Effect of time-series length and resolution on abundance- and trait-based early warning signals of population declines

Alex Arkilanian[1], Christopher F. Clements[2,3], Arpat Ozgul[2], Gaurav Baruah[2]

1 Department of Biology, McGill University, Montreal, Quebec, H3A 1B1, Canada.

2 Department of Evolutionary Biology and Environmental studies, University of Zurich, Winterthurerstrasse 30, 4055 Zurich

3 School of Biosciences, University of Melbourne, Parkville VIC 3052, Melbourne, Australia

Corresponding author: Gaurav Baruah

Email : piyushgkb@gmail.com

Running head: Effect of time-series on EWS

Abstract

Natural populations are increasingly threatened with collapse at the hands of anthropogenic effects. Predicting population collapse with the help of generic early warning signals (EWS) may provide a prospective tool for identifying species or populations at highest risk. However, pattern-to-process methods such as EWS have a multitude of challenges to overcome to be useful, including the low signal to noise ratio of ecological systems and the need for high quality time-series data. The inclusion of trait dynamics with EWS has been proposed as a more robust tool to predict population collapse. However, the length and resolution of available time series are highly variable from one system to another, especially when generation time is considered. As yet it remains unknown how this variability with regards to generation time will alter the efficacy of EWS. Here we take both a simulation- and experimental-based approach to assess the impacts of relative time-series length and resolution on the forecasting ability of EWS. We show that EWS' performance decreased with decreasing length and resolution. Our simulations suggest a relative time-series length between ten and five generations and a resolution of half a generation are the minimum requirements for accurate forecasting by abundance-based EWS. However, when trait information was included alongside abundance-based EWS, we found positive signals at lengths and resolutions half of what was required without them. We suggest that, in systems where specific traits are known to affect demography, trait data should be monitored and included alongside abundance data to improve forecasting reliability.

Keywords: early warning signals, population collapse, sampling, time-series length, trait-based EWS

Introduction

Anthropogenic pressures have long been known to reduce the resilience of ecological systems, leaving them vulnerable to transitioning into undesirable states where the systems' ability to provide valuable ecosystem services is diminished. Such undesirable transitions have occurred in multiple systems such as in the global whale stock collapse of the 20th century due to overfishing (Hilborn et al. 2003, Clements et al. 2017); lake eutrophication through heavy nutrient input (Smith and Schindler 2009); or coral bleaching as a result of increased ocean warming (Hughes et al. 2017). In many cases recovery from such a perturbed state can be difficult as complex systems such as those seen in ecology often show hysteresis (Folke et al. 2004, Scheffer et al. 2009), thus driving a need to minimize impacts on biological systems, as well as a developing effective methods to monitor them (Costanza et al. 1997).

Early warning signals (EWS) have been shown to predict population collapses (Wissel 1984, Drake and Griffen 2010, Dai et al. 2012, Clements and Ozgul 2018) and shifts in ecosystem states (Scheffer et al. 2009, Carpenter et al. 2011). These indicators provide the possibility to intervene and reverse these undesirable events (Biggs et al. 2009, Pace et al. 2017). Classical EWS are statistical signatures which arise as a result of a phenomenon known as critical slowing down (CSD) that occurs prior to an ecosystem transition (Dakos et al. 2008, Scheffer et al. 2009, Clements and Ozgul 2018). CSD occurs as a system loses stability in the face of increasing external stress and takes longer to return to its original equilibrium state (Wissel 1984). Directly measuring CSD requires monitoring the return rate of the system which is challenging to do in natural systems. Alternatively, whether a system is experiencing CSD can be inferred through other statistical metrics measured over state-based data, for example abundance time series. Some proposed statistical signatures related to the return rate of a dynamical system are variance and autocorrelation, although various other metrics have also been

developed (Dakos et al. 2012a). Increases in variance and autocorrelation in abundance time series, known as classical abundance-based EWS (Drake and Griffen 2010, Dai et al. 2012), are shown both theoretically and experimentally to occur in systems approaching transition (Drake and Griffen 2010, Carpenter et al. 2011, Dakos et al. 2012b, Clements and Ozgul 2016b).

Increases in variance and autocorrelation, along with other classical abundance-based EWS, provide an ideal generic method that responds to the dynamics of the system independent of any system-specific data such as intrinsic demographic processes and extrinsic environmental factors. However, the performance of these classical EWS has been questioned in numerous simulations (Hastings and Wysham 2010, Boerlijst et al. 2013, Clements et al. 2015, Burthe et al. 2016, Clements and Ozgul 2016a, Dutta et al. 2018) as well as in experimental and field data (Wilkinson et al. 2017, Pace et al. 2017). Recent studies on data quality have shown that these signals might require high-resolution time-series data to produce reliable forecasting (Clements et al. 2015). Given that EWS are statistics derived from abundance time series, the quality of data available is critical to obtain a strong forecast by EWS and temporal limitations might have many consequences.

In monitoring programs from natural ecological systems high resolution data can be hard to achieve. In addition, data can be spatially and temporally limited due to constraints on resources. Thus, data from ecological systems can often present with short time-series lengths, low sampling resolutions, or both (Clements et al. 2015). Further, abundance time-series data collected in field or experimental populations can vary greatly in temporal quality. For instance, in a laboratory experiment, populations of *Didinium nasutum* were sampled roughly once every two generations for 45 days before a collapse of the population occurred (Clements and Ozgul 2016a) while in a field experiment data on a lake system was collected once a day during the summer season for three years (Carpenter et al. 2011).

Similarly, in wild populations demographic data is generally collected monthly (van Benthem et al. 2017) or annually (Walle et al. 2018). Due to these differences in sampling effort the temporal quality of the time series relative to the process rate of the system (for example the generation time of the species sampled) will vary (Clements and Ozgul 2018). Previous work suggested that the length and resolution of data being analyzed in relation to the process rate, that is: the relative length and resolution, could alter the rate at which a tipping point occurs (Spanbauer et al. 2016). Following this, it has been shown that the speed at which a tipping point occurs can affect the detectability of abundance-based EWS (Clements and Ozgul 2016b). It is thus unknown how the temporal quality of the time series relative to the generation time of the organisms being monitored will affect the detectability of abundance-based EWS. Given that wild populations are monitored using varied sampling efforts, it is important to further our understanding of the relative length and resolution of time series needed to derive reliable EWS. In turn, this would help lay the foundation for generalizable guidelines for the monitoring of populations with the aim of reliably predicting population declines, or whether time series that are currently available will be suitable for detecting abundance-based EWS.

Composite EWS have been proposed as a more reliable method whereby multiple leading indicators are combined to increase overall forecasting ability (Drake and Griffen 2010). Recent work has used this composite approach to drive the inclusion of fitness related phenotypic trait data, specifically body size, alongside abundance-based methods to create trait-based EWS (Clements and Ozgul 2016a, Clements et al. 2017). The motivation behind the use of fitness-related trait data combined with leading indicators comes from a body of work providing evidence that individual traits affected by changes in the external environment are linked with concurrent demographic changes (Ozgul et al. 2009, Pigeon et al. 2017, Baruah et al. 2018b). In particular, individual plasticity in body size has been shown to buffer

environmental change to higher trophic levels for example in the face of reduced prey availability, climate change, and increased pollution (Brown et al. 2004, Cheung et al. 2013), For example, previous study has shown that shifts in body size can accompany a transition in diatom communities (Spanbauer et al. 2016). In addition, trait-based signals derived from body size data have been shown to be more robust than traditional abundance-based leading indicators (Clements and Ozgul 2016a, Clements et al. 2017, Baruah et al. 2018a). There remains a need to fully assess whether the inclusion of body size data leads to any considerable improvement in forecasting population collapses in the face of common data quality issues, such as shortened time-series lengths.

In this paper, we use model-simulations and data from microcosm populations of *Didnium nasutum* (Clements and Ozgul 2016a) to test and compare the effects of sampling length and resolution on the detectability of population collapse by both classical abundance-based EWS as well as trait-based EWS. We first investigated the strength and reliability of abundance-based EWS across a range of sampling lengths and resolutions. Subsequently, we tested whether the inclusion of trait dynamics (body size) can increase forecasting ability even when data is sparse.

Methods

2.1 Simulations: transcritical and fold bifurcation model

We first modeled the logistic growth of a population that moves from an underexploited state to critically exploited state through a non-catastrophic transcritical bifurcation. The population is forced through the transcritical bifurcation via a linear harvesting regime. The dynamics of this population are given by:

$$dN/dt = rN(1 - N/K) + \sigma N dW - c_t N \quad (1)$$

where, r is the growth rate of the population (0.5 individuals/day), K is the carrying capacity (100 individuals), c_t is the harvesting rate, and σNdW is the Gaussian distributed white noise process with mean 0 and standard deviation σ (1.5). Time step dt used was 0.3 for each of the stochastic simulations and was implemented using the Euler approximation. The simulation was run for 100 total time steps. Next, to simulate the dynamics of population collapse via a fold catastrophe, we used a second model where harvesting of individuals of the population followed a non-linear function (May 1977, Scheffer 2009). The parameters in this model are identical to those of the transcritical catastrophe model except for the addition of h, the half-saturation constant:

$$dN/dt = rN(1 - N/K) + \sigma NdW - c_t N^2/(h^2 + N^2) \quad (2)$$

2.2. Population collapse and abundance-based EWS:

We simulated population collapse for the two models by increasing the value of the harvest parameter c_t linearly with time. We used three different levels of forcing (Clements and Ozgul 2016b) : 1) fast forcing: where the forcing parameter c_t increased linearly at the rate of 0.07 and 0.004 in fold and transcritical bifurcation models respectively; 2) moderate forcing: where the forcing parameter c_t increased linearly at the rate of 0.045 and 0.0025 in fold and transcritical models respectively; 3) slow forcing: where the rate of forcing increased linearly from 0.03 and 0.0015 in fold and transcritical models respectively. For each simulated population's time-series data, we estimated the bifurcation time point by fitting GAMS (Generalized Additive Modelling) to $1/N\ (dN/dt)$ over time t. The time point at which $1/N\ (dN/dt) < 0$ is then our estimated bifurcation time point. For abundance-based EWS' analyses we discarded abundance time-series data after the estimated bifurcation point. We then applied generic EWS of population collapse by using the *earlywarnings* package (Dakos et al. 2012a) in R version 3.5.2 (R Core Team 2018). Specifically, we used two early warning indicators: autocorrelation at first-lag (*ar(1)*) and standard deviation (*sd*). Other indicators such as return rate, first-

order autoregressive coefficient, coefficient of variation can theoretically be derived from these two main indicators. We used Gaussian detrending to remove any trends in the abundance time-series data. To quantify the strength of population collapse, we calculated Kendall's Tau correlation coefficients of the statistical indicators over time. Strong positive Kendall's Tau correlation of the statistical indicators (sd, ar(1)) with time would indicate an approaching population collapse (Dakos et al. 2012b). We further quantified the rate of false negatives as the number of times Kendall's Tau value calculated to be less than or equal to zero within the set of replicate time series. While previous work has suggested that a strong trend is indicated by a Kendall's tau correlation approaching one (Dakos et al. 2012b), we use this false negative metric as an alternative visualization to presenting raw Kendall's tau values.

2.3 Effect of sampling in relation to simulated population dynamics on EWS

To study the effect of varying resolutions on the detectability of collapse, we subset the data into four datasets that varied in their resolution: as the generation time for the populations in the simulation models is t=1, sampling of the abundance time series was done every quarter of *t*, every half of *t,* every *t,* and every two *t.* Interpolation between points was not performed. With these four datasets, we explored the effect of different sampling resolutions on the efficacy of EWS forecasting. Specifically, we assessed the resolution required to detect positive EWS. We assessed the effect of these sampling regimes on EWS for the three different levels of environmental forcing as mentioned in section 2.2. Next, we calculated Kendall's tau correlation coefficients as a measure of the strength of EWS and rates of false negatives as a measure of the reliability of EWS for these sampling regimes and for each level of environmental forcing.

2.4 Effect of varying length of simulation abundance time series on EWS

To quantify the effect of varying lengths of abundance time series on EWS we used the quarter generation sampling resolution dataset acquired following section 2.3. We used this resolution as it allowed us to explore the largest range of time-series lengths. Next, we quantified the strength and reliability of EWS on the entire length of the abundance time series and then re-quantified it using the same time series but with the earliest data point removed (or the furthest data point from the bifurcation point). We repeated this process by sequentially removing the earliest data point and re-performing EWS analysis until the time series was too short for any meaningful analysis. We performed this length reduction analysis on time series from the three forcing scenarios mentioned in section 2.2. Finally, we calculated Kendall's tau correlation coefficients as a measure of the strength of EWS and rates of false negatives as a measure of the reliability of EWS for these sampling regimes and for each level of environmental forcing.

2.5. Experimental data

In addition to the model simulations, we also analyzed an experimental microcosm dataset. In this experiment, microcosm populations were forced to collapse by varying the rate of decline in prey availability over time in four different scenarios: 1) fast decline in food availability, 2) moderate decline in prey availability, 3) slow decline in prey availability, 4) constant prey availability as the control treatment. The microcosm populations consisted of protozoan ciliate *Didinum nasutum* feeding on *Paramecium caudatum*. This particular experiment used a total of 60 replicate populations, where 15 replicates were used per treatment. In our study, we used the microcosm data only from the three different deteriorating environments (fast, moderate and slow decline in prey availability). For details of the experimental design refer to Clements and Ozgul 2016a.

2.5.1 Effect of varying length of experimental time series on EWS

For each of the three experimental treatments (fast, moderate, and slow), we estimated the effect of different lengths of time series on the efficacy of EWS as was done with simulation data in section 2.4 to directly compare our simulation results with the experimental data.

2.5.2 Effect of sampling in relation to microcosm population dynamics on EWS

We subset the experimental data into two sampling regimes. Since *Didinum nasutum* has a generation time of roughly 2 days (Beers 1926), sampling of the abundance time-series data was done: 1) every half a generation (everyday), 2) every generation (every 2 days). Subsampling of this kind was done for each of the three experimental treatments of the microcosm population collapse. Next, as was done with simulation data, we quantified Kendall's tau correlation coefficients as a measure of the strength of EWS and rates of false negatives as a measure of the reliability of EWS for these sampling regimes and for each level of environmental forcing.

2.6 Inclusion of body size data: trait-based EWS

We wanted to assess whether including trait dynamic information (body size) alongside abundance-based EWS would improve the predictability of population collapse for the scenarios of time-series length and for the sampling resolutions in the experimental data for the three forcing experimental treatments. To evaluate the utility of trait-based EWS for the different sampling resolutions and lengths of abundance time series, data from mean body size was incorporated with the leading indicators in an additive manner (Clements and Ozgul 2016a). We z-standardized abundance-based EWS (standard deviation (sd), autocorrelation at-lag-1 (ar(1)), and body size time-series data. The length of body-size time series that was used was the same as the corresponding replicate abundance time series. Before a population collapse, sd and ar(1) are anticipated to increase linearly over time, while body-size is expected to decline as food availability in the experimental treatment decreases. As a consequence, z-

standardized body-size time series were multiplied by -1 so that they could be included alongside standardized abundance-based EWS. Next, standardized abundance-based EWS were then added to standardized mean body-size time series to create trait-based EWS. Next, we evaluated two trait-based EWS metrics namely *ar(1)+mean body size, sd+mean body size*. Finally, we compared these two trait-based EWS in terms of strength and reliability of forecasting a population collapse with the abundance-based statistical EWS for both the scenarios of varying time-series length and varying sampling resolution.

Results

3.1 Effect of sampling resolution on abundance-based EWS (simulated data)

Reducing the resolution of the time series used for analysis with abundance-based EWS in simulated populations led to reductions in the strength of abundance-based indicators of collapse. Decreases in the resolution of the time series generally led to small decreases in the median strength of the signal of population collapse (Fig. 1) and these were mostly seen only in *sd* especially when forcing was strong or moderate (Fig. 1 - yellow). While the median strength of signals dropped slightly with decreasing resolution, the maximum predicting strength achieved (the highest Kendall's tau value) dropped more sharply leading to less confident predictions with lower resolutions. This trend of decreased strength of EWS was generally maintained between bifurcation types (Fig. 1). Further, as resolution was decreased, the spread of Kendall's tau values decreased leading to a smaller number of false negative signals. However, the proportion of these false negatives followed an upward trend with decreasing resolution in both bifurcation types indicating a decrease in the reliability of forecasting in addition to reduced strength (Table S1).

3.2 Effect of length of time series on abundance-based EWS (simulated data):

In simulations, decreasing the length of the sampling time series before either a fold or transcritical bifurcation negatively affected the performance of abundance-based EWS *ar(1)* and *sd* across all three intensities (Fig. 2). Moreover, as forcing increased, the decline of signal reliability (Fig. 2B) and signal strength (Fig. 2A), as indicated by the rate of false negatives and the Kendall's tau value respectively, saw a steeper decline. With strong forcing, when time series dropped below approximately ten generations long both EWS performed poorly with Kendall's tau values near zero (Fig. 2A) and approximately 50% false negatives (Fig. 2B). With moderate forcing, this minimum length drops closer to 5 generation long timeseries. Finally, with slow forcing the minimum timeseries length becomes less clear.

3.3 Effect of sampling resolution on abundance and trait-based EWS (experimental data):

Decreasing the resolution of the time series in microcosm populations led to decreases in the strength of abundance-based EWS (Fig. 3), corroborating the results of model simulations (Fig. 1). The decrease in strength of was most noticeable in *sd*, while *ar(1)* remained nearly constant as resolution was decreased (Fig. 3).

Including information from mean body size data alongside abundance-based EWS to create trait-based EWS led to significant increases in strength and reliability of predicting population collapse regardless of forcing strength, or time-series resolutions (Fig. 3, dark colors).

3.4 Effect of length of time series on abundance and trait-based EWS (experimental data):

In agreement with the results from simulations, abundance-based EWS derived from microcosm data saw an increase in false negatives and a decrease in signal strength as the length of the sampling time

series was decreased (Fig. 4A, 4B, dotted lines). This was most noticeable with fast forcing but also occurred in all forcing conditions with time series of six or less generations long. However, the drop in performance seen in microcosm data was lower than in the simulations, which is partially due to the fact that the former performs the analysis on a smaller range of time-series lengths.

Furthermore, trait-based EWS i.e., including body size information alongside abundance EWS, significantly improved the strength as well as reliability of population collapse across different length in time-series data for different levels of environmental forcing. With trait-based EWS, Kendall's Tau value always remained positive regardless of very short time series and forcing strength (Fig 4A, solid lines). Similarly, the number of false negatives was also significantly lower than the abundance-based EWS alone across different time-series lengths (Fig 4B, solid lines).

Discussion

Generic EWS would provide a unique tool for conservation prioritization and management of populations facing increased stress with changes to their abiotic environment if they are detectable prior to their collapse (Scheffer et al. 2009, Dakos et al. 2012b). The attraction of generic EWS is their relative simplicity; they are easy to calculate and require only state data such as the abundance of a population. Alternative approaches such as trait-based EWS have been developed with the goal of providing more reliable predictions of population collapse but require additional data to calculate (Clements and Ozgul 2016a, Clements et al. 2017). For both complimentary approaches how the relative length and resolution of a time series may affect their performance has thus far remained unknown. This question is critical for our understanding of the utility and applicability of these methods. Here, using simulations and experimental data, we show that time-series length and resolution

relative to the process rate of the system in question significantly influence the performance of abundance-based EWS. Further, we found that including average body size with abundance-based EWS leads to stronger and more reliable signals even when temporal length and resolution of the time series are low.

4.1 Effect of resolution of sampling time series on abundance-based and trait-based EWS

Reducing the resolution of time series in simulated and experimental populations subjected to different forcing levels led to a decrease in the performance of abundance-based EWS (Fig. 1 and 3). There were important differences in the behavior of abundance-based EWS calculated between simulated and experimental populations. While in simulated population we found *sd* to produce the most robust signal of population collapse (Fig. 1), in experimental populations *ar(1)* performed better across varying resolutions (Fig. 3). Our simulated populations were likely subjected to smaller amounts of stochasticity in population size compared to our experimental populations. Given that standard deviation-based EWS are more sensitive to rising stochasticity in a system (Boettiger and Hastings 2012), this difference between experimental and simulated populations is expected. In addition, abundance-based EWS showed a greater sensitivity to system stochasticity than trait-based EWS did; the latter did not present any obvious differences in performance between metrics used (Fig. 3). This is possibly an effect of adding phenotypic data to abundance data which stabilizes the signal through time and reduces stochasticity. This is an important finding if abundance-based EWS are to be used to monitor natural populations which likely have more stochasticity than laboratory populations.

As observed with abundance-based EWS, trait-based EWS also saw a decrease in their performance with decreasing sampling resolution (Fig. 3). However, these drops were in most cases very small and the performance of trait-based EWS was much better than abundance-based EWS alone. Trait-based

EWS calculated with the lowest resolution in either simulated or experimental populations still derived a more reliable forecast than abundance-based EWS calculated with the highest possible resolution.

Our findings suggest that abundance-based EWS derived from autocorrelation at-first-lag would perform better in natural populations subjected to high levels of stochasticity when compared with standard deviation derived signals when the resolution of available abundance time series is low. On the other hand, the inclusion of trait data with either metric led both to an increase in EWS performance when faced with reduced sampling time-series resolutions and minimized any important differences between the two time-series metrics used. Thus, whenever possible, trait-based EWS will likely outperform abundance-based EWS in natural populations (but see Baruah et al. 2018a) and, when trait data is not available, abundance-based EWS derived from autocorrelation at-first-lag are more powerful and reliable indicators.

4.2 Effect of length of sampling time series on abundance-based and trait-based EWS

Reducing the length of the sampling time series in either simulated or experimental data negatively affected the performance of abundance-based EWS. In our simulation study, time-series lengths of less than 10 generations long and less than 5 generations long when forcing is fast and moderate respectively, returned on average, incorrect forecasts indicating no population collapse (Fig. 2). The choice of fold or transcritical bifurcation models had little influence on predictability of population collapse with the exception of the fast forcing scenario. When forcing was fast, transcritical models showed a shift in trend indicating the length of time series no longer had a strong effect on the predictability of collapse (Fig. 2 – dashed lines). In the context of a ciliate population such as the one used in the microcosms of this paper (*D. nasutum*) 10 generations corresponds to roughly 20 days. Twenty days' worth of sampling might seem feasible to generate reliable forecasting for a ciliate

population. However, for this same reliability in a population experiencing rapid forcing, 22 to 35 years of sampling would be required in the case of *Thynnus thynnus* (Atlantic bluefin tuna): an endangered species of high economic concern with a generation time between 2.2 to 3.5 years (Collette et al. 2011). For organisms with longer generation times such as *T. thynnus,* the required investment of resources and time becomes an important consideration if there is a desire to monitor populations using abundance-based EWS. In addition, we found that abundance time series require a minimum of 10 generations worth of data when forcing is fast and a minimum of 5 when forcing is moderate (fig. 2). This finding suggests that in situations where forcing is more intense and, thus, populations are most at risk the requirement for good quality data is extended. This requirement of long time series is a clear shortcoming of abundance-based EWS for organisms with long generation times and populations experiencing rapid forcing.

In our microcosm study, the effect of time-series length on the Kendall's tau metric was less clear than in the simulation. However, a general decreasing trend in forecasting strength was still observed in most forcing scenarios. In addition, abundance-based EWS rarely had the length of sampling required to make a confident forecast of collapse. It appears that regardless of improvements brought about by the increased length of the sampling time series, most abundance-based EWS did not provide a confident forecast even with the maximum length of sampling time series (between 10 and 12.5 generations). In our simulations we showed that abundance time series need a minimum of 10 generations of data to provide an accurate forecast. These microcosm results show that the required minimum length of sampling exceeds 10 generations in controlled laboratory settings. Thus, we can expect that in natural populations subjected to higher levels of stochasticity the minimum length of sampling might exceed even that which is required in the microcosms. However, in natural populations there is the possibility to measure phenotypic data and derive trait-based EWS which we have

additionally assessed. The addition of phenotypic data with abundance data improved forecasting in the microcosms. Trait-based EWS, that included body size information alongside abundance-based EWS, offered a significant improvement over abundance-based EWS, providing a positive forecast of collapse for all time-series lengths tested.

Our findings on the effect of sampling length on the forecasting of population collapse by EWS suggest that the length of sampling required to calculate confident forecasts by abundance-based EWS are very long and likely impractical from a conservation and monitoring standpoint. On the other hand, trait-based EWS can give accurate and reliable forecasts with much less sampling data compared to their abundance-based counterparts.

Conclusion

Our results stemming from simulations and a laboratory microcosm study make a case for the preferential use of trait-based EWS over classical abundance-based signals based on their increased reliability and strength when faced with varying lengths and resolutions of sampling time series. We attempted here to evaluate these EWS with data which might more realistically resemble data gathered from field monitoring programs. However, true field data arising from monitoring programs are likely to be noisier than the data we have presented and analyzed here. Nonetheless, our results suggest that abundance-based EWS perform poorly when the length or resolution of abundance time series with regards to the process rate of the system is decreased. We found that including trait dynamics alongside abundance-based EWS to generate trait-based EWS leads to more reliable and confident forecasts of population collapse. We further found in our simulations that a length of ten generations and a resolution of a half generation are the minimum for deriving confident forecasts with abundance-based

EWS. If trait-based EWS are used, this length and resolution requirement is much more relaxed. With these results in mind, we recommend considerations be made for the length and resolution of sampling time series required to accurately forecast populations in the design of monitoring programs. Further, we recommend the additional monitoring of phenotypic traits in populations, which have been shown to vary with increasing levels of environmental forcing such that trait-based EWS which have more forecasting power and reliability can be derived.

Acknowledgments

AA acknowledges support and funding from the University of Zurich BUSS program.

Literature cited

Baruah, G., C. F. Clements, F. Guillaume, and A. Ozgul. 2018a. When do shifts in trait dynamics precede population declines? bioRxiv:424671.

Baruah, G., C. F. Clements, and A. Ozgul. 2018b. Eco-evolutionary processes underlying early warning signals of population declines. bioRxiv:422915.

Beers, C. D. 1926. The life-cycle in the ciliate Didinium nasutum with reference to encystment. Journal of Morphology 42:1–21.

van Benthem, K. J., H. Froy, T. Coulson, L. L. Getz, M. K. Oli, and A. Ozgul. 2017. Trait-demography relationships underlying small mammal population fluctuations. Journal of Animal Ecology 86:348–358.

Biggs, R., S. R. Carpenter, and W. A. Brock. 2009. Turning back from the brink: Detecting an impending regime shift in time to avert it. Proceedings of the National Academy of Sciences 106:826–831.

Boerlijst, M. C., T. Oudman, and A. M. de Roos. 2013. Catastrophic Collapse Can Occur without Early Warning: Examples of Silent Catastrophes in Structured Ecological Models. PLOS ONE 8:e62033.

Boettiger, C., and A. Hastings. 2012. Early warning signals and the prosecutor's fallacy. Proceedings of the Royal Society of London B: Biological Sciences 279:4734–4739.

Brown, J. H., J. F. Gillooly, A. P. Allen, V. M. Savage, and G. B. West. 2004. Toward a Metabolic Theory of Ecology. Ecology 85:1771–1789.

Burthe, S. J., P. A. Henrys, E. B. Mackay, B. M. Spears, R. Campbell, L. Carvalho, B. Dudley, I. D. M. Gunn, D. G. Johns, S. C. Maberly, L. May, M. A. Newell, S. Wanless, I. J. Winfield, S. J.

Thackeray, and F. Daunt. 2016. Do early warning indicators consistently predict nonlinear change in long-term ecological data? Journal of Applied Ecology 53:666–676.

Carpenter, S. R., J. J. Cole, M. L. Pace, R. Batt, W. A. Brock, T. Cline, J. Coloso, J. R. Hodgson, J. F. Kitchell, D. A. Seekell, L. Smith, and B. Weidel. 2011. Early Warnings of Regime Shifts: A Whole-Ecosystem Experiment. Science 332:1079–1082.

Cheung, W. W. L., J. L. Sarmiento, J. Dunne, T. L. Frölicher, V. W. Y. Lam, M. L. D. Palomares, R. Watson, and D. Pauly. 2013. Shrinking of fishes exacerbates impacts of global ocean changes on marine ecosystems. Nature Climate Change 3:254–258.

Clements, C. F., J. L. Blanchard, K. L. Nash, M. A. Hindell, and A. Ozgul. 2017. Body size shifts and early warning signals precede the historic collapse of whale stocks. Nature Ecology & Evolution 1:0188.

Clements, C. F., J. M. Drake, J. I. Griffiths, and A. Ozgul. 2015. Factors Influencing the Detectability of Early Warning Signals of Population Collapse. The American Naturalist 186:50–58.

Clements, C. F., and A. Ozgul. 2016a. Including trait-based early warning signals helps predict population collapse. Nature Communications 7:10984.

Clements, C. F., and A. Ozgul. 2016b. Rate of forcing and the forecastability of critical transitions. Ecology and Evolution 6:7787–7793.

Clements, C. F., and A. Ozgul. 2018. Indicators of transitions in biological systems. Ecology Letters 21:905–919.

Collette, B. B., K. E. Carpenter, B. A. Polidoro, M. J. Juan-Jordá, A. Boustany, D. J. Die, C. Elfes, W. Fox, J. Graves, L. R. Harrison, R. McManus, C. V. Minte-Vera, R. Nelson, V. Restrepo, J. Schratwieser, C.-L. Sun, A. Amorim, M. B. Peres, C. Canales, G. Cardenas, S.-K. Chang, W.-C.

Chiang, N. de O. Leite, H. Harwell, R. Lessa, F. L. Fredou, H. A. Oxenford, R. Serra, K.-T. Shao, R. Sumaila, S.-P. Wang, R. Watson, and E. Yáñez. 2011. High Value and Long Life—Double Jeopardy for Tunas and Billfishes. Science 333:291–292.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton, and M. van den Belt. 1997. The value of the world's ecosystem services and natural capital. Nature 387:253–260.

Dai, L., D. Vorselen, K. S. Korolev, and J. Gore. 2012. Generic Indicators for Loss of Resilience Before a Tipping Point Leading to Population Collapse. Science 336:1175–1177.

Dakos, V., S. R. Carpenter, W. A. Brock, A. M. Ellison, V. Guttal, A. R. Ives, S. Kéfi, V. Livina, D. A. Seekell, E. H. van Nes, and M. Scheffer. 2012a. Methods for Detecting Early Warnings of Critical Transitions in Time Series Illustrated Using Simulated Ecological Data. PLOS ONE 7:e41010.

Dakos, V., E. H. van Nes, P. D'Odorico, and M. Scheffer. 2012b. Robustness of variance and autocorrelation as indicators of critical slowing down. Ecology 93:264–271.

Dakos, V., M. Scheffer, E. H. van Nes, V. Brovkin, V. Petoukhov, and H. Held. 2008. Slowing down as an early warning signal for abrupt climate change. Proceedings of the National Academy of Sciences 105:14308–14312.

Drake, J. M., and B. D. Griffen. 2010. Early warning signals of extinction in deteriorating environments. Nature 467:456–459.

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

Folke, C., S. Carpenter, B. Walker, M. Scheffer, T. Elmqvist, L. Gunderson, and C. S. Holling. 2004. Regime Shifts, Resilience, and Biodiversity in Ecosystem Management. Annual Review of Ecology, Evolution, and Systematics 35:557–581.

Hastings, A., and D. B. Wysham. 2010. Regime shifts in ecological systems can occur with no warning. Ecology Letters 13:464–472.

Hilborn, R., T. A. Branch, B. Ernst, A. Magnusson, C. V. Minte-Vera, M. D. Scheuerell, and J. L. Valero. 2003. State of the World's Fisheries. Annual Review of Environment and Resources 28:359–399.

Hughes, T. P., M. L. Barnes, D. R. Bellwood, J. E. Cinner, G. S. Cumming, J. B. C. Jackson, J. Kleypas, I. A. van de Leemput, J. M. Lough, T. H. Morrison, S. R. Palumbi, E. H. van Nes, and M. Scheffer. 2017. Coral reefs in the Anthropocene. Nature 546:82–90.

Ozgul, A., S. Tuljapurkar, T. G. Benton, J. M. Pemberton, T. H. Clutton-Brock, and T. Coulson. 2009. The Dynamics of Phenotypic Change and the Shrinking Sheep of St. Kilda. Science 325:464–467.

Pace, M. L., R. D. Batt, C. D. Buelo, S. R. Carpenter, J. J. Cole, J. T. Kurtzweil, and G. M. Wilkinson. 2017. Reversal of a cyanobacterial bloom in response to early warnings. Proceedings of the National Academy of Sciences 114:352–357.

Pigeon, G., T. H. G. Ezard, M. Festa-Bianchet, D. W. Coltman, and F. Pelletier. 2017. Fluctuating effects of genetic and plastic changes in body mass on population dynamics in a large herbivore. Ecology 98:2456–2467.

R Core Team. 2018. R: A Language and Environment for Statistical Computing. Vienna, Austria.

Scheffer, M., J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. van Nes, M. Rietkerk, and G. Sugihara. 2009. Early-warning signals for critical transitions. Nature 461:53–59.

Smith, V. H., and D. W. Schindler. 2009. Eutrophication science: where do we go from here? Trends in Ecology & Evolution 24:201–207.

Spanbauer, T. L., C. R. Allen, D. G. Angeler, T. Eason, S. C. Fritz, A. S. Garmestani, K. L. Nash, J. R. Stone, C. A. Stow, and S. M. Sundstrom. 2016. Body size distributions signal a regime shift in a lake ecosystem. Proc. R. Soc. B 283:20160249.

Walle, J. V. de, G. Pigeon, A. Zedrosser, J. E. Swenson, and F. Pelletier. 2018. Hunting regulation favors slow life histories in a large carnivore. Nature Communications 9:1100.

Wilkinson, G. M., S. R. Carpenter, J. J. Cole, M. L. Pace, R. D. Batt, C. D. Buelo, and J. T. Kurtzweil. 2017. Early warning signals precede cyanobacterial blooms in multiple whole-lake experiments. Ecological Monographs 88:188–203.

Wissel, C. 1984. A Universal Law of the Characteristic Return Time near Thresholds. Oecologia 65:101–107.

Figure legends

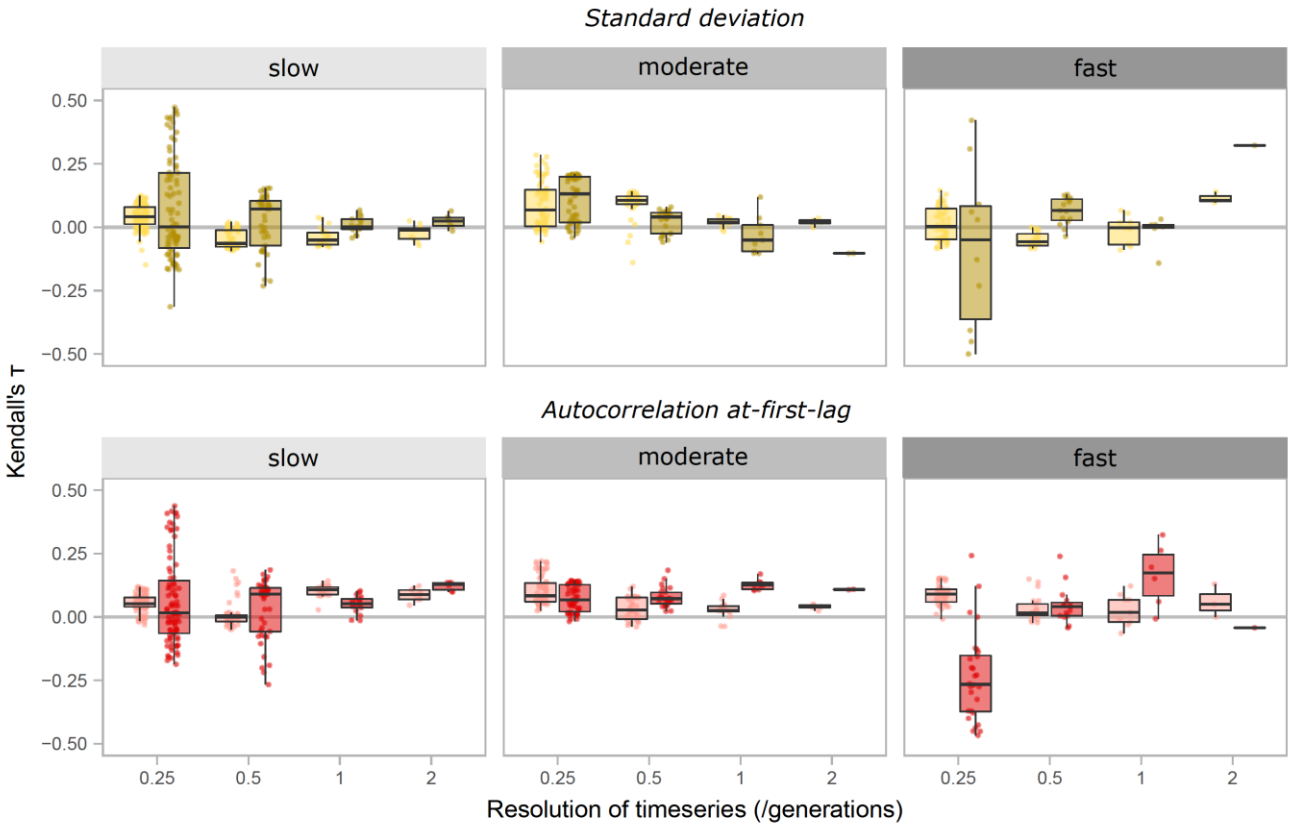
Figure 1. Boxplots of the strength of abundance-based early warning signals of population collapse across decreasing resolutions of sub-sampling in simulated populations subjected to collapse by harvesting. The data is split into subplots based on the bifurcation model simulated (fold, transcritical) and forcing level (slow, medium, fast). Boxplots are further split by metric used (yellow: standard deviation, red: autocorrelation at-first-lag). Each box represents the median Kendall's tau value (shown on the y-axis) across replicate time series across three forcing level with lengths ranging from 1,000 to 40 time steps. The x-axis shows the resolution of sampling in numbers of generations such that a value of 2 indicates that sampling was performed every 2 generations and 0.25 denotes sampling every quarter generation.

Figure 2. Performance of abundance-based early warning signals of population collapse across decreasing sampling time-series lengths in simulated populations subjected to collapse by harvesting. The data is split into subplots based on forcing intensity (slow, moderate, fast) and further split by metric used (yellow: standard deviation, red: autocorrelation at-lag-1) and bifurcation model simulated (solid: fold model, dashed: transcritical model). Each point represents (A) the mean Kendall's tau correlation coefficient or (B) rate of false negatives on the y-axis of 100 replicate simulations of population collapse. X-axis represents length of time series analyzed.

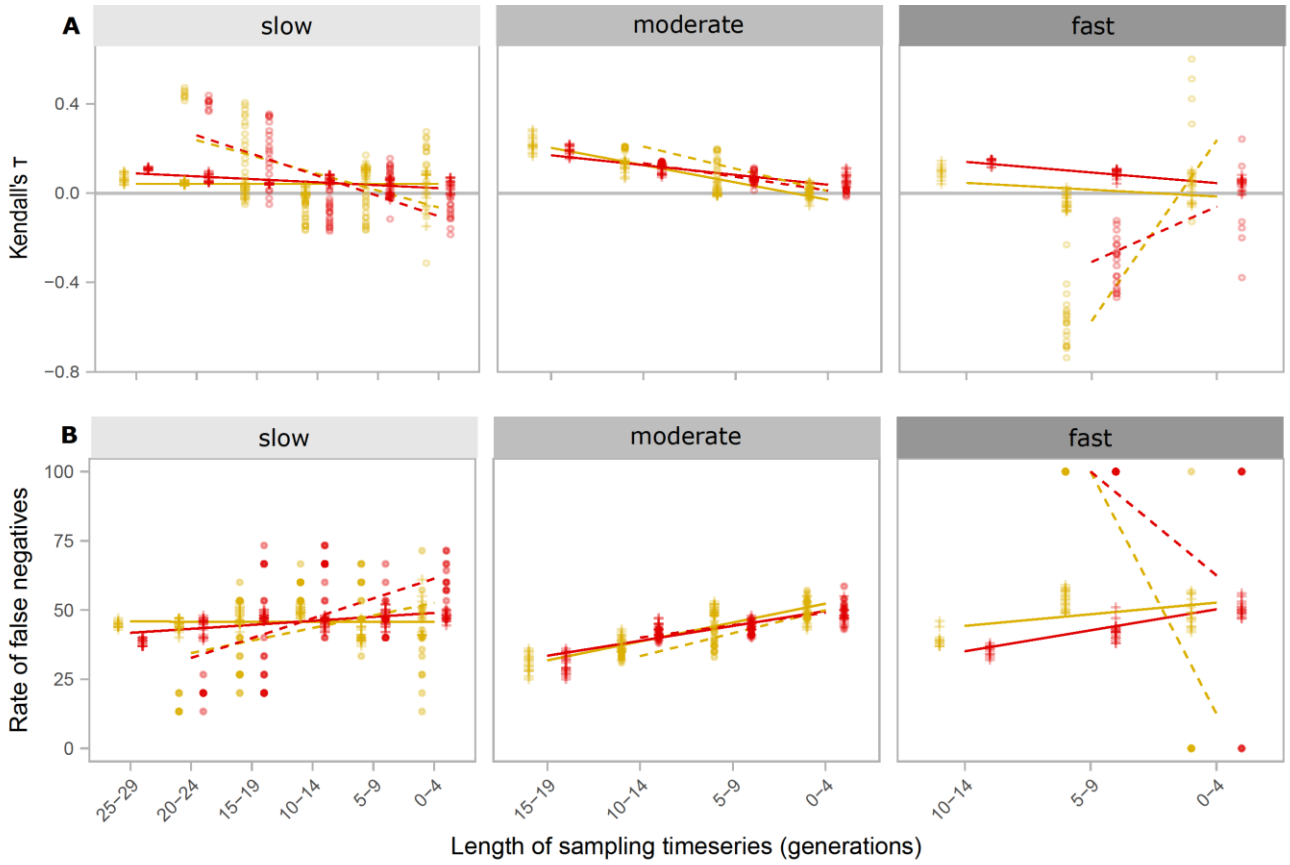
Figure 3. Performance of abundance-based and trait-based EWS of population collapse across decreasing sampling time-series resolutions for the experimental data of Clements & Ozgul 2016. Plots are organized into subplots based on the forcing intensity (slow, medium, fast) and further split by metric used (blue: standard deviation, purple: autocorrelation at-first-lag). Each box represents the

Kendall's tau value (on the y axis) for each resolution studied (on the x axis). Lighter colors represent classical abundance-based EWS while the darker colours represent trait-based EWS.

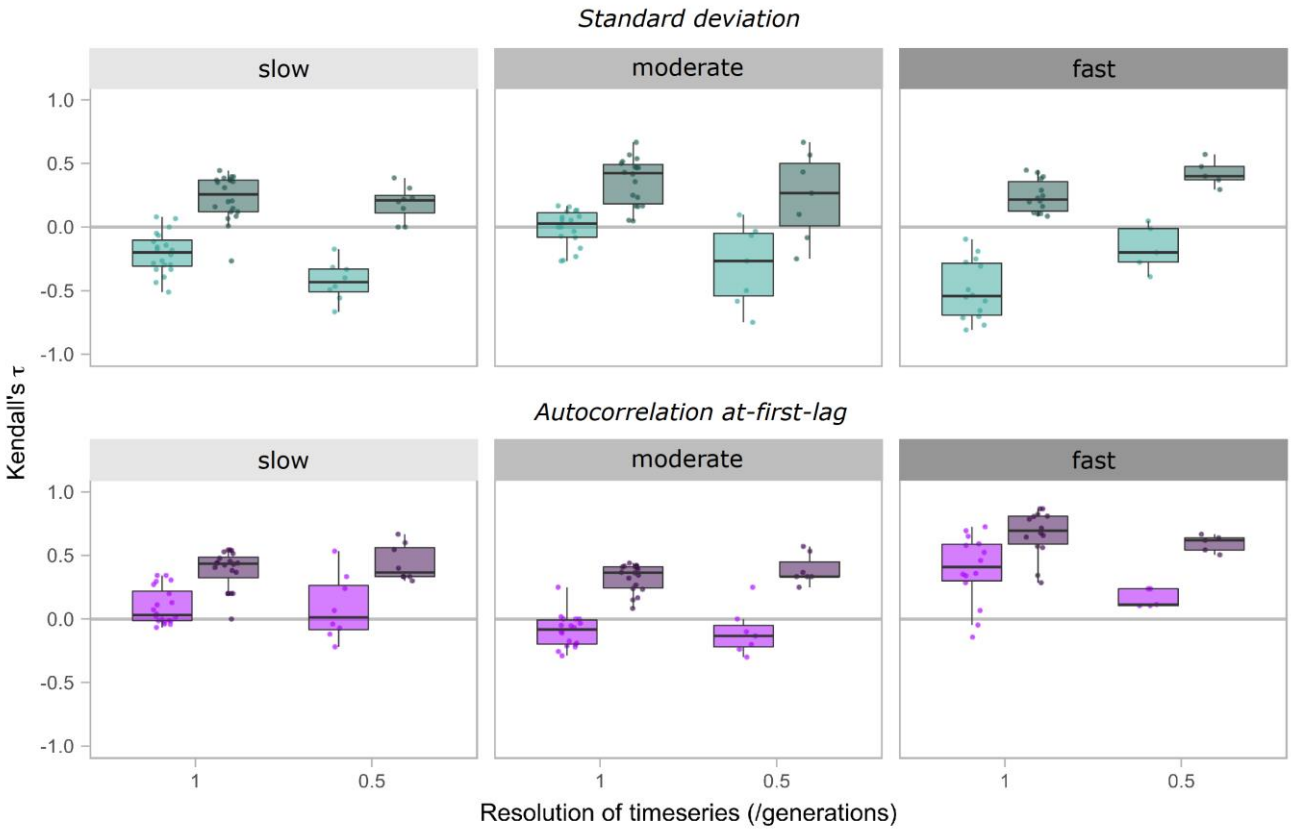
Figure 4. Performance of abundance-based and trait-based EWS of population collapse across decreasing sampling time-series lengths for the experimental data of Clements & Ozgul 2016. Plots are organized into subplots based on the forcing intensity (slow, medium, fast). X-axis is the length of time-series data analyzed and Y-axis denotes in (A) mean Kendall's Tau value and in (B) rate of false negatives. Solid lines indicate EWS with body size information added (trait-based EWS) while dotted lines indicate classical abundance-based EWS. Colours indicate metric used (blue: standard deviation, purple: autocorrelation at-first-lag). False negative plots were produced using loess smoothing.


Standard deviation
slow
moderate
fast
Autocorrelation at-first-lag
slow
moderate
fast
Kendall's τ
0.50
0.25
0.00
−0.25
−0.50
0.25
0.5
1
2
Resolution of timeseries (/generations)

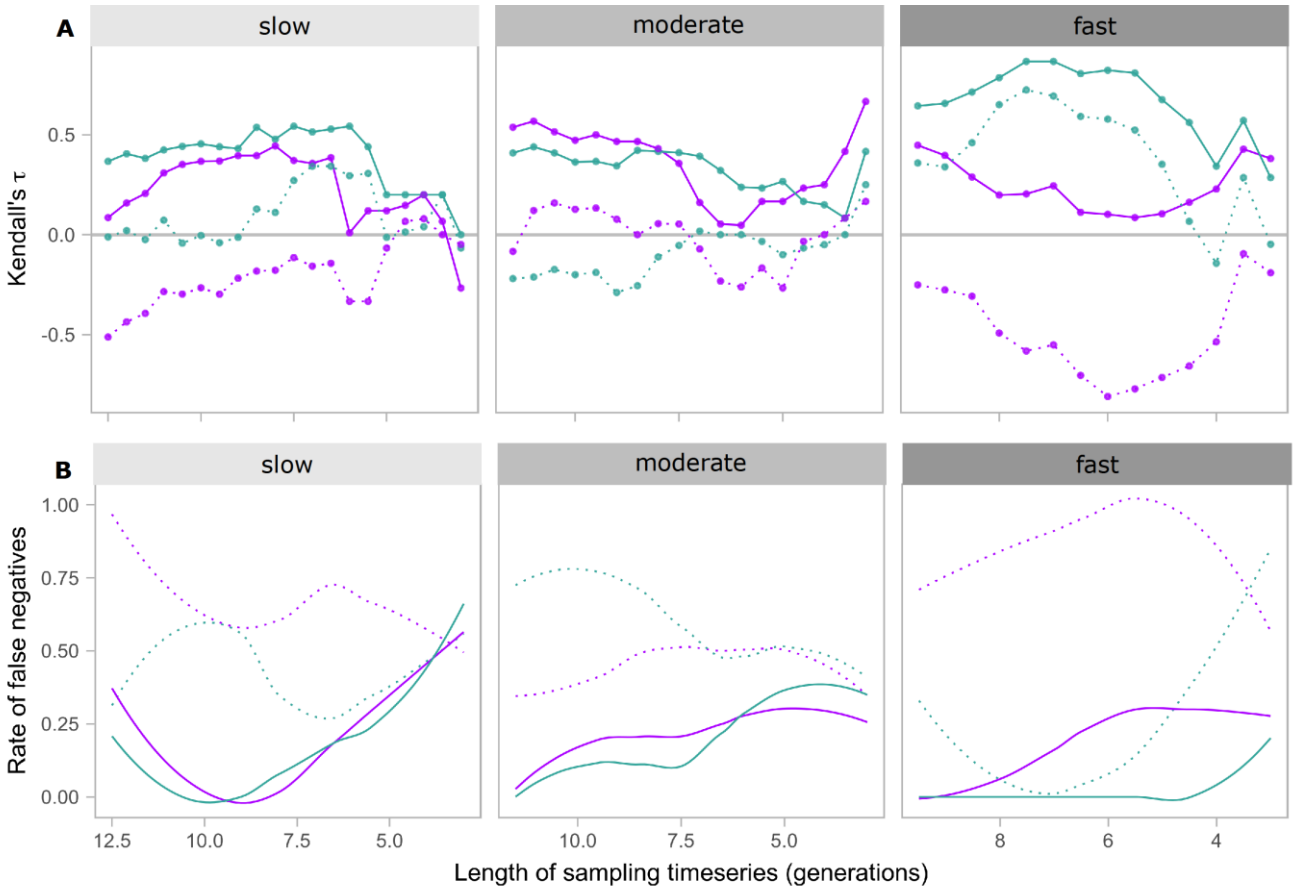
572


A
slow
moderate
fast
Kendall's τ
0.4
0.0
−0.4
−0.8
B
slow
moderate
fast
Rate of false negatives
100
75
50
25
0
25–29
20–24
15–19
10–14
5–9
0–4
15–19
10–14
5–9
0–4
10–14
5–9
0–4
Length of sampling timeseries (generations)

573


Standard deviation
slow
moderate
fast
Autocorrelation at-first-lag
slow
moderate
fast
Kendall's τ
1.0
0.5
0.0
-0.5
-1.0
1
0.5
Resolution of timeseries (/generations)

574


A
slow
moderate
fast
Kendall's τ
0.5
0.0
-0.5
B
slow
moderate
fast
Rate of false negatives
1.00
0.75
0.50
0.25
0.00
12.5
10.0
7.5
5.0
10.0
7.5
5.0
8
6
4
Length of sampling timeseries (generations)

575

Appendices

Table S1. Rate of false negatives from abundance-based early warning signals (EWS) of population collapse in simulated populations led to collapse through harvesting at three different rates: weak, medium, strong. The rate of false negatives reported is the number of false negative signals returned divided by the number of true positive signals. Results are further split by four time-series resolutions tested (a value of 2 indicates sampling was performed every 2 generations) and by the metric used as an EWS.

Table S1: Mean rates of false negatives (%) for varying forcing levels and timeseries resolutions

Strength of forcing	Slow				Moderate				Fast			
Resolution	2	1	50	25	2	1	50	25	2	1	50	25
Standard deviation	52.5	52.4	51.5	45.1	51.4	50.4	44.5	41.7	39.5	51	50.4	58.7
Autocorrelation at-lag-1	46.5	46.1	49.7	46.6	46.4	47.8	48	42.6	52.5	46.5	48.2	60.7