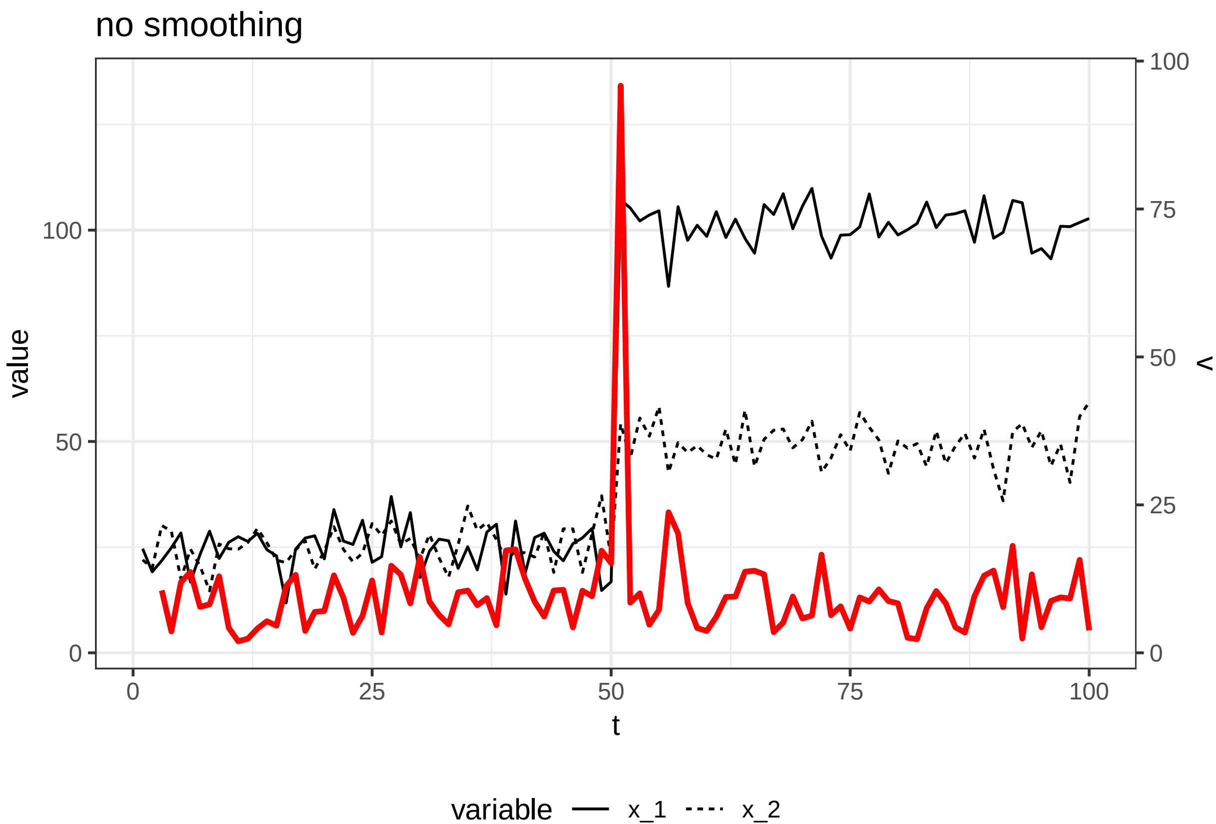
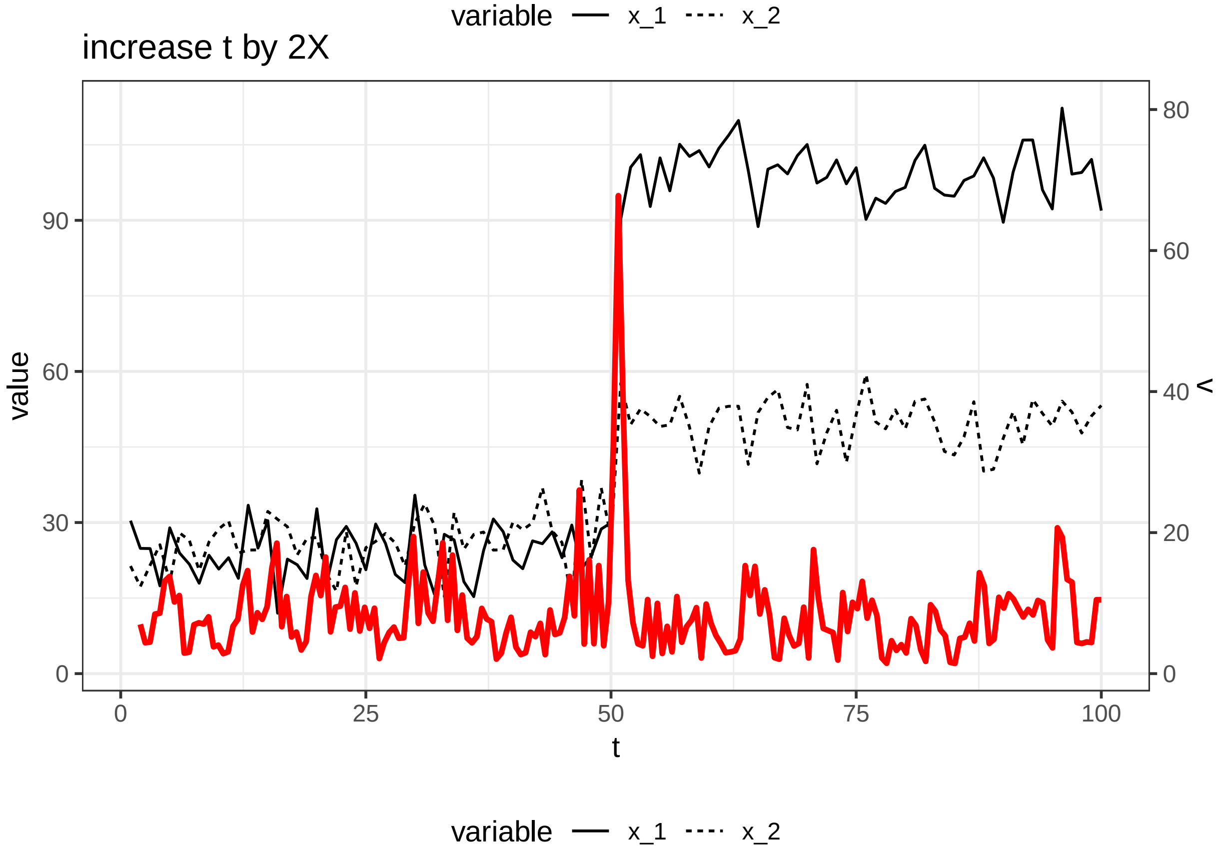
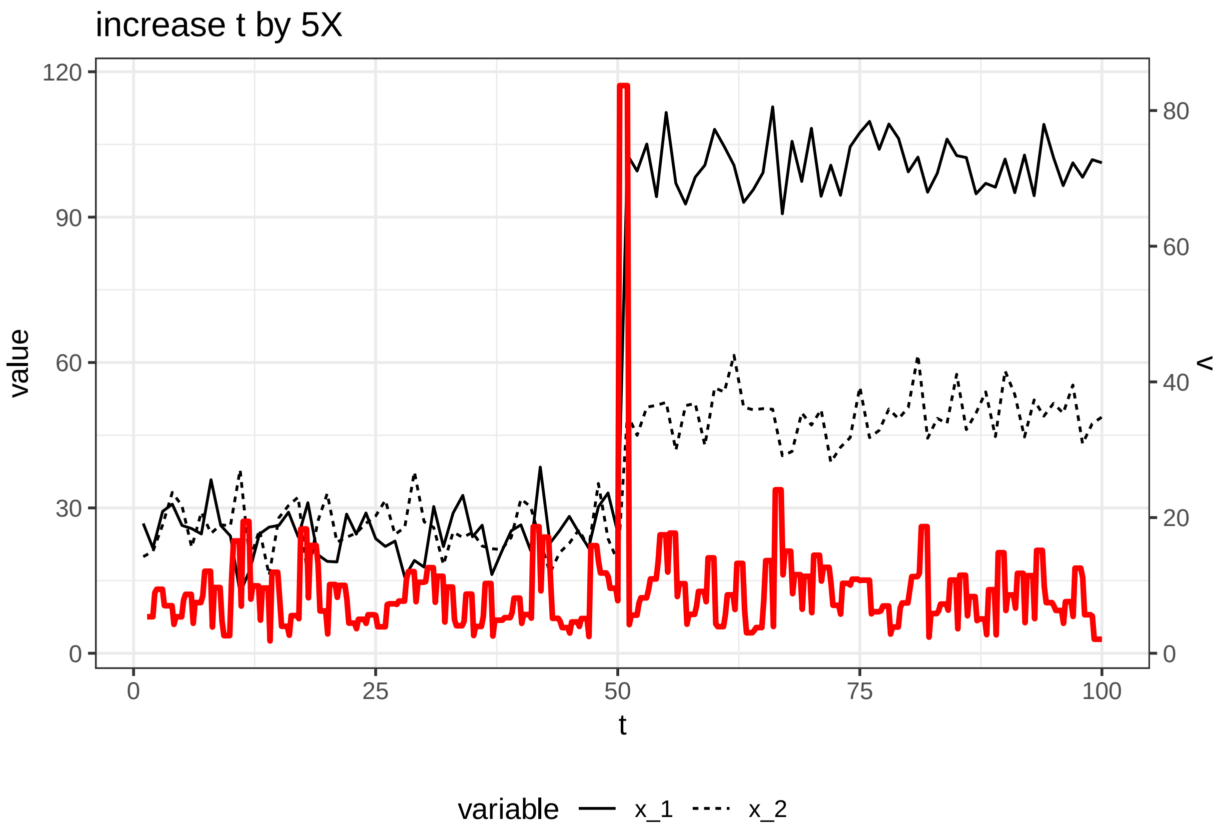
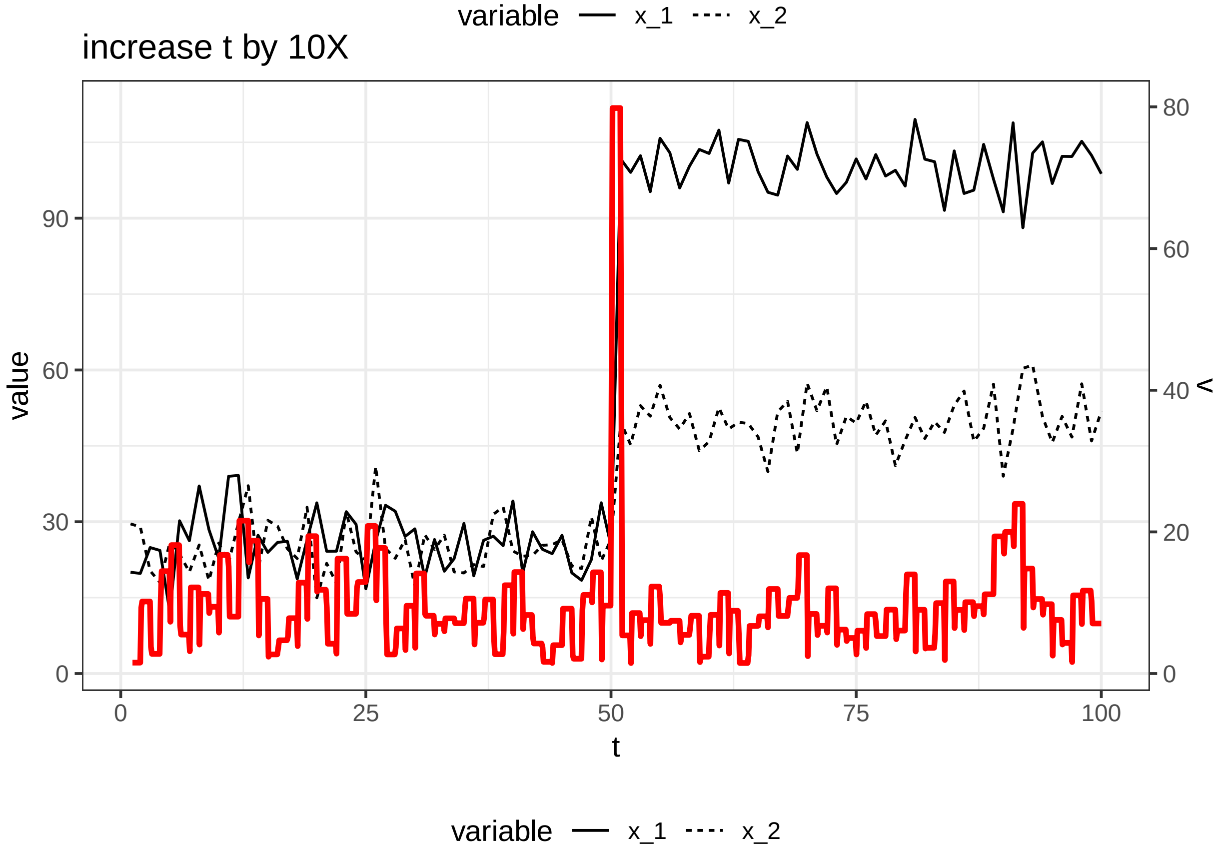


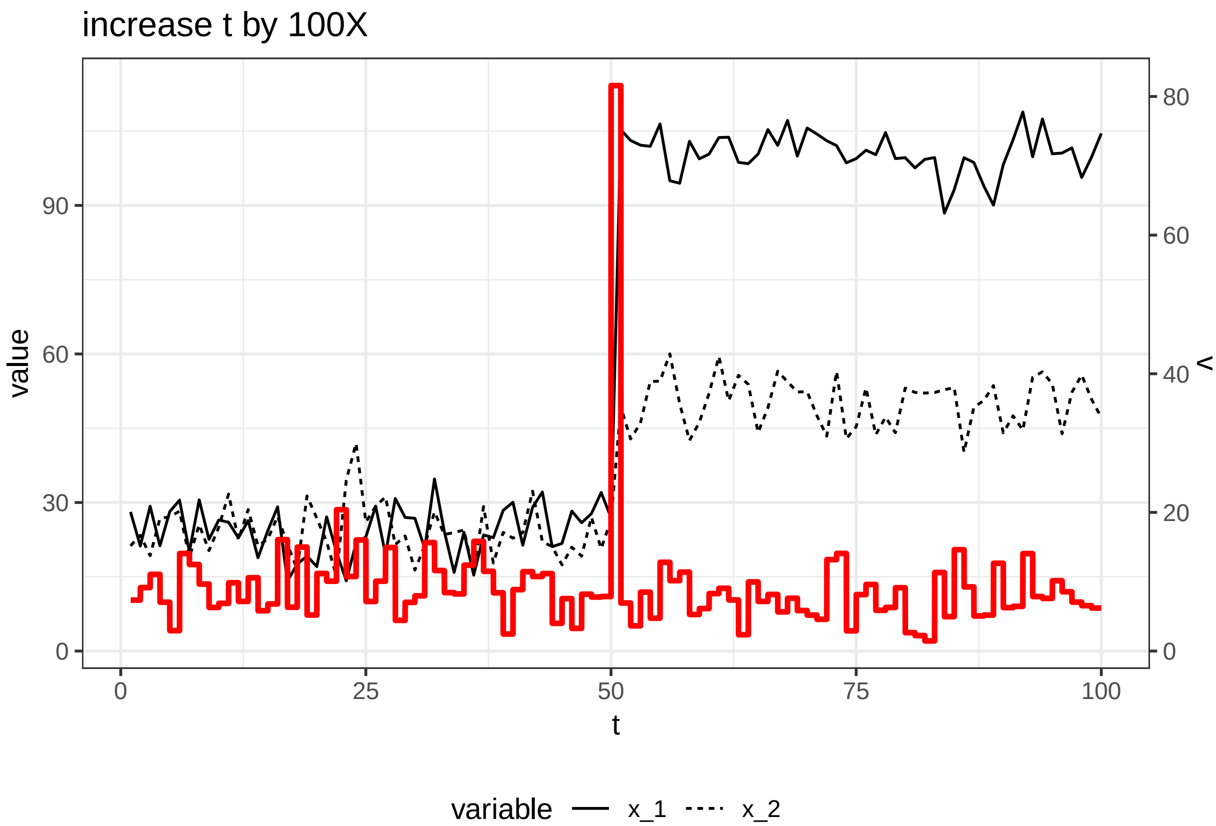
Figure 3: Average (± 2 SD) velocity (v) worsens as the variance of $\bar{x}_{2_{t=50(post)}}$ (post shift) increases. $\bar{x}_{1_{pre}} = 25$, $\bar{x}_{1_{post}} = 100$, $\bar{x}_{2_{pre}} = 25$, $\bar{x}_{2_{post}} = 50$, $\sigma_{1_{pre}} = 5$, $\sigma_{2_{pre,post}} = 5$


no smoothing
value
t
v
variable x_1 x_2


increase t by 2X
value
t
v
variable x_1 x_2


increase t by 5X
value
t
variable
x_1
x_2


increase t by 10X
value
t
variable
x_1
x_2



Performance of velocity using empirical data: paleodiatom community example

To gather baseline information on the use of velocity in empirical systems data, I calculated velocity for the paleodiatom system described in Chapter @ref(resampling) (see also Appendix @ref(appPaleo). Briefly, the paleodiatom community comprises 109 time series over a period of approximately 6936 years (Fig. @ref(fig:paleoTurnover)). As elaborated in @spanbauer_prolonged_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 points in the data (Fig. @ref(fig:paleoTurnover)). Using Fisher Information and climatological records, @spanbauer_prolonged_2014 suggest that regime shifts in this system at approximately 1,300 years before present (where present is equal to year 1950).

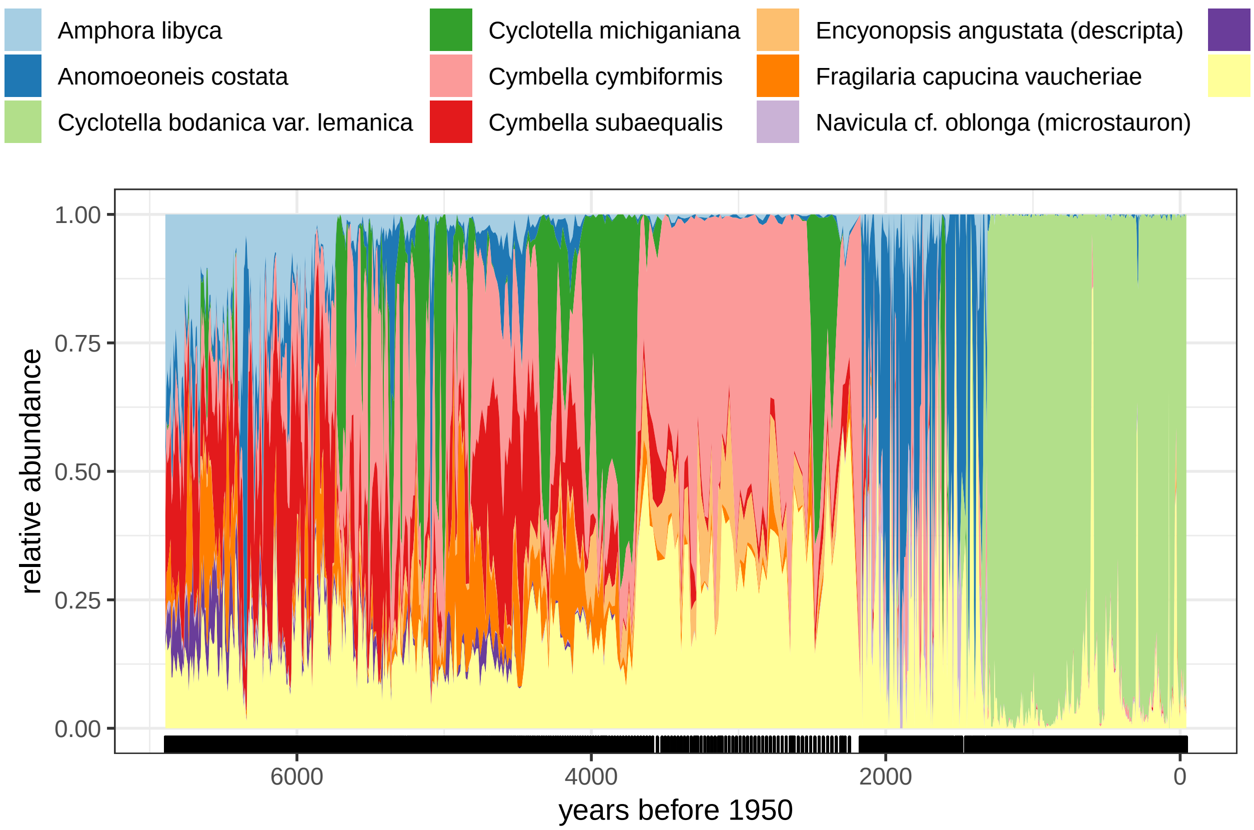
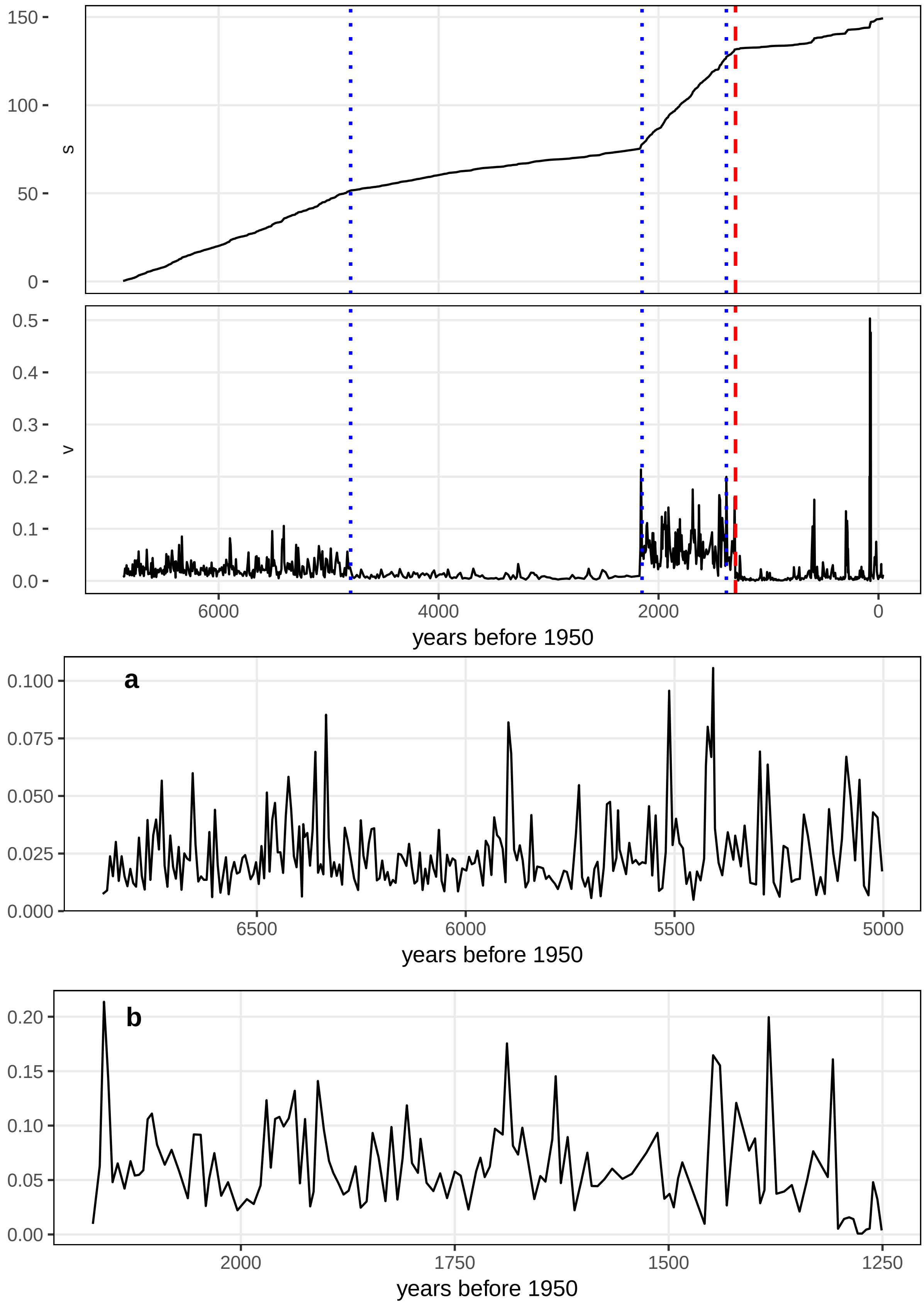


Figure 4: Relative abundances of the most common diatom species in the time series. Few species dominate the data over the entire time series, and turnover is apparent at multiple observations.


s
150
100
50
0
v
0.5
0.4
0.3
0.2
0.1
0.0
6000
4000
2000
0
years before 1950
a
0.100
0.075
0.050
0.025
0.000
6500
6000
5500
5000
years before 1950
b
0.20
0.15
0.10
0.05
0.00
2000
1750
1500
1250
years before 1950

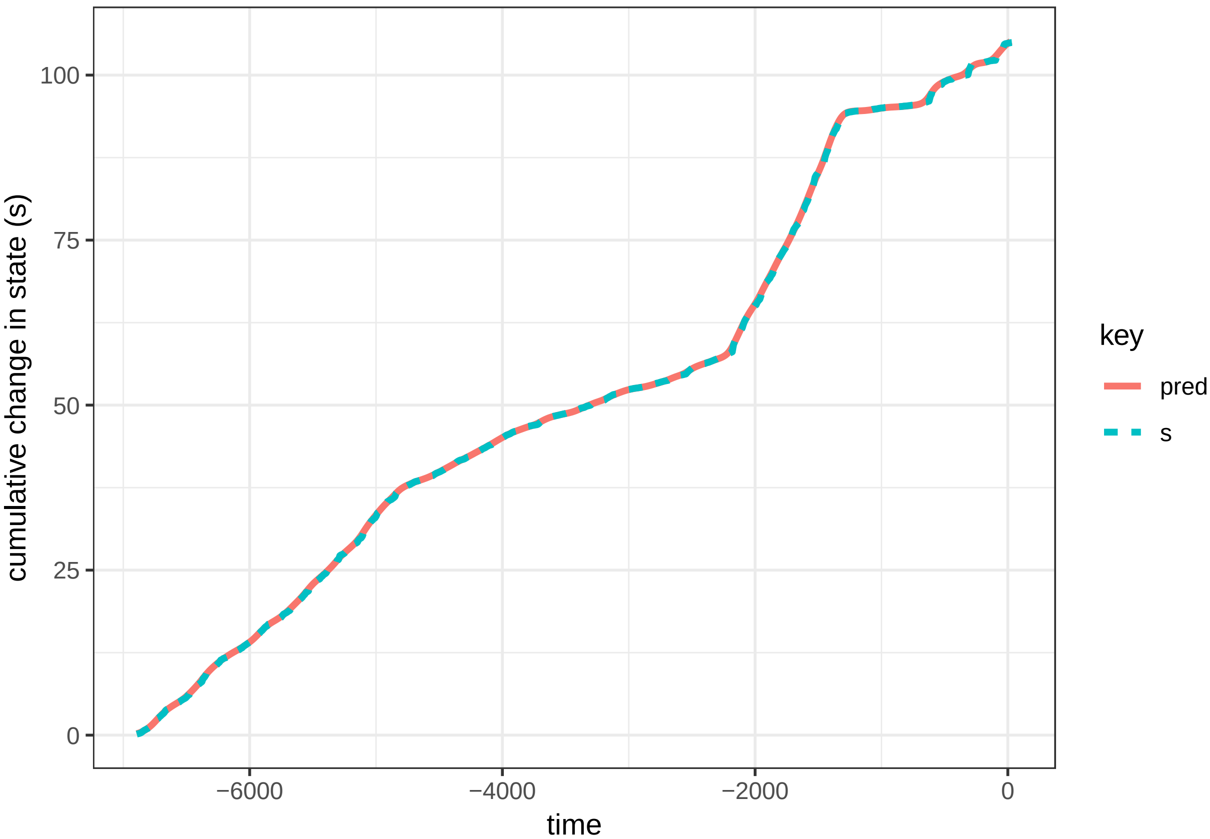


Figure 5: The regularized differentiation of s was best fit using $\alpha = 100$. Higher overlap of s and pred indicates a good fit of the regularized differentiated metric to the non-smoothed metric, s.

@spanbauer_prolonged_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. @ref(fig:paleoV)). The results of v and s (@ref(fig:paleoVelocity)) on the relative abundance data correspond with both the large shifts in species dynamics (see Fig @ref(fig:paleoTurnover), and also with the regime shift identified by @spanbauer_prolonged_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_prolonged_2014 (): 1. Two additional large shifts occurred at approximately 2,500, 4,800 and years before 1950
1. The periods before the first and after the second large shifts appear oscillatory (Fig. @ref(fig:paleoRegime1and3)).

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 $\alpha = 100$)[1].. This method of regularized differentiation is an ideal approach to smoothing s because it assumes the data are non-smooth, unlike other popular smoothing techniques e.g., Generalized Additive Models.

The smoothed velocity (@ref(fig:paleoV)) provides a similar but smoother picture of the velocity of the system trajectory. Comparing the smoothed (@ref(fig:paleoV)) to the non-smoothed velocity (@ref(fig:paleoVelocity)) yields similar inference regarding the location of the regime shifts at 2,200 and 1,300 years before present, but more clearly identifies the inter-regime dynamics (e.g., between 7,000 and 4,800 years before present).

[1]*We created the R-wrapper `tvdiff` as a Python wrapper for the tvdiff MatLab package @price2019tvdiff

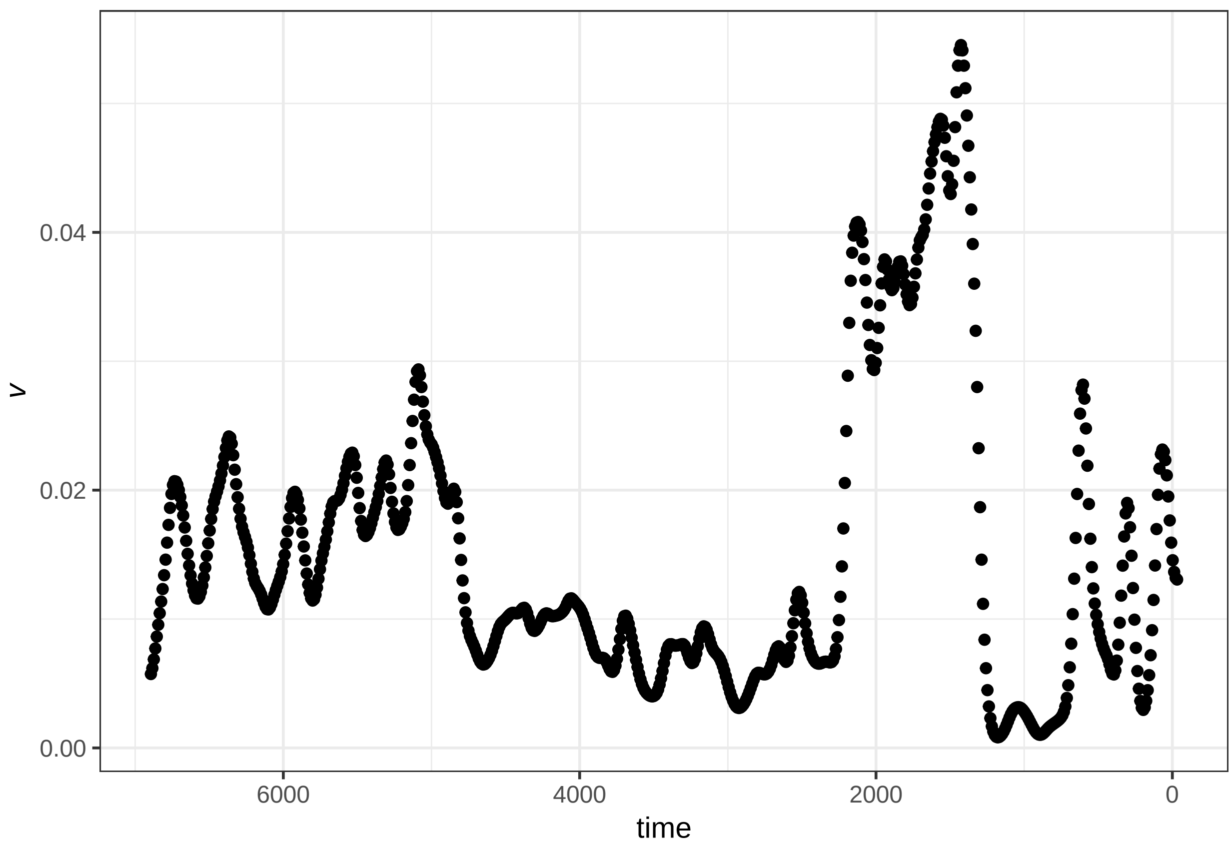


Figure 6: Need a cpation here!!!

Discussion

In this chapter, I described the steps for calculating a novel regime detection metric, system velocity (v). First described in @fath_regime_2003, v is used as a single step for caclulating a more complicated regime detection metric, Fisher Information (see also Chapter @ref(fiGuide)). System velocity is arguably simple to calculate, as shown in this chapter, 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. @ref(simVarPlot2)), and smooothing the original data does not reduce the noise surrounding this metric when variance is moderate (Fig. @ref(smoothV)).

Variance is a commonly-used indicator of ecological regime shifts (@brock_variance_2006), however, fails to perform when the number of variables is » a few. System velocity, v, may be useful in situations where the number of state variables is much greater than a few, and appears especially useful when the magnitude of change in one or more state variables is high (Fig. @ref(fig:simVplot2)). For example, this method will likely identify signals of regime shifts where the shift is defined as high species turnover within a community.

I tested the efficacy of this metric as an indicator of abrupt change in a two-variable system. Although a useful first step, this metric should be considered in a multi-species context, and particularly in community-level empirical data which is difficult to simulate. I demonstrate a compelling case study in materials associated with my R Package, **regimeDetectionMeasures**, and in Appendix @ref(appPaleo) in which multiple species turnover events are apparent in a paleodiatom community time series. In this case study, the 'distance travelled', s (Eq. @ref(eq:diffXsq2)), 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 the 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. @ref(fig:simVplot2).

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 clear shifts in variable means. For multispecies data you will typically need to reduce dimensionality before you can proceed with analyses, for example using some sort of ordination. 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 (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.

Supplementary Materials

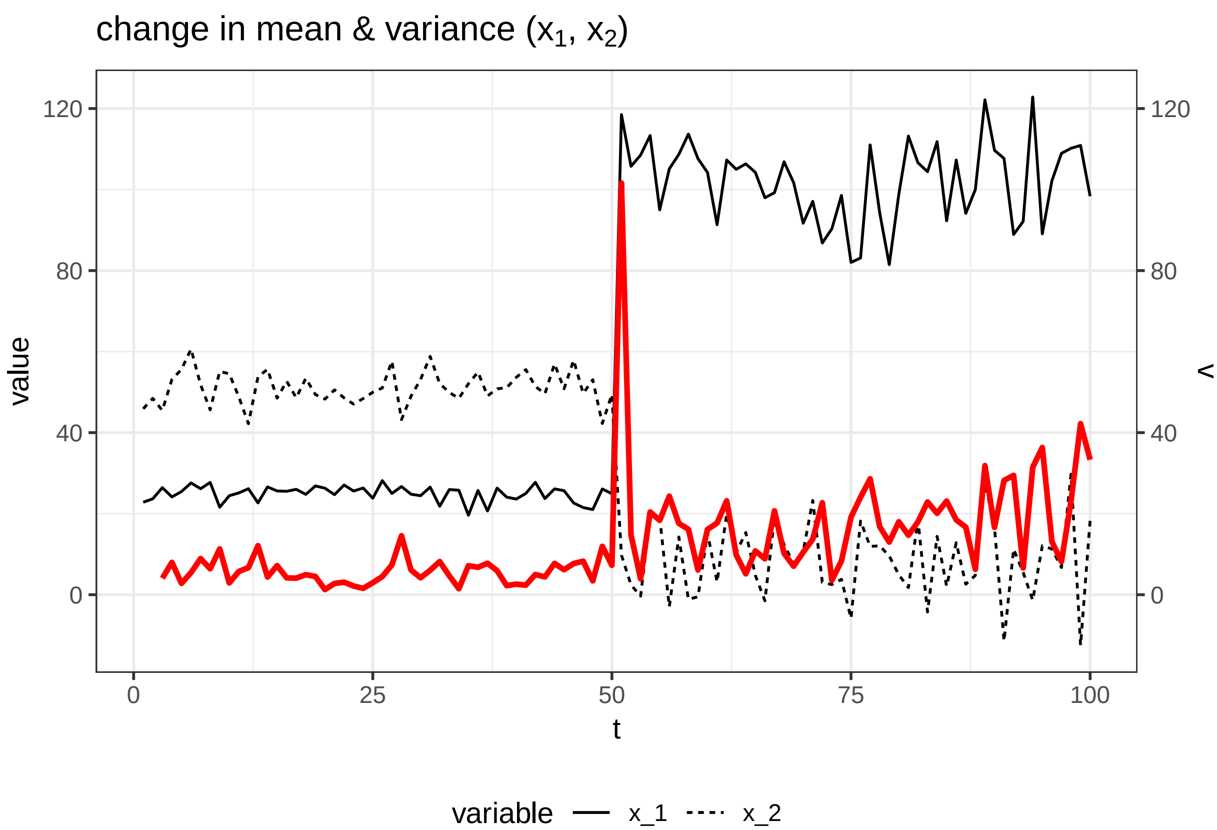


Figure 7: System change (s) and velocity (v) of the model system over the time period. Change in means ($\bar{x}_{1_{pre}} = 25$, $\bar{x}_{1_{post}} = 100$, $\bar{x}_{2_{pre}} = 50$, $\bar{x}_{2_{post}} = 10$) and an increase in variance ($\sigma_{1_{pre}} = 2$, $\sigma_{1_{post}} = 10$, $\sigma_{2_{pre}} = 5$, $\sigma_{2_{post}} = 10$).

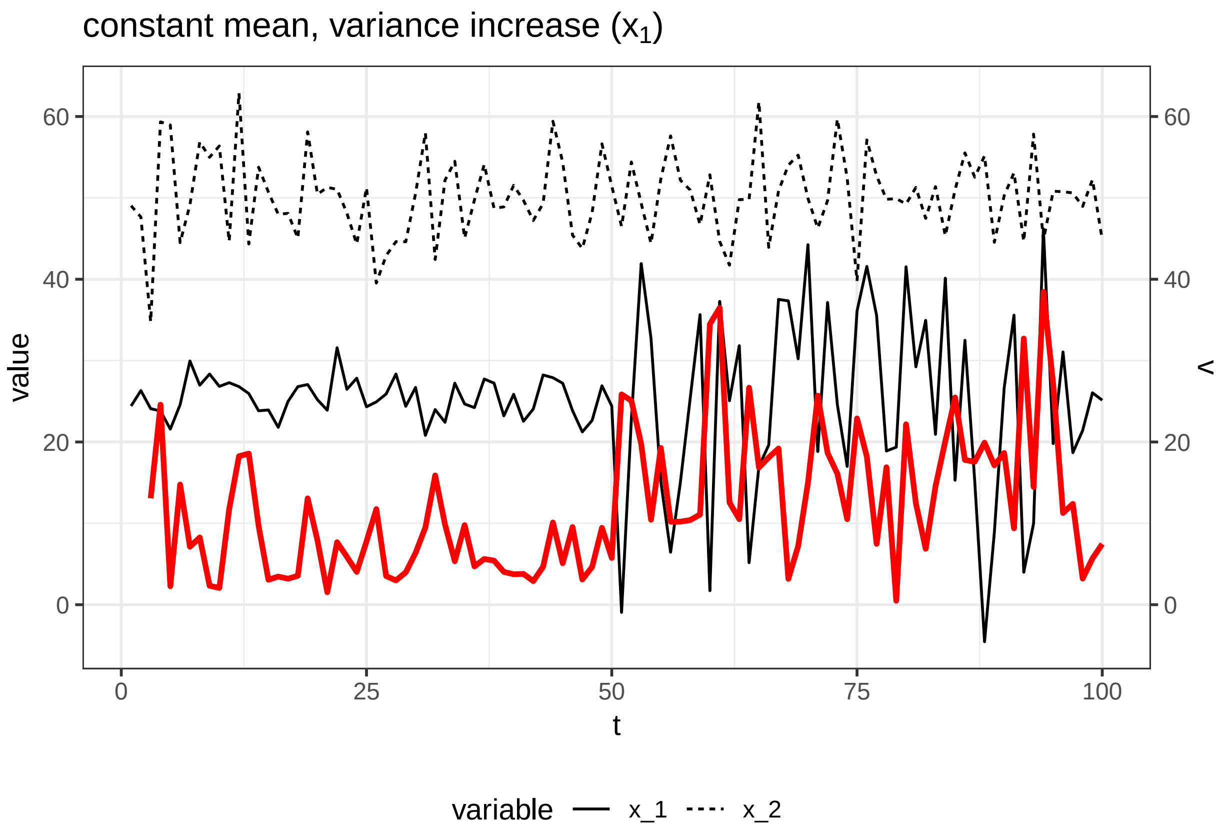


Figure 8: System change (s) and velocity (v) of the model system over the time period. Constant means ($\bar{x}_1 = 25$, $\bar{x}_2 = 50$) and sharp change in variance for one state variable $\sigma_{1_{pre}} = 2$, $\sigma_{1_{post}} = 12$, $\sigma_{2_{pre,post}} = 5$

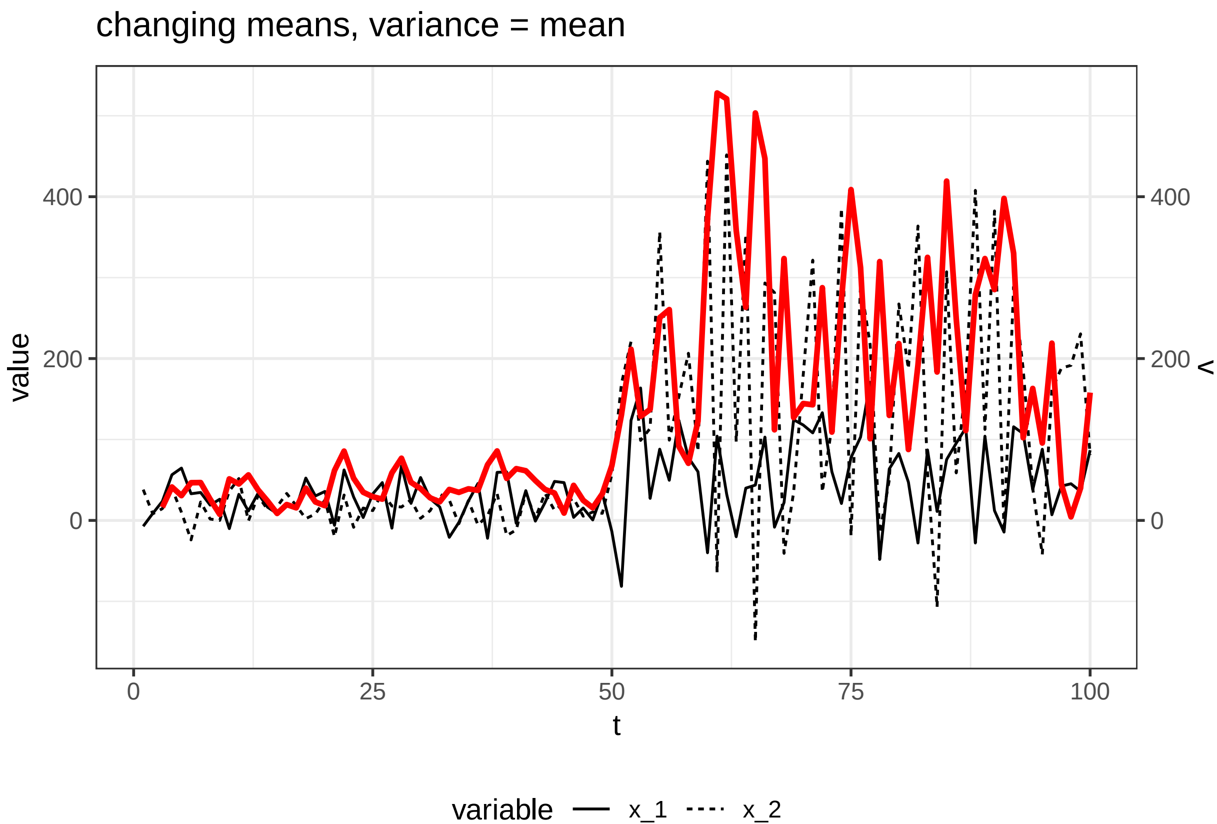


Figure 9: System change (s) and velocity (v) of the model system over the time period. Variance equal to mean ($/bar x_i = /sigma_i$), where means ($/bar x_{1_{pre}} = 25$, $/bar x_{1_{post}} = 50$, $/bar x_{2_{pre}} = 15$, $/bar x_{2_{post}} = 150$).