

1 *Title:*

2 **Weighted trait-abundance early warning signals better predict**  
3 **population collapse**

4

5 **Authors:** Christopher F. Clements<sup>1,2\*</sup>, Martijn van de Pol<sup>3,4</sup>, Arpat Ozgul<sup>1</sup>

6 **Affiliations:**

7 <sup>1</sup>Department of Evolutionary Biology and Environmental Studies, University  
8 of Zurich, Zurich 8057, Switzerland.

9 <sup>2</sup>School of BioSciences, University of Melbourne, Melbourne 3010, Australia.

10 <sup>3</sup> Department of Ecology & Evolution, Research School of Biology,  
11 Australian National University, Canberra 2601, Australia.

12 <sup>4</sup> Department of Animal Ecology, Netherlands Institute of Ecology (NIOO-  
13 KNAW), Wageningen 6708PB, the Netherlands.

14 **Keywords:** Body size, extinction, population dynamics, tipping points, trait dynamics

15

16 **Word count:** 2489

17

18 **Acknowledgements:**

19 This work was made possible by an SNSF Post-doctoral fellowship  
20 (P300PA\_174359/1) awarded to C.C. and an ERC Starting grant (#337785) to A.O.

21 M.vd.P. was supported by a Future fellowship from the Australian Research Council  
22 (FT12010020).

23 **Abstract:**

24 Predicting population collapse in the face of unprecedented anthropogenic  
25 pressures is a key challenge in conservation. Abundance-based early warning signals  
26 have been suggested as a possible solution to this problem; however, they are known  
27 to be susceptible to the spatial and temporal subsampling ubiquitous to abundance  
28 estimates of wild population. Recent work has shown that composite early warning  
29 methods that take into account changes in fitness-related phenotypic traits - such as  
30 body size - alongside traditional abundance-based signals are better able to predict  
31 collapse, as trait dynamic estimates are less susceptible to sampling protocols.  
32 However, these previously developed composite early warning methods weighted the  
33 relative contribution of abundance and trait dynamics evenly. Here we present an  
34 extension to this work where the relative importance of different data types can be  
35 weighted in line with the quality of available data. Using data from a small-scale  
36 experimental system we demonstrate that weighted indicators can improve the  
37 accuracy of composite early warning signals by >60%. Our work shows that non-  
38 uniform weighting can increase the likelihood of correctly detecting a true positive  
39 early warning signal in wild populations, with direct relevance for conservation  
40 management.

41

42

43 **Keywords:**

44 Body size, extinction, population dynamics, tipping points, trait dynamics

45

## 46 **Introduction**

47           Statistical early warning signals (EWSs) calculated from abundance time  
48 series data have been suggested as a possible method for predicting approaching  
49 population collapses and regime shifts (Drake & Griffen, 2010; Carpenter et al., 2011;  
50 Dakos et al., 2012; Kéfi et al., 2013). However, abundance-based early warning  
51 signals are known to be susceptible to the spatial and temporal subsampling  
52 ubiquitous to wild population abundance estimates (Clements et al., 2015), and have  
53 been criticized for not reliably predicting significant declines in natural populations  
54 (Burthe et al., 2016). Recent work has sought to resolve these issues by incorporating  
55 data on the dynamics of fitness-related phenotypic traits alongside abundance data  
56 (Clements & Ozgul, 2018). Traits such as body size are highly responsive to  
57 environmental perturbations and changes in the dynamics of these traits often precede  
58 demographic responses to deteriorating environments (Anderson et al., 2008; Ozgul et  
59 al., 2014; Clements & Ozgul, 2016a). Incorporating information on the shift in the  
60 body-size distribution of a population can not only provide an additional measure of  
61 stability (Anderson et al., 2008), but has the potential to improve the predictive  
62 accuracy of EWS as trait dynamic estimates may be less susceptible to sampling  
63 protocols than population abundance estimates are when the distribution of ages and  
64 sexes is assumed to be random (spatial partitioning between ages or sexes may affect  
65 this) (Clements et al., 2015, 2017; Clements & Ozgul, 2016a). Previous work has  
66 shown composite early warning metrics that include data on both abundance and trait  
67 dynamics better predict population collapse than those that incorporate abundance-  
68 only or trait-only data (Clements & Ozgul, 2016a).

69           Recently developed trait-abundance composite early warning indicators have  
70 been based upon the method proposed by Drake & Griffen (2010), whereby multiple

71 statistical signals are normalized and then summed to create a single composite signal.  
72 Clements & Ozgul (2016a) used this approach to incorporate shifts in mean body size  
73 and variance in body size along with concurrent changes in the statistical properties of  
74 an abundance time series, and demonstrated that such an approach can significantly  
75 improve the reliability of early warning signals in both experimental (Clements &  
76 Ozgul, 2016a) and natural (Clements et al., 2017) populations. However, in this  
77 method the relative importance of abundance versus trait data in the composite  
78 indicators was weighted evenly. Given the known issues with abundance data, a  
79 logical extension to this method is to non-evenly weight the relative importance of  
80 abundance and trait data in the composite indicators.

81 Non-uniform weighting of model parameters has a history of use in  
82 conservation biology, particularly in determining optimal management strategies to  
83 maximize outputs from limited resources (Joseph, Maloney & Possingham, 2009).  
84 For example habitat conservation may be prioritized based on the suitability of the  
85 habitat for certain species, with such weightings often being determined by expert  
86 opinion (Lehtomäki et al., 2009). Such approaches have also been used to assess  
87 trade-offs, for example between conservation and carbon sequestration (Thomas et al.,  
88 2013). As well as expert opinion, weighting may be based on more quantitative  
89 measures of data quality; for example by the frequency of sampling of a population to  
90 estimate abundances, or the percentage of a habitat sampled when counting  
91 individuals, both of which have been shown to affect the reliability of early warning  
92 signals (Clements et al., 2015). However, practitioners must first discern if non-  
93 uniform weightings convey an advantage before implementing such an approach for  
94 wild populations.

95           Here we assess whether non-uniform weightings improve the predictive ability  
96 of composite EWS of population collapse using data from an experimental protozoa  
97 study. We take the most reliable composite early warning metric (as identified by  
98 Clements & Ozgul (2016a)), and alter the relative weighting of the abundance and  
99 trait data when calculating whether an early warning signal is present or not. We then  
100 reanalyze the data from an experimental protist microcosm system, presented in  
101 Clements & Ozgul (2016a), and show that alternate weightings can improve the  
102 predictive ability of composite EWS by decreasing the frequency of false positive  
103 signals, and increasing the frequency of true positive signals.

104

## 105 **Methods**

### 106 *Experimental Data*

107           Data on the population dynamics and body-size (width,  $\mu\text{m}$  – a proxy for  
108 mass) of individuals of a predatory ciliate protozoa (*Didinium nasutum*) feeding on a  
109 bacterivorous ciliate protozoa (*Paramecium caudatum*) were collected over a 47-day  
110 period (Fig. 1). Populations of *D. nasutum* were subjected to four different treatments  
111 (15 replicates per treatment), where the number of *P. caudatum* fed to each population  
112 per day was manipulated. In one treatment (“Constant”) populations of *D. nasutum*  
113 were fed 300 *P. caudatum* per day for the 47 days of the experiment, whilst in the  
114 other three treatments the number of *P. caudatum* declined through time at three  
115 different rates (“Slow”, “Medium”, “Fast”) driving the populations of *D. nasutum* to  
116 extinction at varying points in time, and with varying population dynamics prior to  
117 extinction (Fig. 1). None of the populations in the Constant treatment went extinct.  
118 For each population the time at which it passed through a tipping point, if at all, was  
119 calculated (as in Drake & Griffen, 2010), and early warning signals were then

120 calculated prior to the occurrence of each of these tipping points. Because of the size  
121 of the microcosms it was impractical to count every individual of a population, hence  
122 a subsample was taken (10% of the habitat, a volume that allowed all individuals to  
123 be easily counted with close to zero error) and we assumed that the total number of  
124 individuals in each microcosm was reflected by the abundance in the subsample.  
125 Whilst this undoubtedly introduced some minor error into the abundance estimates,  
126 EWSs were still detectable using this uncorrected subsample data (Clements & Ozgul,  
127 2016a). We believe that this uncertainty in abundances is very representative of the  
128 ubiquitous spatial subsampling associated with the monitoring of all wild populations,  
129 and hence applying such methods to this data is a reliable reflection of the challenges  
130 of applying them to real world population dynamics. For full details of the  
131 experimental design and protocols see Clements & Ozgul (2016a).

132

### 133 *Early warning signals*

134 Previous work has identified a composite index comprised of the coefficient of  
135 variation of the abundance time series (*cv*), shifts in mean body size of the individuals  
136 in the population (*size*), and shifts in the standard deviation of mean body size  
137 (*sd.size*) as producing the most reliable estimates of whether a population was at risk  
138 of collapse in these experimental data (Clements & Ozgul, 2016a). Here we test this  
139 composite index by systematically altering the weighting of these three competent  
140 parts as a proof of concept of non-uniform weighting increasing the predictive  
141 accuracy of the composite metric.

142 Here we implement the approach developed by Clements & Ozgul (2016a).  
143 Each of the three leading indicators (*cv*, *size*, *sd.size*) was calculated at each day  
144 observations were made, and for each experimental population independently. Each

145 leading indicator was then normalized by subtracting the long-run mean of that  
146 indicator from the value of that indicator at each time point, and dividing it by the  
147 long run standard deviation (Drake & Griffen, 2010; Clements & Ozgul, 2016a)  
148 (Supplementary Information). The value of the composite early warning signal was  
149 then calculated by summing the value of each leading indicator (*cv*, *size*, *sd.size*) at  
150 each time point. Previous work has suggested an EWS could be considered present  
151 when the value of this composite EWS exceeds its running mean by either 1 or  $2\sigma$   
152 (Drake & Griffen, 2010). Recent evidence suggesting a  $2\sigma$  threshold provides more  
153 reliable results (Clements & Ozgul, 2016a) and consequently here we consider a  
154 signal present at a  $2\sigma$  threshold.

155 The weighting of each of the three leading indicators was altered by  
156 multiplying the normalized value of each metric prior to summing them together to  
157 calculate the composite EWS. Each leading indicator was weighted from 1 to 10, with  
158 every combination of weightings tested (e.g.  $cv_w=1:size_w=2:sd_w=5$ ,  
159  $cv_w=8:size_w=4:sd_w=1$ ). The performance of each weighting was assessed by using  
160 a “normalized metric score” (Clements & Ozgul, 2016a), calculated by subtracting the  
161 proportion of false positives (EWS present in data from the constant treatment) from  
162 the proportion of true positives (EWS present in data from the slow, medium, and fast  
163 treatments). The highest scoring weighting for each of the slow, medium, and fast  
164 treatments was compared to uniform weighting in each of these treatments (Fig. 2a).  
165 The best metric when data from all three treatments were grouped together was  
166 calculated as the weighting with the highest normalized metric score, and the  
167 minimum difference in normalized metric scores between treatments (Fig. 2b). This  
168 gave an indication as to the weighting that was most robust to different rates of

169 environmental change, and thus potentially most widely applicable to different  
170 scenarios.

171 All analyses were carried out using the statistical software R (R Development  
172 Core Team, 2016), and the code to implement weighted trait-abundance early warning  
173 signals is available as supplementary information.

174

## 175 **Results**

### 176 *Experimental data*

177 Non-uniform weighting of the importance of abundance and trait data in early  
178 warning indicators can improve the reliability of these methods in predicting  
179 population collapses (Fig. 2). The largest improvement (62.5%) was seen when data  
180 from the Medium treatment was analyzed, possibly because uniform weighting  
181 performed relatively poorly (Fig. 2a). The highest achieved normalized metric score  
182 was 0.8 (in the slow treatment), suggesting very high numbers of true positive EWS,  
183 and low numbers of false positive EWS (Fig. 2a).

184 When data from three deteriorating treatments was grouped together the  
185 weighting that produced the greatest improvement in predictive accuracy weighted  
186 the relative importance of *cv*, *size.sd*, *size* as 4:7:4, although the improvement over  
187 uniform weighting was not large (Fig. 2b), suggesting that the how fast the pressure  
188 on the system changes (known as the rate of forcing (Clements & Ozgul, 2016b)) may  
189 be an important factor in determining not only the correct weighting to apply, but  
190 also our ability to reliably predict population declines. To highlight this, the 4:7:4  
191 weighting performed as well as the best weighting in the fast treatment, average in the  
192 medium treatment, and worse than both the uniform and best weighting in the slow  
193 treatment (Fig. 2a).



194

195 **Discussion**

196 Predicting population collapse is a key but challenging goal in conservation  
197 biology. Because previously developed EWS that take into account both trait and  
198 abundance data are non-system specific, and thus widely applicable, they may be of  
199 particular interest. Here we analyze data from a small-scale experimental system and  
200 show that non-uniform weighting can improve the reliability and strength of trait-  
201 abundance early warning signals, but that the strength of this improvement is not  
202 uniform across different rates of environmental change.

203 Previous work in small-scale experimental systems has identified a composite  
204 metric of *cv*, *size*, and *size.sd* as the most reliable predictor of population collapse in  
205 experimental microcosm populations. This method provides improved reliability over  
206 methods that are based on either abundance-only or size-only data; however, the  
207 method still produces false positive and false negative signals in some populations.  
208 Because of the known susceptibility of abundance-based early warning signals to poor  
209 quality data (Clements et al., 2015), non-uniformly weighting the components of  
210 composite metrics provides an obvious extension to this previous work.

211 Here we demonstrate, using the same experimental data with which the  
212 original trait-abundance method was developed, that a weighting of 4:7:4  
213 (*cv:size.sd:size*) provides the greatest overall improvement across all three treatments,  
214 with the minimum between-treatment variation in this result (Fig. 2b). Resilience to  
215 treatment variation in performance in the experiment is important, as it maximizes the  
216 reliability of applying such methods to systems where the rate of change remains  
217 unknown. However, the among-treatment variation in the potential advantages of  
218 non-uniform weighting should not be ignored, as with weightings other than 4:7:4

219 there were significantly higher normalized metric scores in the medium and slow  
220 treatments (Fig 3a). Such a result is likely to be driven by the rate of forcing, known  
221 to potentially alter the detectability of EWSs (Clements & Ozgul, 2016b), of the  
222 system altering the rates of change of the mean and  $\sigma$  body size of individuals. For  
223 example, mean body size rapidly declines in the fast treatment (Fig. 1), whilst in the  
224 medium treatment body size decline is more gradual and a weighting towards the  
225 coefficient of variation of abundance, rather than towards body size, improves  
226 predictive accuracy (Fig. 2). These results suggest that the rate of forcing a system  
227 can alter the weighting that produces the most reliable predictions of an approaching  
228 population collapse.

229         Generalizing such a result to real-world systems may be problematic, as we  
230 cannot assume that the population and trait dynamics of the microcosm system  
231 analyzed here are truly representative of all real-world population collapses. Ideally  
232 one would select the weighting based on an estimate of the reliability of the available  
233 abundance or trait data, and possibly based on the rate of forcing of the system,  
234 although doing so is likely to be non-trivial. If, for example, available abundance data  
235 are known to be estimated from a survey conducted on a small proportion of the  
236 known range of a species, or are temporally limited, it may be prudent to calculate the  
237 presence of early warning signals with a bias in favor of trait-based data. A less  
238 quantitative option would be to weight metrics based on expert opinion of the  
239 reliability of the available data. However, whilst criticism has been leveled at the use  
240 of expert opinion in conservation management, it has been shown to be useful if  
241 approached with caution (Johnson & Gillingham, 2004; Martin et al., 2005, 2012).  
242 Similar caution must be applied to the non-uniform weighting of trait and abundance  
243 data in the models presented here, however the significant improvements in the

244 predictive accuracy of these approaches when weighting is non-uniform mean that  
245 canvassing expert opinion may be a relatively simple and cost-effective solution to  
246 improve predictive accuracy.

247 In conclusion, we demonstrate the possible advantages of non-uniform  
248 weighting in an early warning signals framework. This work provides a first step to  
249 improving the reliability of recently proposed abundance-trait methods (Clements &  
250 Ozgul, 2016a), and may be used to negate some of the known issues that affect  
251 abundance-based EWSs (Clements et al., 2015). Future work may seek to make more  
252 concrete recommendations on weightings based on qualitative measures such as  
253 expert opinion, or more quantitative measures such as measures of data quality, the  
254 known level of threat to a species or a population, the trophic level of the species, or  
255 its connectedness in a network. One option to tackle this is to use complex size-  
256 structured community models, such as those commonly used in fisheries (Blanchard  
257 et al., 2012; Scott, Blanchard & Andersen, 2014), to simulate shifts in the trait  
258 dynamics and abundances of multiple interacting species, allowing alternative  
259 weightings of data from various trophic levels to be tested on communities where  
260 collapse can be invoked by, for example, overfishing or changing climatic variables.

261

## 262 **References**

- 263 Anderson CNK., Hsieh C., Sandin S a., Hewitt R., Hollowed A., Beddington J., May  
264 RM., Sugihara G. 2008. Why fishing magnifies fluctuations in fish abundance.  
265 *Nature* 452:835–9.
- 266 Blanchard JL., Jennings S., Holmes R., Harle J., Merino G., Allen JI., Holt J., Dulvy  
267 NK., Barange M. 2012. Potential consequences of climate change for primary  
268 production and fish production in large marine ecosystems. *Philosophical*

- 269        *transactions of the Royal Society of London. Series B, Biological sciences*  
270        367:2979–89. DOI: 10.1098/rstb.2012.0231.
- 271    Burthe SJ., Henrys PA., Mackay EB., Spears BM., Campbell R., Carvalho L., Dudley  
272        B., Gunn IDM., Johns DG., Maberly SC., May L., Newell MA., Wanless S.,  
273        Winfield IJ., Thackeray SJ., Daunt F. 2016. Do early warning indicators  
274        consistently predict nonlinear change in long-term ecological data? *Journal of*  
275        *Applied Ecology* 53:666–676. DOI: 10.1111/1365-2664.12519.
- 276    Carpenter SR., Cole JJ., Pace ML., Batt R., Brock W a., Cline T., Coloso J., Hodgson  
277        JR., Kitchell JF., Seekell D a., Smith L., Weidel B. 2011. Early Warnings of  
278        Regime Shifts: A Whole-Ecosystem Experiment. *Science (New York, N.Y.)*  
279        332:1079–1082.
- 280    Clements CF., Blanchard JL., Nash K., Hindell M., Ozgul A. 2017. Body size shifts  
281        and early warning signals precede the historic collapse of whale stocks. *Nature*  
282        *Ecology & Evolution* 1:188.
- 283    Clements CF., Drake JM., Griffiths JI., Ozgul A. 2015. Factors influencing the  
284        detectability of early warning signals of population collapse. *Am Nat* 186:50–58.  
285        DOI: in press.
- 286    Clements CF., Ozgul A. 2016a. Including trait-based early warning signals helps  
287        predict population collapse. *Nature Communications* 7:10984. DOI:  
288        10.1038/NCOMMS10984.
- 289    Clements CF., Ozgul A. 2016b. Rate of forcing and the forecastability of critical  
290        transitions. *Ecology and Evolution* 6:7787–7793.
- 291    Clements CF., Ozgul A. 2018. Indicators of transitions in biological systems. *Ecology*  
292        *Letters* In press.
- 293    Dakos V., Carpenter SR., Brock WA., Ellison AM., Guttal V., Ives AR., Kéfi S.,

- 294 Livina V., Seekell DA., van Nes EH., Scheffer M. 2012. Methods for detecting  
295 early warnings of critical transitions in time series illustrated using simulated  
296 ecological data. *PloS one* 7:e41010. DOI: 10.1371/journal.pone.0041010.
- 297 Drake J., Griffen B. 2010. Early warning signals of extinction in deteriorating  
298 environments. *Nature* 467:456–459.
- 299 Johnson CJ., Gillingham MP. 2004. Mapping uncertainty: Sensitivity of wildlife  
300 habitat ratings to expert opinion. *Journal of Applied Ecology* 41:1032–1041.
- 301 Joseph LN., Maloney RF., Possingham HP. 2009. Optimal allocation of resources  
302 among threatened species: a project prioritization protocol. *Conservation Biology*  
303 23:328–338. DOI: 10.1111/j.1523-1739.2008.01124.x.
- 304 Kéfi S., Dakos V., Scheffer M., Van Nes EH., Rietkerk M. 2013. Early warning  
305 signals also precede non-catastrophic transitions. *Oikos* 122:641–648. DOI:  
306 10.1111/j.1600-0706.2012.20838.x.
- 307 Lehtomäki J., Tomppo E., Kuokkanen P., Hanski I., Moilanen A. 2009. Applying  
308 spatial conservation prioritization software and high-resolution GIS data to a  
309 national-scale study in forest conservation. *Forest Ecology and Management*  
310 258:2439–2449. DOI: 10.1016/j.foreco.2009.08.026.
- 311 Martin TG., Burgman MA., Fidler F., Kuhnert PM., Low-Choy S., McBride M.,  
312 Mengersen K. 2012. Eliciting Expert Knowledge in Conservation Science.  
313 *Conservation Biology* 26:29–38.
- 314 Martin TG., Kuhnert PM., Mengersen K., Possingham HP. 2005. The power of expert  
315 opinion in ecological models using Bayesian methods: Impact of grazing on  
316 birds. *Ecological Applications* 15:266–280. DOI: 10.1890/03-5400.
- 317 Ozgul A., Bateman AW., English S., Coulson T., Clutton-Brock TH. 2014. Linking  
318 body mass and group dynamics in an obligate cooperative breeder. *Journal of*

- 319        *Animal Ecology*.
- 320    R Development Core Team. 2016. R: A language and environment for statistical  
321        computing.
- 322    Scott F., Blanchard JL., Andersen KH. 2014. *mizer*: an R package for multispecies,  
323        trait-based and community size spectrum ecological modelling. *Methods in*  
324        *Ecology and Evolution* 5:1121–1125. DOI: 10.1111/2041-210X.12256.
- 325    Thomas CD., Anderson BJ., Moilanen A., Eigenbrod F., Heinemeyer A., Quaipe T.,  
326        Roy DB., Gillings S., Armsworth PR., Gaston KJ. 2013. Reconciling  
327        biodiversity and carbon conservation. *Ecology Letters* 16:39–47.
- 328

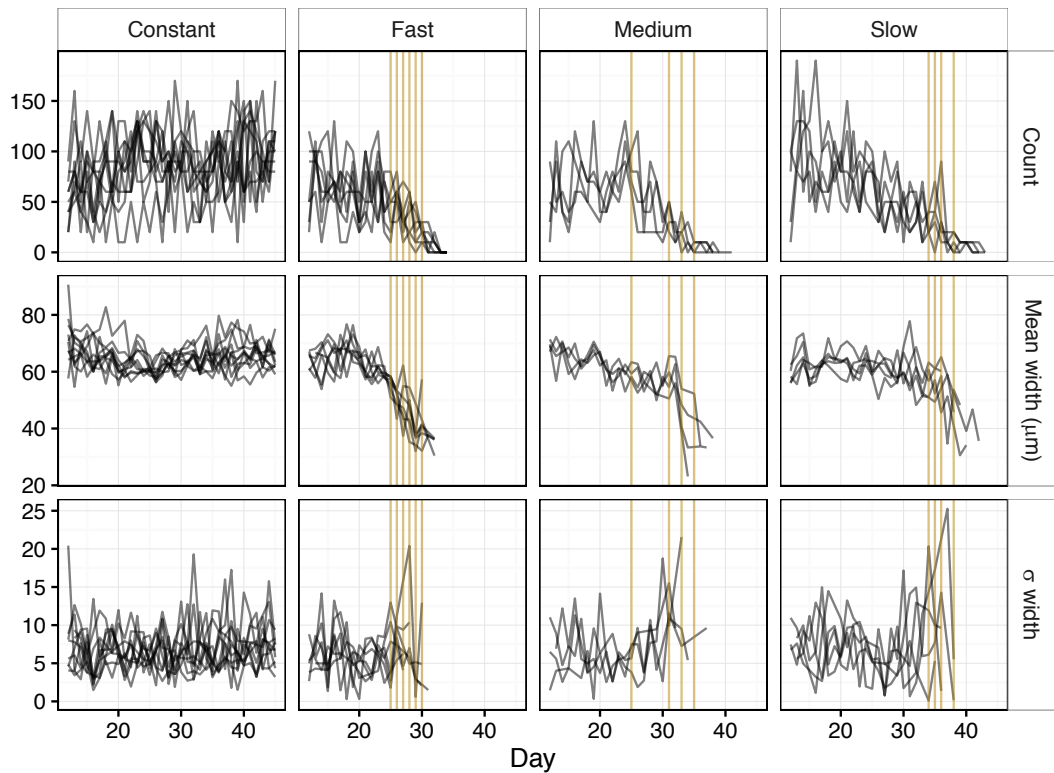


Figure 1. Black lines describe the population and body size dynamics of individual populations of *Didinium nasutum* subjected to four different experimental treatments (constant, fast, medium, and slow rates of decline in prey availability). Data from day 0 to 12 were removed to minimize the effects of transitory dynamics. Each vertical gold line indicates an inferred tipping points for a collapsing population.

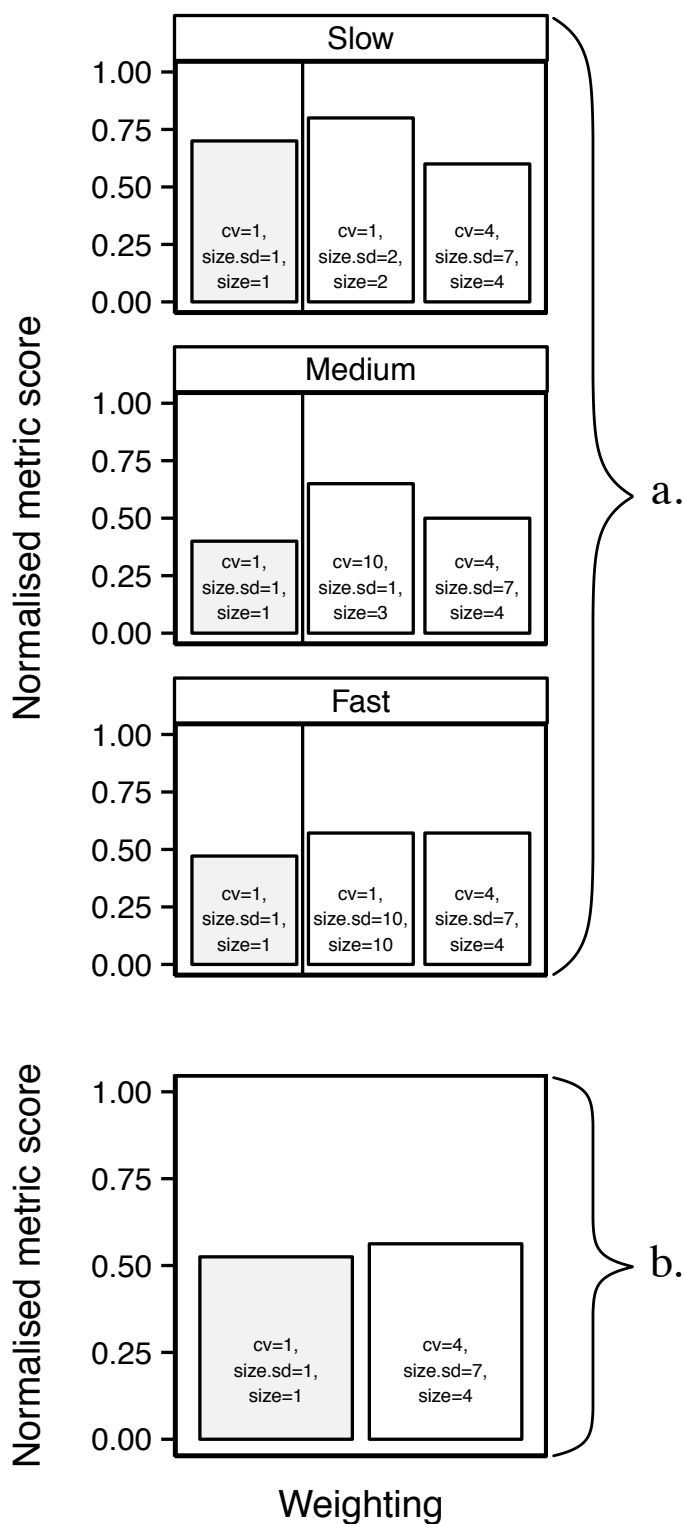


Figure 2. (a) The highest scoring weighting across each of the experimental treatments compared to even weighting and the best weighting when data from all treatments were combined, and (b) the weighting with the highest normalized metric score across all three treatments, and the lowest difference in normalized metric score amongst treatments.