

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

## 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., ecosystem services, biodiversity, and human well-being. Quantitatively detecting and forecasting these changes, however, has yet to be accomplished for most ecological systems [Chapter \@ref(rdmReview); @ratajczak2018abrupt]. 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, @mumby2013evidence]. 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\_variance\_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\_leading\_2008], but are unreliable in other empirical systems [e.g., @perretti2012regime; @dutta2018robustness; @dakos2012robustness]. The last type of approach is the model-free approach [@dakos\_methods\_2012; Ch. \@ref(rdmReview)]. 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.

## Tracking ecosystem trajectory through time to explore system dynamics

```
knitr::include_graphics(here::here("chapterFiles/velocity/figsCalledInDiss/lorenz3D.png"))
knitr::include_graphics(here::here("chapterFiles/velocity/figsCalledInDiss/lorenz3D_timeseries.png"))
```

A classic example of state-switching by a system is demonstrated in the Lorenz ('butterfly') attractor [Fig. \@ref(fig:lorenz3D); @takens1981detecting]. This phase plot (Fig. \@ref(fig:lorenz3D)) provides an informative visual of the behavior of a chaotic system manifesting two attractors. Although the periodic, attractor behaviors are made clear when examining the time series of each dimension (Fig \@ref(fig:lorenz3Dts)), 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. @sugihara2012detecting 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; @sugihara2012detecting; @ye2015equation], 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 dimension reduction techniques and clustering algorithms (e.g., Principal Components Analysis, K-means clustering). In fact, this is true of many abrupt change and regime shift indicators.

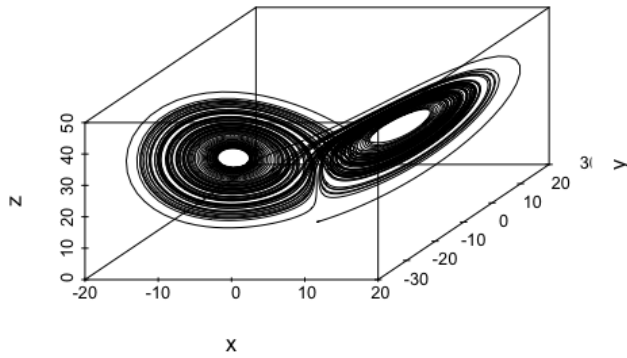


Figure 1: An example solution of the Lorenz ('butterfly') represented in 3-dimensional phase-space. Phase plots are typically used to visualize stable areas within a system's trajectory but reconstruction requires the difference models to be known and parameterized.

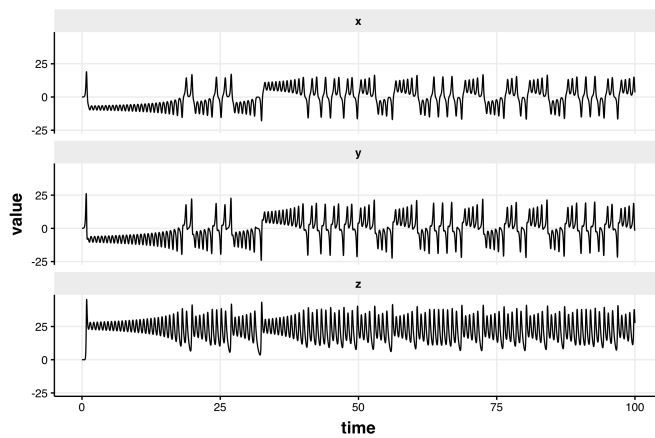


Figure 2: An example solution of the Lorenz ('butterfly') represented in individual system components.

## 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; @ref(fig:lorenz3D)] 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\_variance\_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  $\Delta$ ) is a term used for various measures which describe the relationship among 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  $\geq 0$  and **V** can take any value between  $-\infty$  and  $\infty$ . 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 sinusoidal fashion, where **V** is  $V > 0$  when ascending,  $V < 0$  when descending, and  $V \approx 0$  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.

## 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\_towards\_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 phase 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. @ref(fiGuide)], first introduced in @fath\_regime\_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.

## 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.

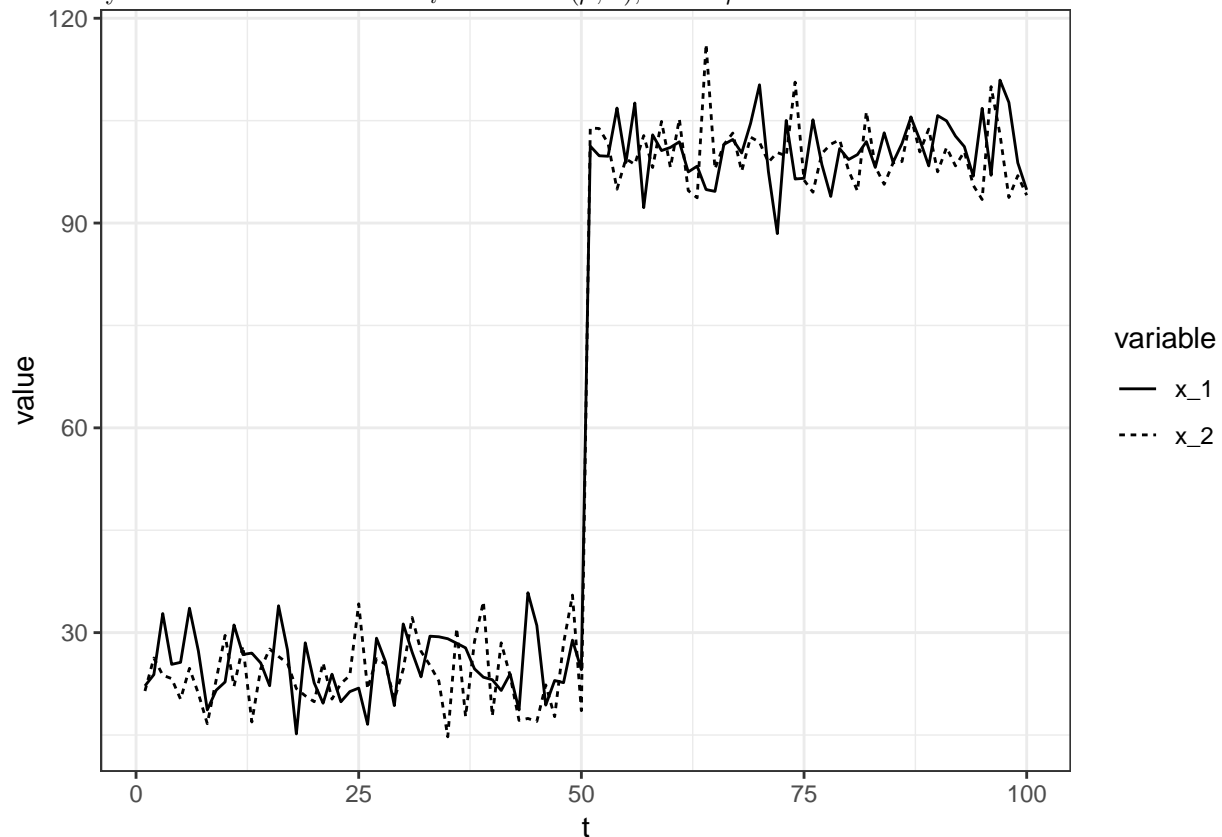
Next, I construct a second model system similar to the first, but one which exhibits a non-discontinuous rapid change in the state variables. The purpose of this section is three-fold. First, I demonstrate how velocity behaves when the system undergoes varying degrees of change (e.g., slow change versus nearly discontinuous, rapid). Second, I concurrently identify baseline expectations of velocity under varying conditions of mean and variability of the state variables 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.

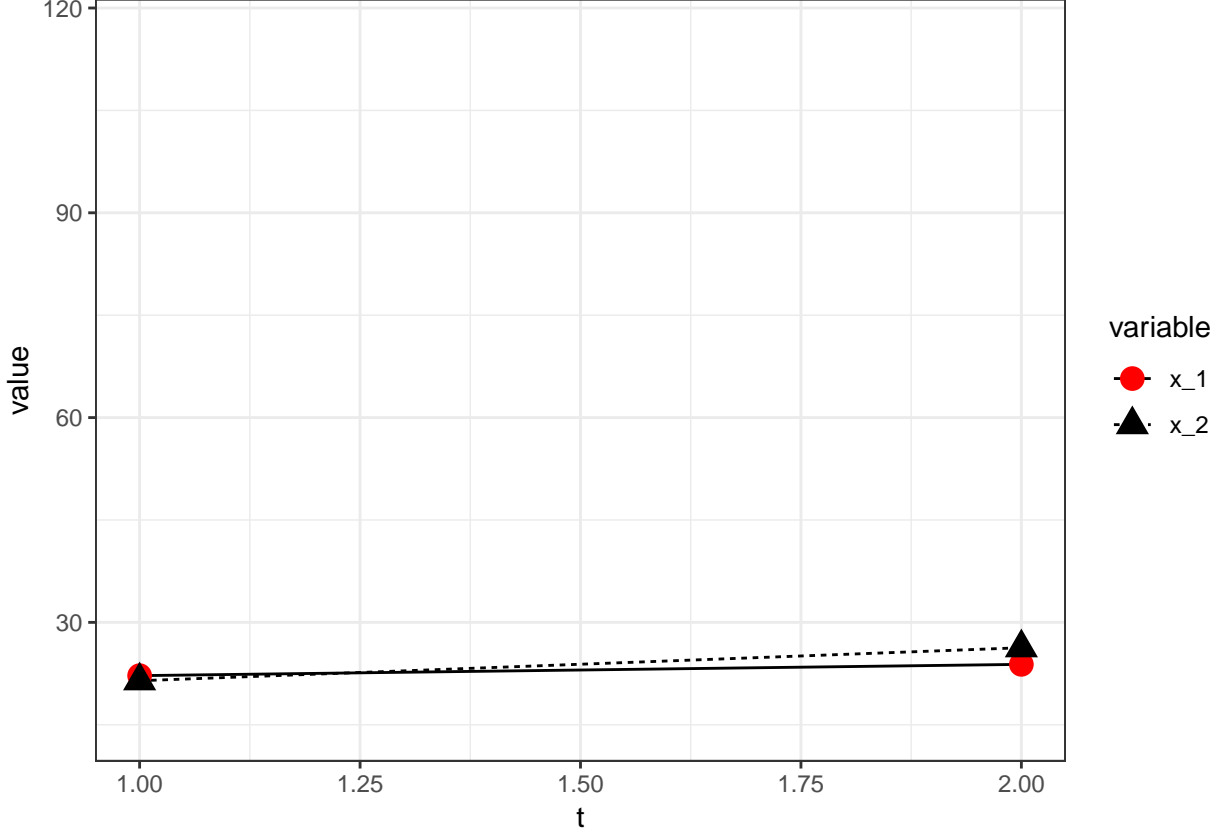
Finally, I calculate the velocity of an empirical, paleolitic freshwater diatom community time series to demonstrate the utility of the velocity metric in highly noisy, high dimensional, and irregularly-sampled data.

## Calculating velocity, $v$ using a simple toy system

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  $x_i \sim \text{Normal}(\mu, \sigma)$ , where  $\mu$  is the mean and  $\sigma$  is the standard deviation.



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,  $\{x_1 \dots x_N\}$ , 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  $t = 1, 2, 3, \dots, 100$ . 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. @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  $\sigma_i$  (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



$\sigma_i$ .

### Steps for calculating  $v$  ### Step 1: Calculate  $\Delta x_i$  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. @ref(fig:sysEx2); Eq. @ref(eq:diffX)]:

$$\Delta x_i = x_{i(t+1)} - x_{it} \text{ (\#eq : diffX)} \quad (1)$$

Note that  $\Delta x_i$  can take any value between  $-\infty$  and  $\infty$ .

### Step 2: Calculate distance travelled, $s$

Next, we calculate the total change in the system between two points in time as  $\Delta s$ . This is achieved by taking this square root of the sum of squares of the changes in all state variables per Pythagora's theorem (Eq. @ref(eq:ds)):

$$\Delta s = \sqrt{\sum \Delta x_i^2} \text{ (\#eq : ds)} \quad (2)$$

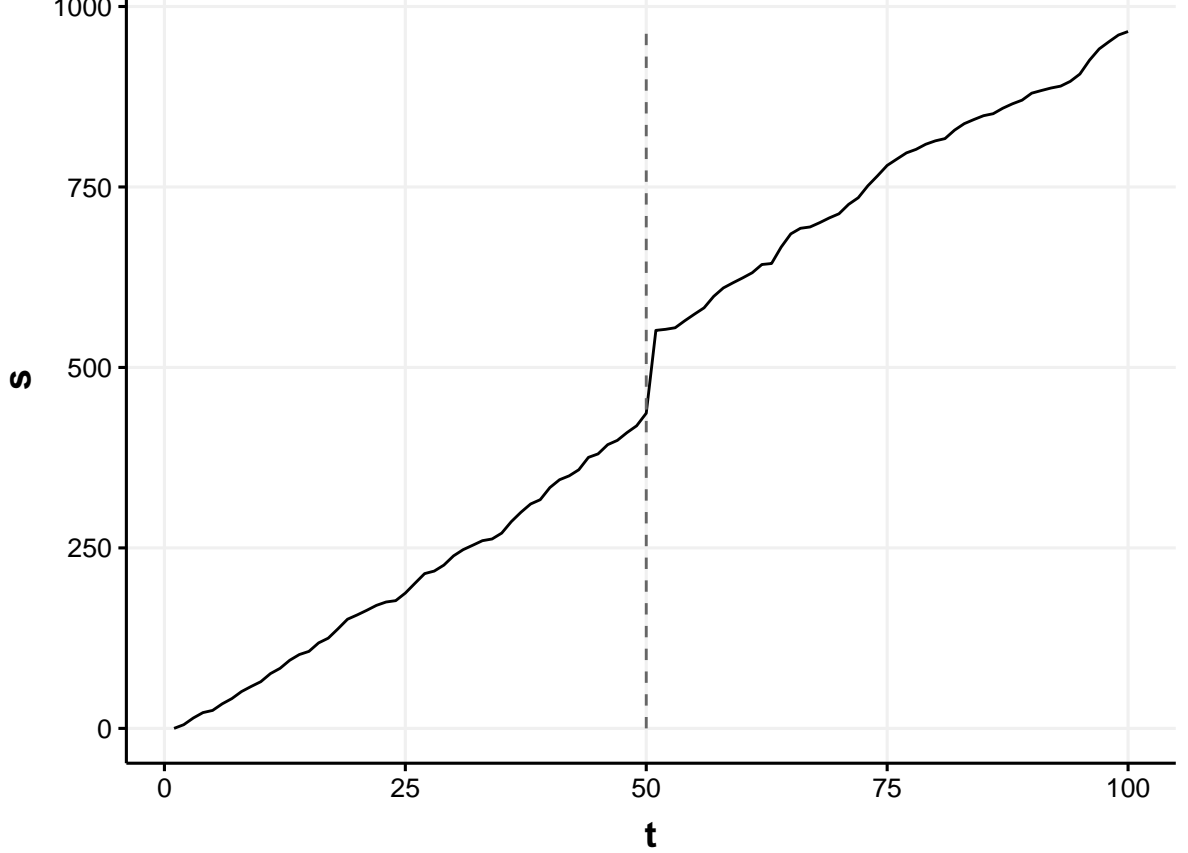
Although  $\Delta s$  represents the absolute change in the system between consecutive points in time, the measure is not yet relative to the trajectory. Finally, because we are interested in identifying relative changes along the system trajectory, we calculate the distance travelled,  $s$ , as the cumulative sum of  $\Delta s$  (Eq. @ref(eq:ds)) over some time interval,  $(0, T)$  (Eq. @ref(eq:s)):

$$s_T = \sum_{t=0}^T \Delta s \text{ (\#eq : s)} \quad (3)$$

We now have a single measure,  $s$  (Eq. @ref(eq:s)), for each pair of time points in our  $N$ -dimensional system (Fig. @ref(fig:sysExs)). It should be noted that  $s$  (Fig. @ref(fig:sysExs)) is monotonically increasing since the value of  $\Delta s$  (Eq. @ref(eq:ds)) is a sum of 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,

Table 1: Steps outlined for calculating system velocity,  $v$ , using the 2-variable toy data as an example.

$t$	$x_1$	$x_2$	$\Delta x_1$	$\Delta x_2$	$\Delta t$	$\sqrt{(\sum_{i=1}^N \Delta x_i^2)}$	$s$	$v$
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



$x_i$ .

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

$$v = \frac{s_{t+1} - s_t}{\Delta t} (\#eq : velocity) \quad (4)$$

The numerical results for each step in the calculation of velocity [Eq. @ref(eq:velocity)] is 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

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 parameters at the regime shift:  $\bar{x}_1$ , increased the mean value of  $x_1$   $\sigma_1$ , change in variance of  $x_1$ . Simulations consisted of 10,000 random samples drawn from the normal distribution for each parameter, I randomly drew the toy system samples 10,000 times under increasing values of  $\bar{x}_1$  and  $\sigma_1$ . To identify patterns in the influence of parameter values on velocity, I present the mean values of  $v$  across all

changing means, constant variance

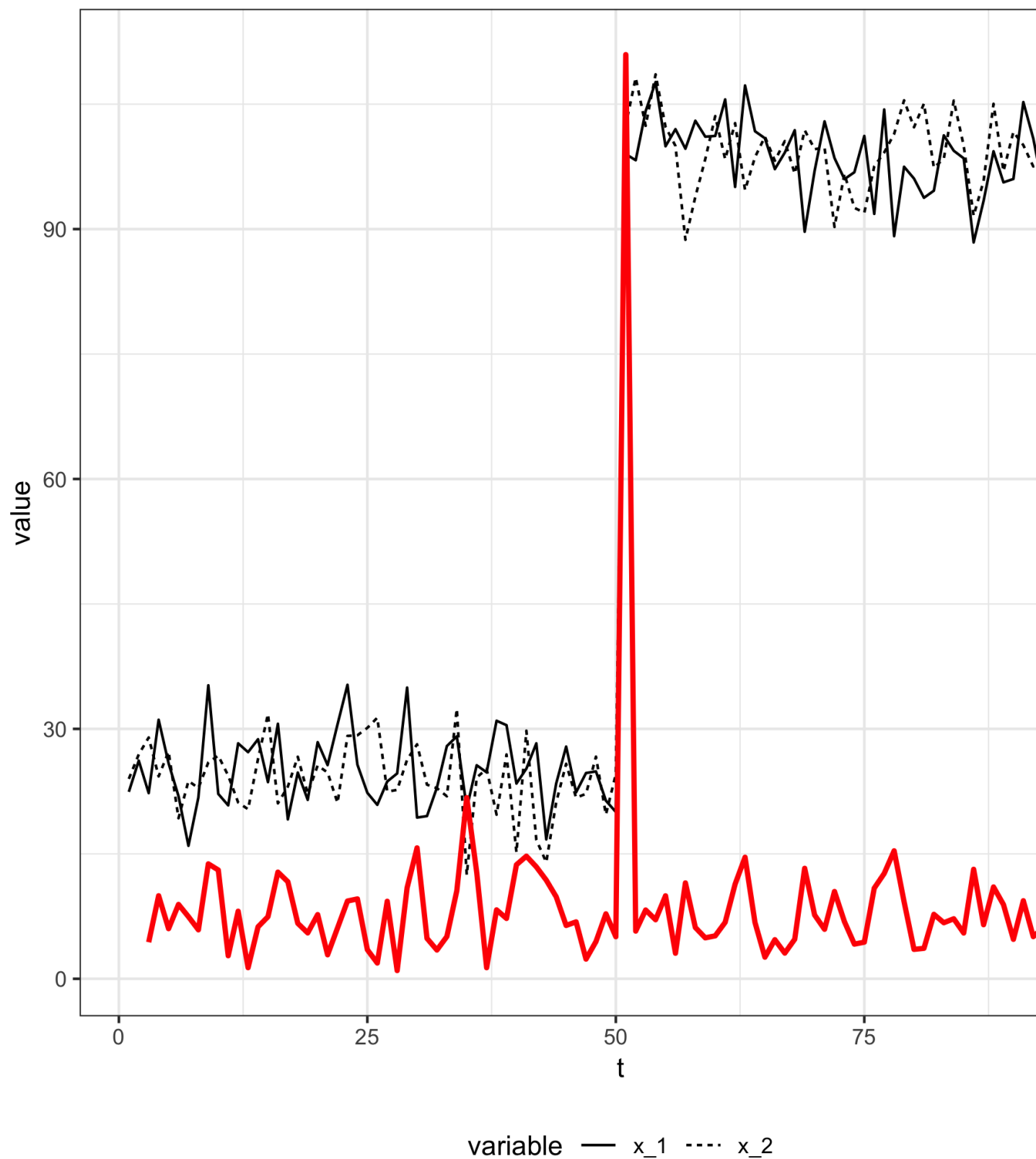


Figure 3: System change ( $s$ ) and velocity ( $v$ ) of the  $\tilde{\gamma}$  model system over the time period. Constant means ( $\bar{x}_{pre} = 25$ ,  $\bar{x}_{post} = 10$ ) and sharp change in variance for both state variables,  $\sigma = 5$ .

simulations, with confidence intervals of  $\pm 2$  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  $\sigma_i$  (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  $W = \{25, 30, 35, \dots, 100\}$  (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 system 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,  $\sigma_1$  in the set  $W = \{1, 2, 3, \dots, 25\}$ . The variance for both state variables prior to the regime shift,  $\sigma_1$  and  $\sigma_2$ , was 5, with the change occurring in  $\sigma_{1_{post}}$ . System velocity  $v$  appears sensitive 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  $\sigma_{post}$  (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)).

## Velocity performance in a toy system with smooth transitions

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 this type of rapid change, velocity will approach infinity regardless of the effect size. In other words, if the system exhibits turnover in e.g. 25% of the state variables, the value of velocity will be equal to that of turnover in e.g. 75% of the variables. Removing the possibility of infinite values provides more relative measures within the community time series.

### 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 explored six conditions (Table @ref(tab:sysParams))



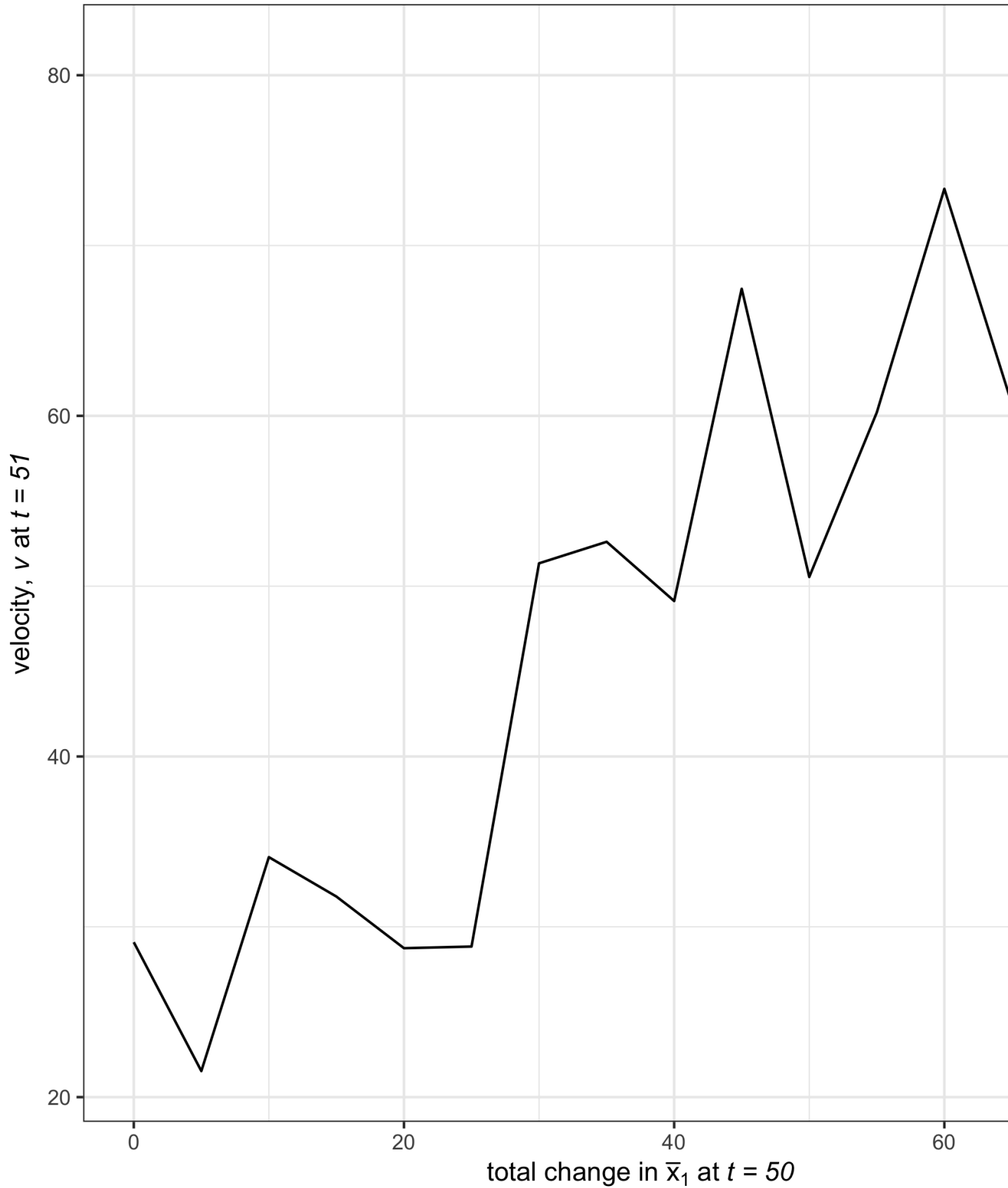


Figure 4: Velocity ( $v$ ) generally increases as the total change in the mean value of  $\bar{x}_{1_{t=50}}$  increases in a single iteration of our toy system ( $N_{iter} = 1$ , seed = 123). This 2-variable system exhibits a regime shift at  $t = 50$ , where variance is constant  $\sigma = 5$ ,  $\bar{x}_1 = 25$  when  $t < 50$ ,  $\bar{x}_2 = 50$  when  $t \geq 50$ ,  $\bar{x}_1 = 25$  when  $t < 50$ .

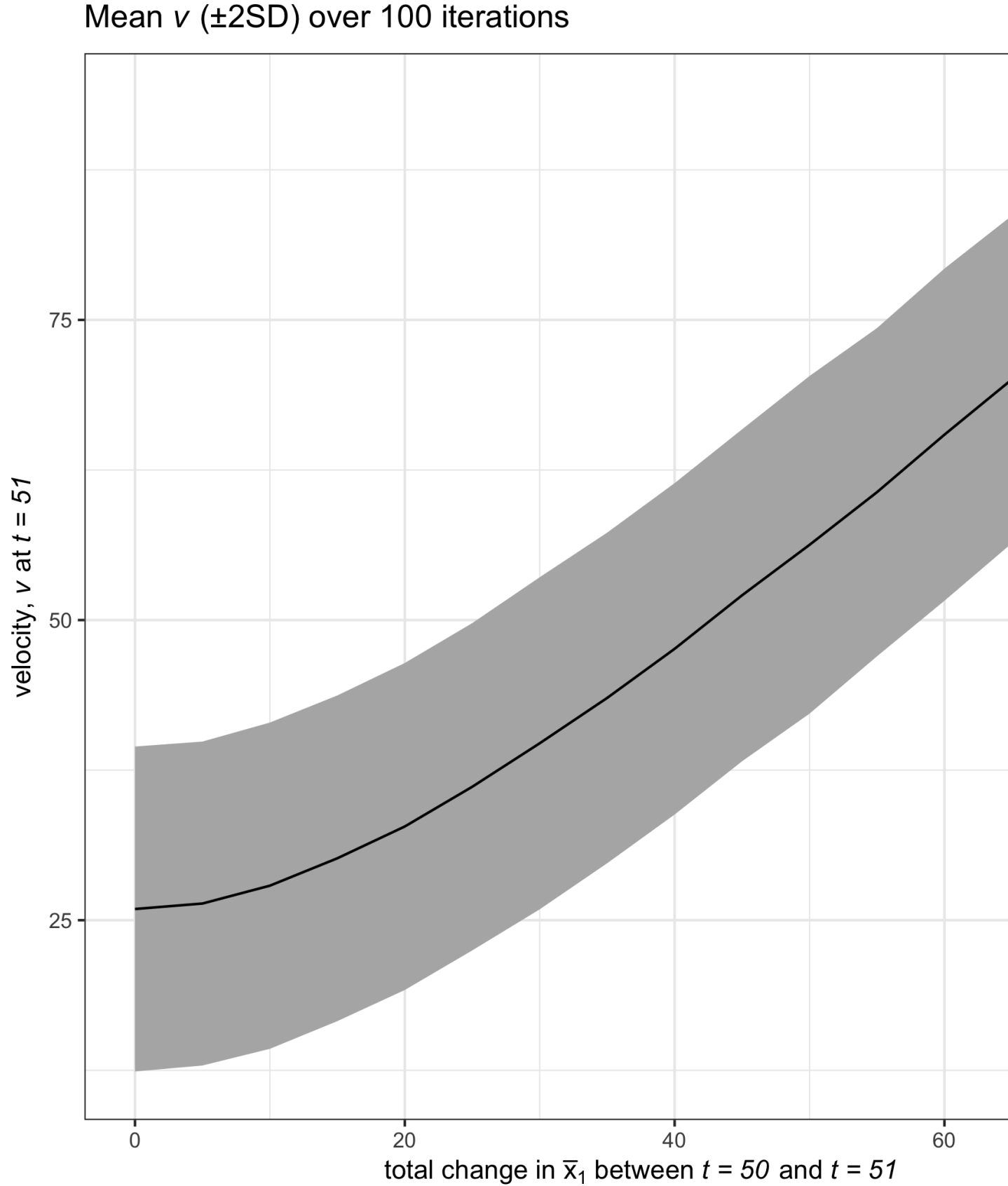


Figure 5: Change in velocity ( $v$ ) as the total change<sup>10</sup> in the mean value of  $\bar{x}_{2,t=50}$  over 10,000 simulations. A regime shift was induced at  $t = 50$  with constant variance  $\sigma = 5$ ,  $\bar{x}_2 = 25$  when  $t < 50$ , and changes in variable mean values,  $\bar{x}_2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

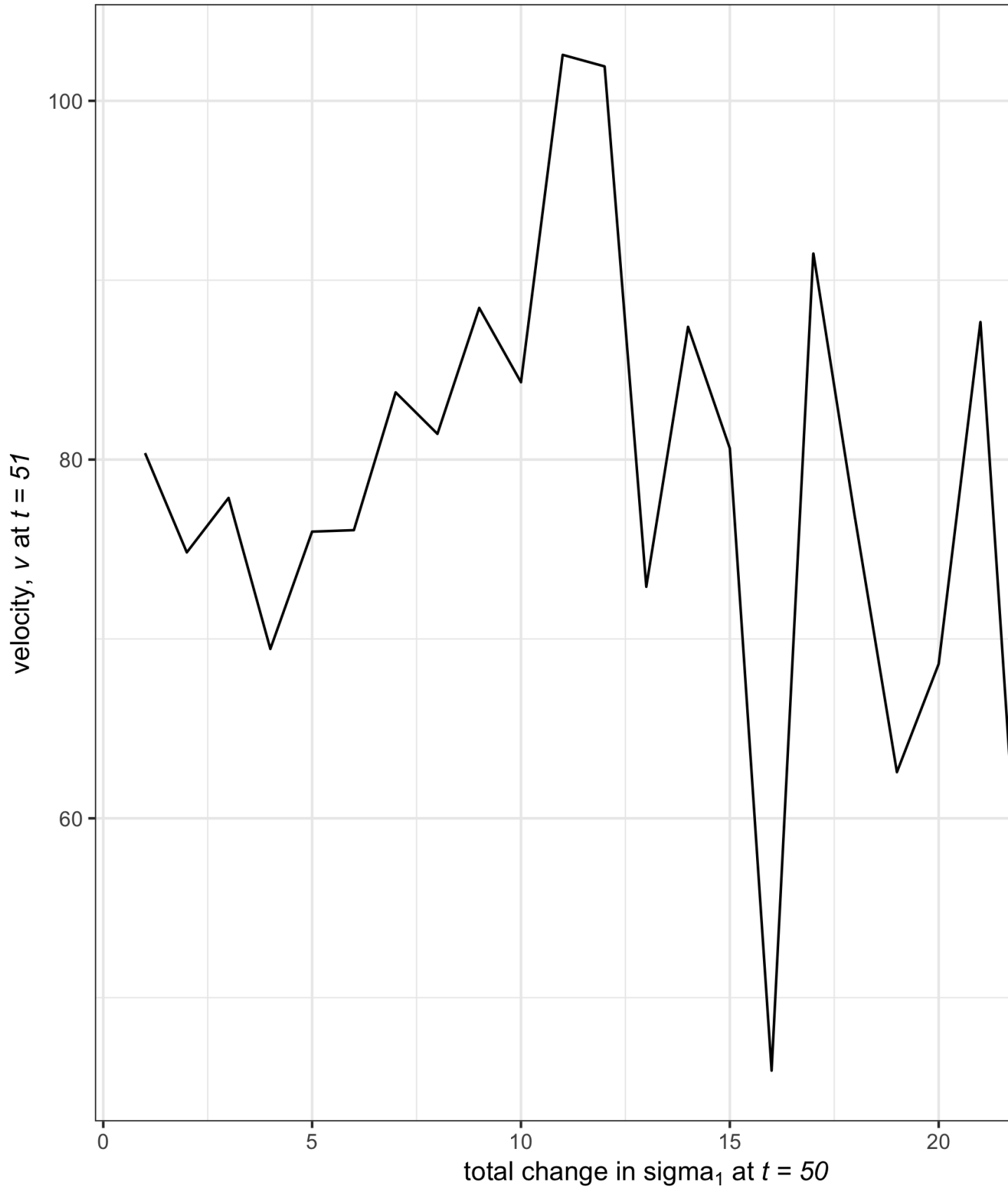


Figure 6: High variance of velocity ( $v$ ) in a single iteration ( $N_{iter} = 1$ , seed = 123) of simulations as we increase  $\sigma_1$  at  $t = 50$ .

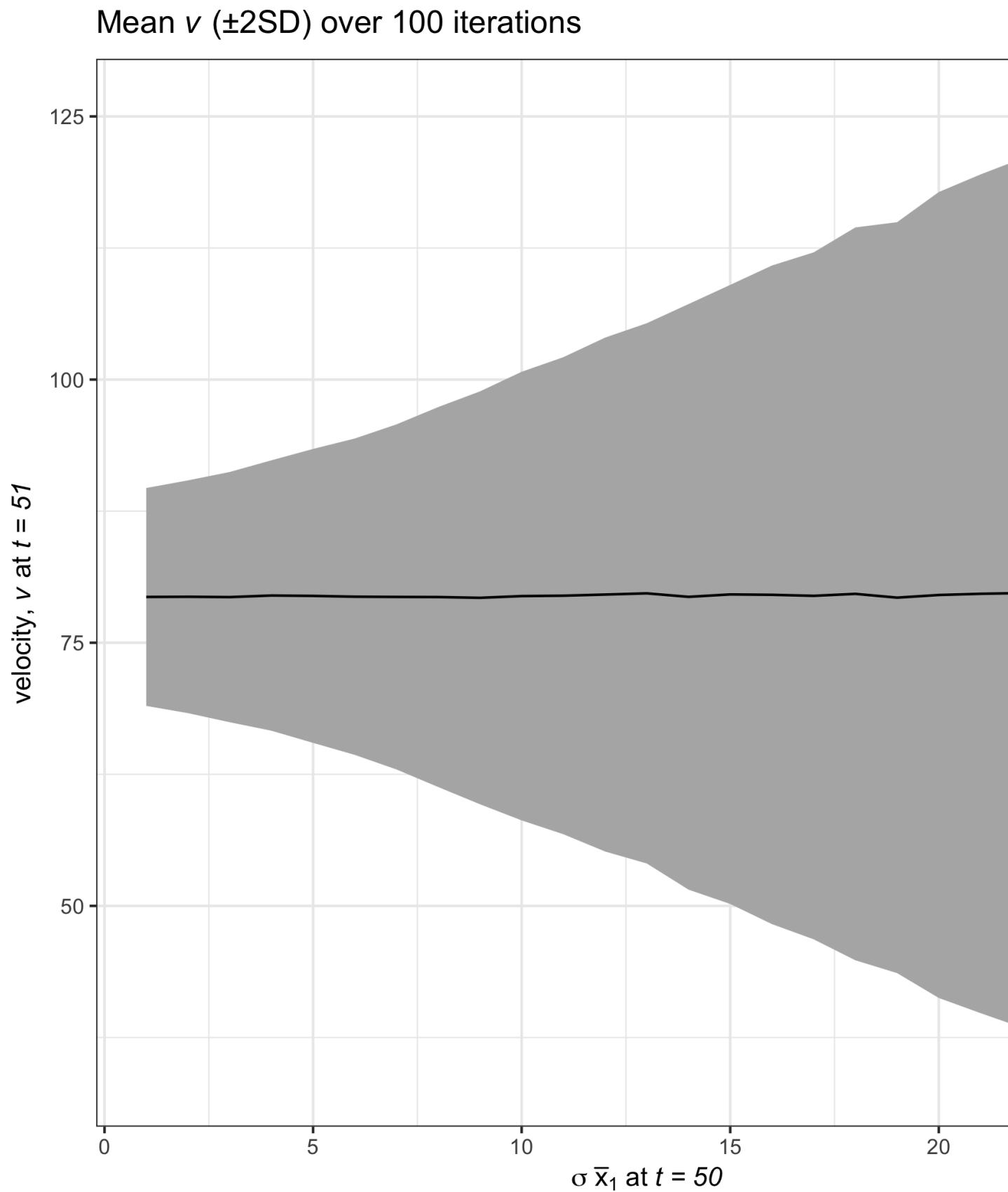


Figure 7: Average ( $\pm 2$  SD) velocity ( $v$ ) worsens as the variance of  $\bar{x}_{t=50(post)}$  (post shift) increases.  $\bar{x}_{1_{pre}} = 25$ ,  $\bar{x}_{1_{post}} = 100$ ,  $\bar{x}_{2_{pre}} = 25$ ,  $\bar{x}_{2_{post}} = 50$ ,  $\sigma_{1_{pre}} = 5$ ,  $\sigma_{2_{pre,post}} = 5$

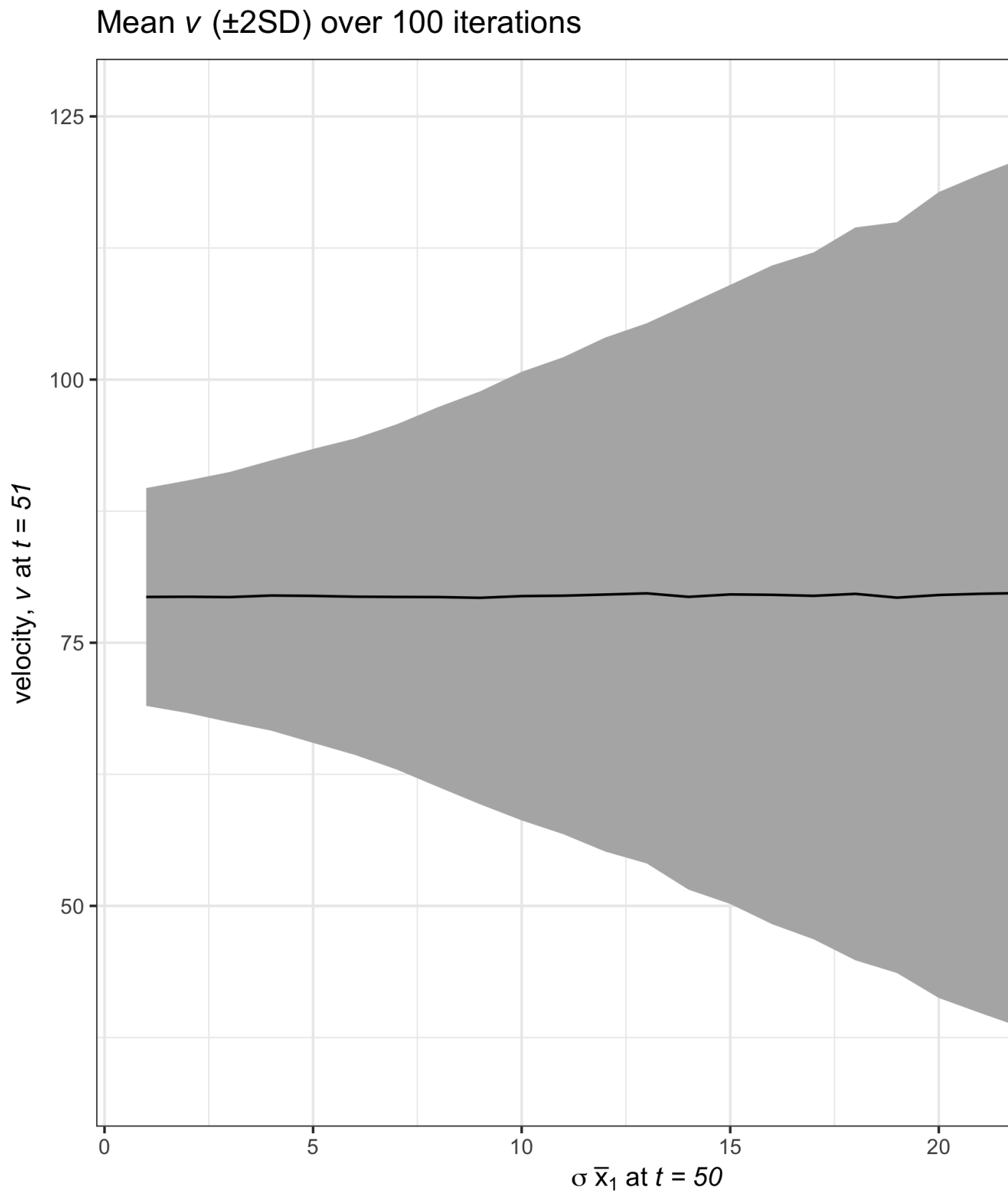


Figure 8: The noise in system velocity ( $v$ ) is not obviously reduced in this system as the original data ( $x_1$ ,  $x_2$ ) is increasingly smoothed.

Table 2: Conditions for generating various scenarios of the hyperbolic tangent-induced abrupt change.  $\sigma_i$  represents the standard deviation of  $\mu_{x_i}$  as the percent of  $\mu_{x_i}$ ,  $\mu_{x_i}$  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	$\sigma_{x_{1pre}}$	$\sigma_{x_{1post}}$	$\sigma_{x_{2pre}}$	$\sigma_{x_{2post}}$	$\mu_{x_{1pre}}$	$\mu_{x_{1post}}$	$\mu_{x_{2pre}}$	$\mu_{x_{2post}}$
$\mu_{x_1}, \mu_{x_2}, \sigma_{x_1}, \sigma_{x_2}$	0.05	0.10	0.05	0.10	10	55	15	44
$\mu_{x_1}, \sigma_{x_1}$	0.05	0.10	0.05	0.05	10	55	15	15
$\mu_{x_1}, \mu_{x_2}$	0.05	0.05	0.05	0.05	10	55	15	44
$\mu_{x_1}$	0.05	0.05	0.05	0.05	10	55	15	15
$\sigma_{x_1}, \sigma_{x_2}$	0.05	0.10	0.05	0.10	10	10	15	15
$\sigma_{x_1}$	0.05	0.10	0.05	0.05	10	10	15	15

## Application to empirical data: 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 @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 abundances 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).

@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  $iter = 2000$ ,  $scale = small$ ,  $ep = 1 \times 10^{-6}$ , and  $\alpha = 100$ )<sup>1</sup>.. 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 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, 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 system [@spanbauer\_prolonged\_2014].

## 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 calculating a more complicated regime detection metric, Fisher Information (see also Chapter @ref(fiGuide)). System velocity is arguably simple to calculate,

<sup>1</sup>\*We created the R-wrapper `tvdiff` as a Python wrapper for the `tvdiff` MatLab package @price2019tvdiff

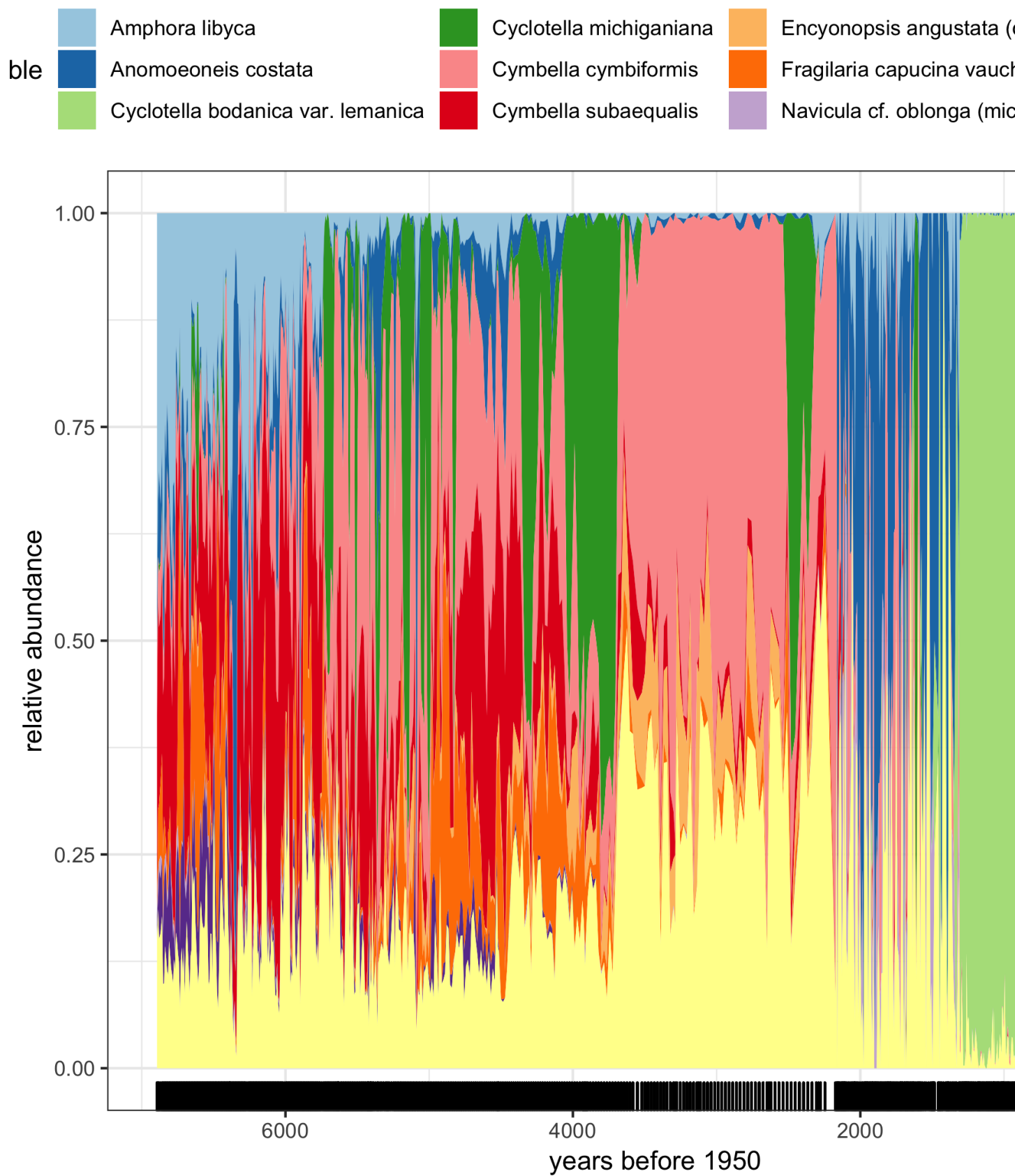


Figure 9: Relative abundances of the most common diatom species in the time series. Few species dominate the data over the entire time series, and turnover is apparent at multiple observations.

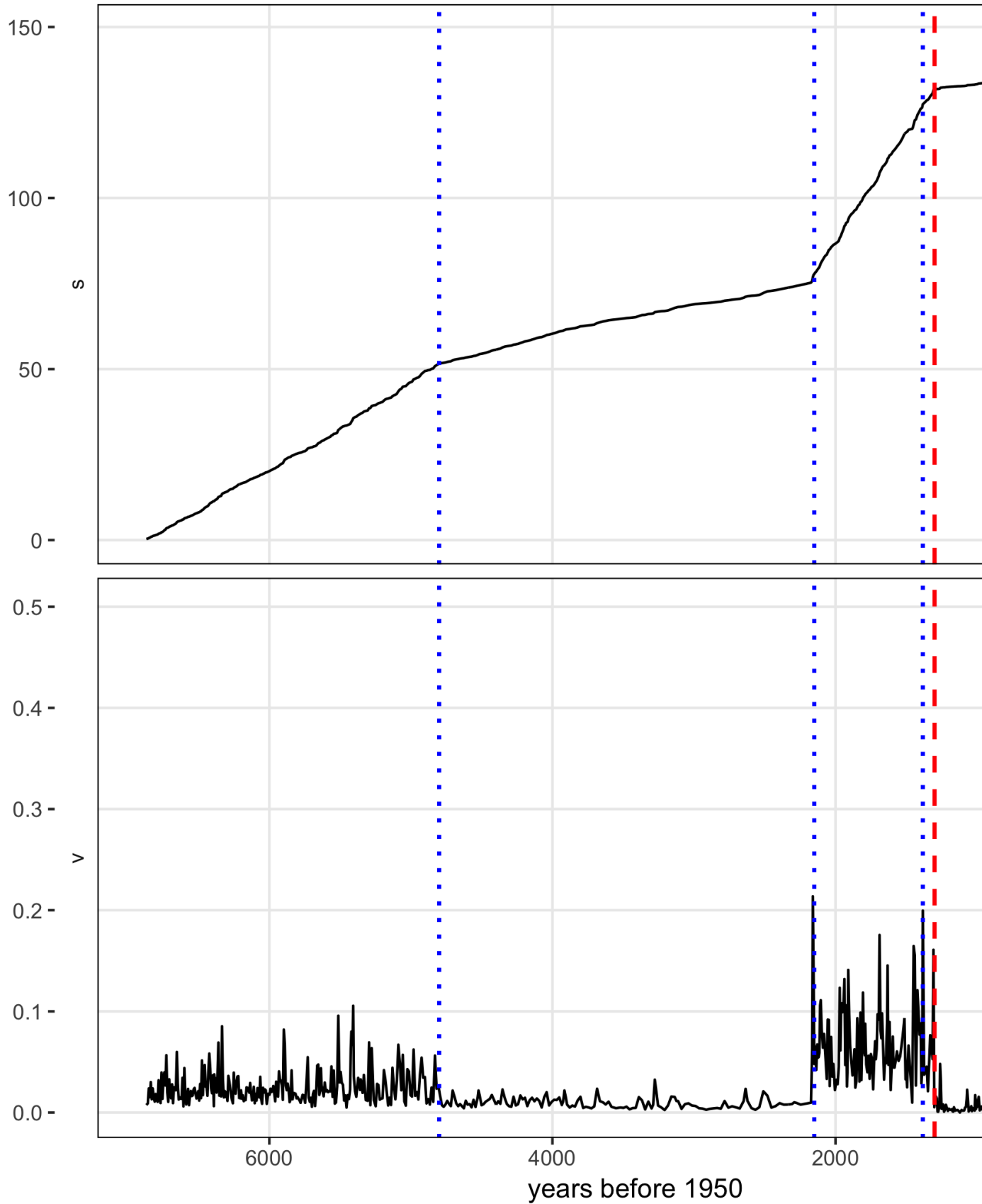


Figure 10: Velocity  $v$  and distance travelled  $s$  of the <sup>16</sup>paleodiatom time series. Dashed line at 1,300 years before 1950 indicates the regime shift identified in Spanbauer et al. (2014). Dotted lines indicate regime shifts as visually identified on metrics  $s$  and  $v$ .



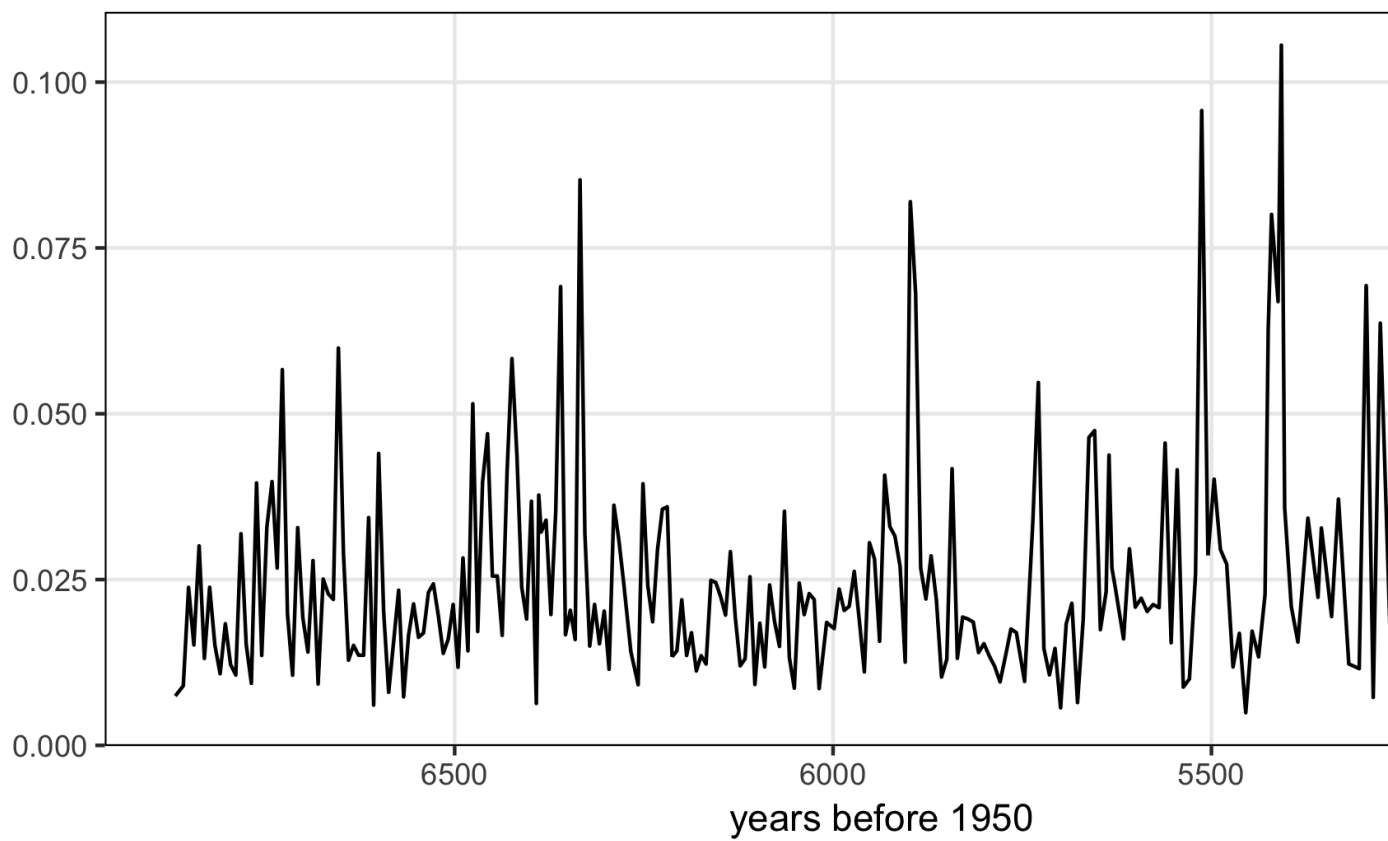
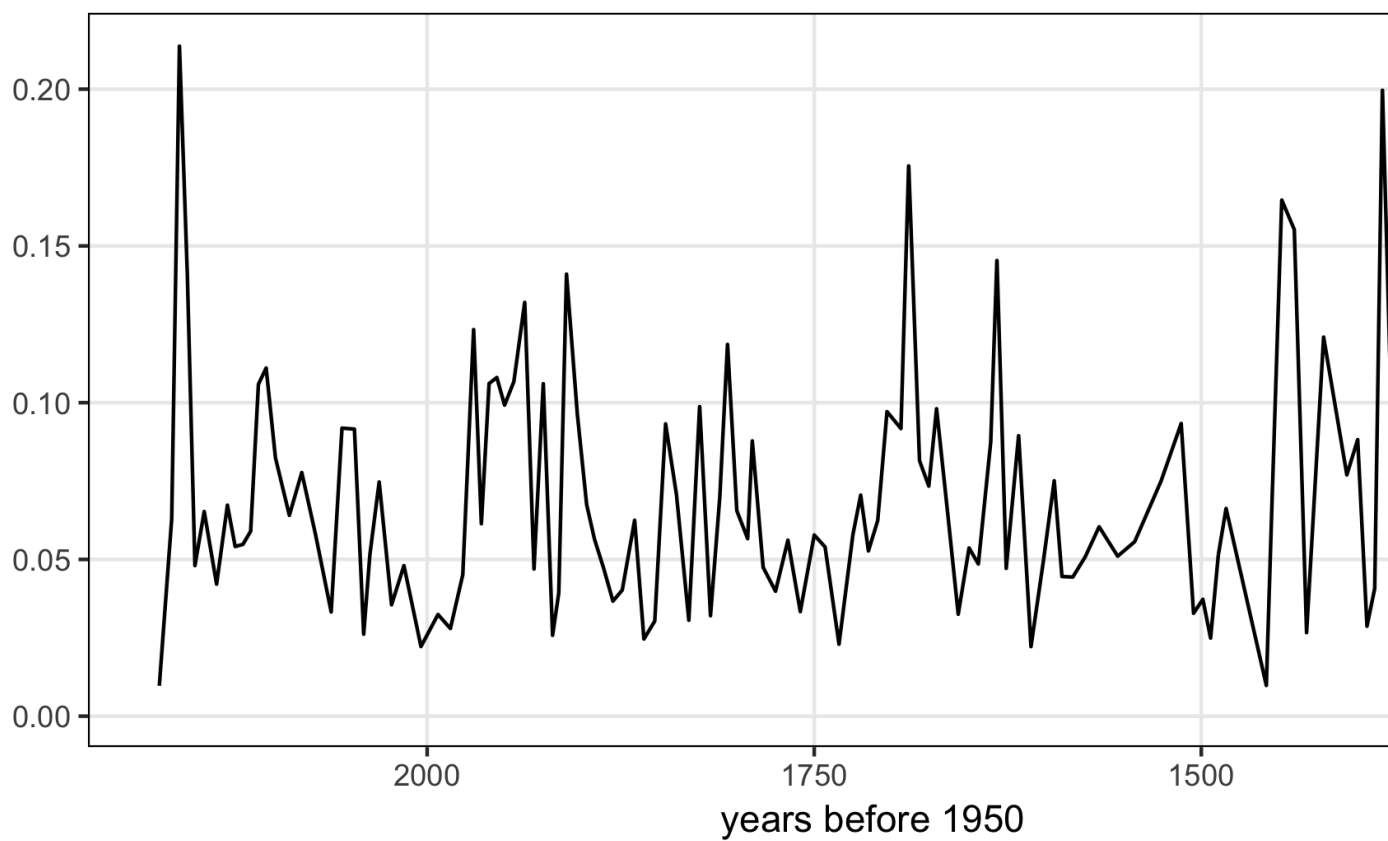
**A****B**

Figure 11: Velocity ( $v$ ) indicates periodic-like conditions in the first (A) and second (B) regimes.

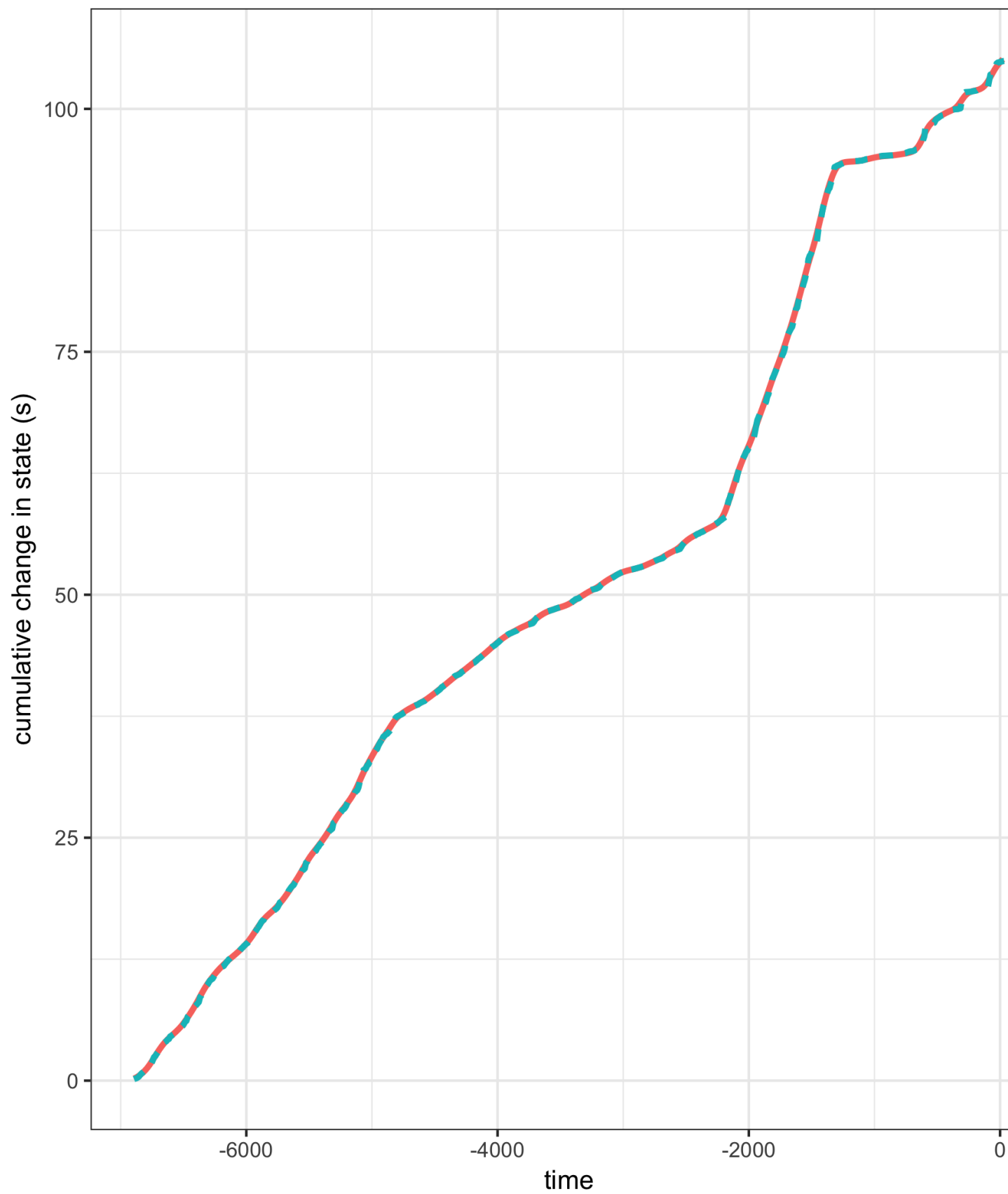


Figure 12: The regularized differentiation of  $s$  was best fit using  $\alpha = 100$ . Higher overlap of  $s$  and  $\text{pred}$  indicates a good fit of the regularized differentiated metric to the non-smoothed metric,  $s$ .

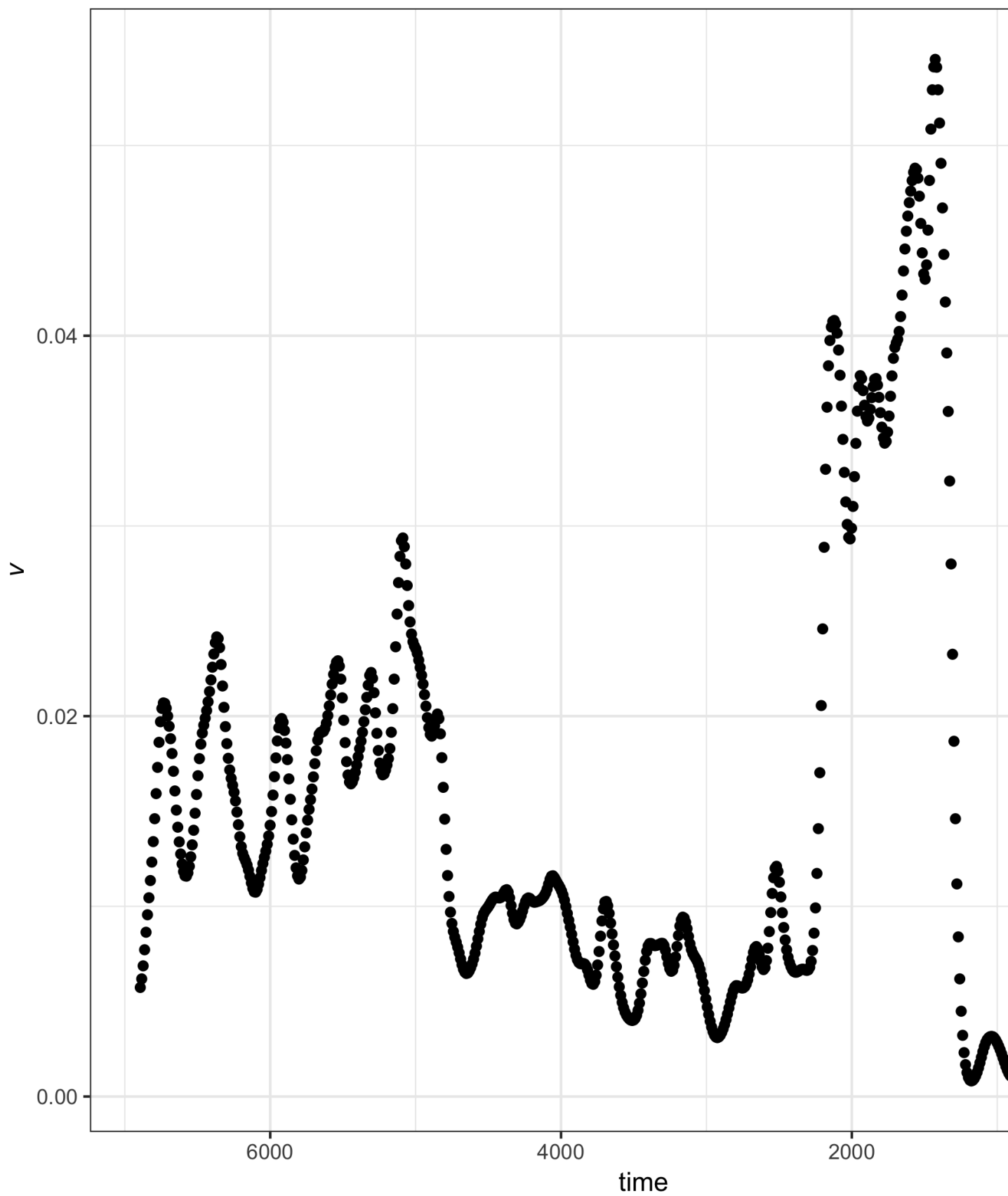


Figure 13: The velocity metric ( $v$ ) signals potential periodicities in the paleo diatom time series data when the distance travelled metric,  $s$ , is smoothed using regularized differentiation methods [price2019tvdif].

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 smoothing 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 predominant 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.