Supplementary Material: Minimum time required to detect population trends: the need for long-term monitoring programs

Easton R. White*

Address:
Center for Population Biology, University of California-Davis, Davis, California 95616 USA

*Corresponding author: eawhite@ucdavis.edu

Contents

Data and code for all the figures and tables can be found at (https://github.com/erwhite1/time-series-project). All analyses were run using R (R Core Team 2016).

S1 Detailed example of subsampling and power calculations

Here, I illustrate how we performed the subsampling and power calculations for a specific population. As an example, I examine a 35-year time series of Bigeye tuna (*Thunnus obesus*), one species in the Global Population Dynamics Database (NERC Centre for Population Biology 2010). Simple linear regression indicates a significant decrease for this population with an estimated slope coefficient of -0.0271717. I assume that this significant increase over 35 years is in fact the "true trend". In statistical jargon, the 35-year trend is an effect that is actually present; we can reject the null hypothesis of no trend. We can then use this as a benchmark to see if subsamples of the time series also indicate a significant increase.
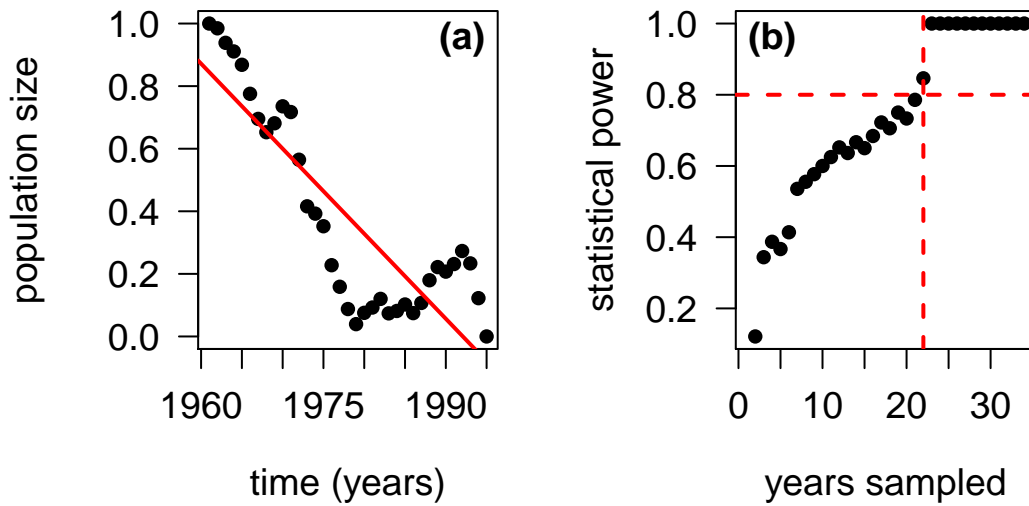

Figure A1: (a) Population size of Bigeye tuna (*Thunnus obesus*) over time. The line is the best fit line from linear regresssion. (b) Statistical power for different subsets of the time series in panel a.

I then perform a subsampling routine to estimate the minimum time required T_{min} (similar to Gerber, DeMaster, and Kareiva (1999)). This is the same routine I used for the results in the main manuscript.

1. We first extract all contiguous subsamples of the time series. This leads to 34 two-year subsamples, 33 two-year subsamples, and so forth until a single 35-year subsample.

2. For each subsample, we conduct linear regression and extract model coefficients and p-values.

3. We can call each set of subsamples, of the same length, a set. The proportion of subsamples within a set that show significant trends (significant slope coefficient α less than 0.05) is the statistical power. It is important to note that we only consider subsamples to be significant if they are significant in the same direction as the complete 35-year time series. In other words, we are conducting a one-tailed test.

4. We can then plot statistical power as a function of time series length (Fig. A1b). As expected, we can see that power increases with the more years that are sampled.

5. Then, we determine an appropriate level of statistical power that we find acceptable. Traditionally, this has been at 0.8, however, this is purely historical. Statistical power of 0.8 implies if a true trend is present, or there is a real change in abundance, then we will detect this trend 0.8 proportion of the time.

6. We then determine the minimum time series length (T_{min}) required to achieve that level of statistical power. Here, T_{min} is the first point in Fig. A1b where following points are also above 0.8. In this example, T_{min} is 22.

Therefore, a minimum of 22 years of continuous monitoring are required (for 0.8 statistical power at 0.05 significance level) to determine long-term changes in abundance.

S2 Additional results from the main manuscript

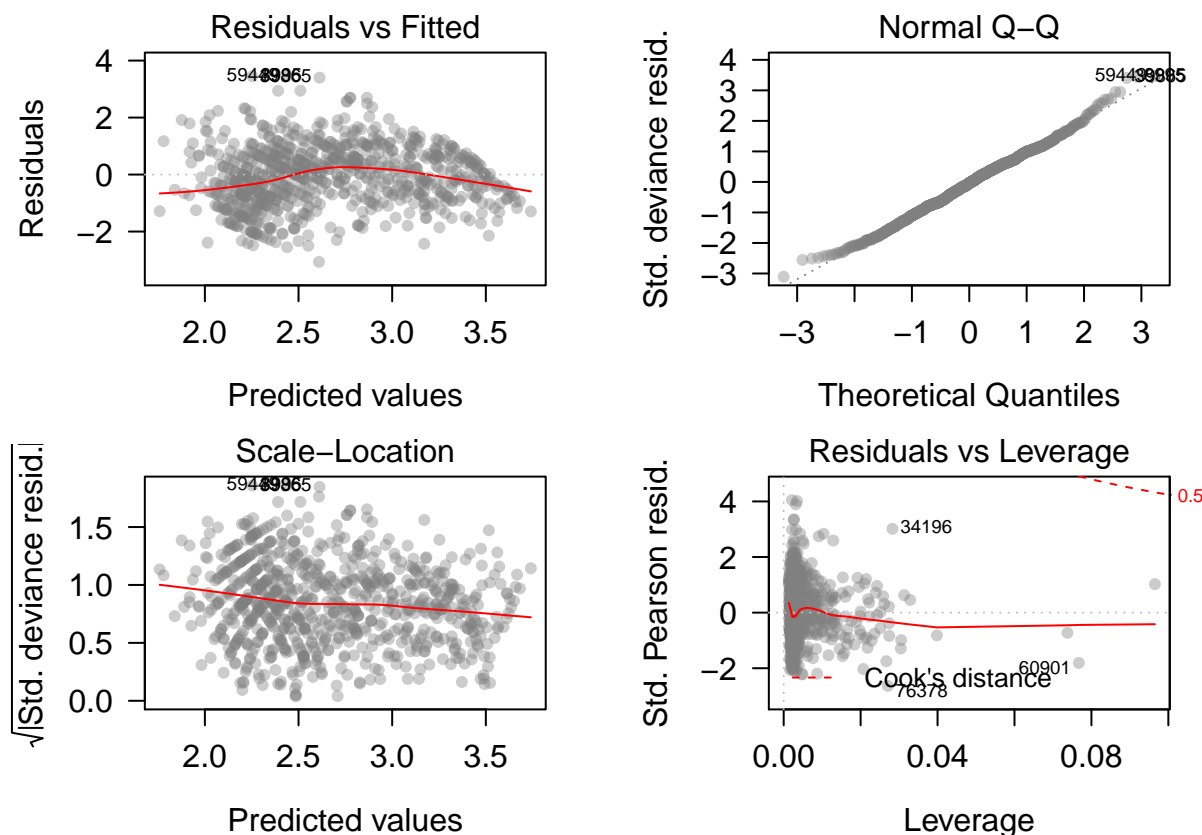## S2.1 Correlates of minimum time required

Figure A2: Output of generalized linear model with a Poisson error structure for predicting the minimum time required with explanatory variables of the absolute value of the slope coefficient (or trend strength), temporal autocorrelation, and coefficient of variation of population size.

In the main text, I explained how the minimum time required strongly correlated with the trend strength, temporal autocorrelation, and coefficient of variation in population size. Here, I use a generalized linear model framework with a Poisson error structure to determine explanatory variables of the minimum time required. I use the same 822 populations as in the main text. In figure A2, I show a set of residual plots for the regression analysis. I then show the coefficient estimates and levels of significance in table A1.

Table A1: Output of generalized linear model to examine time series characteristics as correlates of the minimum time required for determining long-term population trends.

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	4.2658002	0.0369853	115.337838	0.0000000

	Estimate	Std. Error	z value	Pr($>$|z|)
abs_overall_trend	-70.0413235	2.0027094	-34.973283	0.0000000
autocorrelation	0.0601596	0.0550282	1.093249	0.2742843
coefficient_variation	-0.1928445	0.0387392	-4.978022	0.0000006

Table A2: Output of generalized linear model to examine life-history trait correlates of the minimum time required for determining long-term population trends.

	Estimate	Std. Error	z value	Pr($>$|z|)
(Intercept)	2.8131074	0.1297488	21.681181	0.0000000
litter_or_clutch_size_n	0.0473894	0.0058279	8.131439	0.0000000
gen_len	0.0640424	0.0087382	7.328999	0.0000000
adult_body_mass_g	-0.0000809	0.0000177	-4.563594	0.0000050
maximum_longevity_y	-0.0074775	0.0019534	-3.827907	0.0001292
incubation_d	-0.0127828	0.0031681	-4.034864	0.0000546
trophic2	-0.2872678	0.1267799	-2.265878	0.0234589
trophic2.5	-0.3604948	0.1260791	-2.859274	0.0042461
trophic3	-0.2936987	0.1274698	-2.304064	0.0212190

Table A3: Output of generalized linear model to examine both time series characteristics and life-history trait correlates of the minimum time required for determining long-term population trends.

	Estimate	Std. Error	z value	Pr($>$|z|)
(Intercept)	4.3236120	0.1361628	31.7532519	0.0000000
abs(overall_trend)	-70.8497548	2.8424505	-24.9255898	0.0000000
coefficient_variation	-0.2089905	0.0531591	-3.9314122	0.0000844
autocorrelation	-0.0038698	0.0790400	-0.0489597	0.9609515
litter_or_clutch_size_n	0.0008944	0.0058450	0.1530248	0.8783787
gen_len	0.0087069	0.0088167	0.9875484	0.3233738
adult_body_mass_g	-0.0000214	0.0000174	-1.2311406	0.2182703
maximum_longevity_y	-0.0023131	0.0019647	-1.1773197	0.2390679
incubation_d	-0.0004509	0.0031244	-0.1443118	0.8852542
trophic2	0.0302480	0.1273238	0.2375677	0.8122164
trophic2.5	-0.0101150	0.1263433	-0.0800595	0.9361899
trophic3	-0.0049709	0.1276051	-0.0389552	0.9689261

The model with trend strength, autocorrelation, and the coefficient of variation in population size 75.13% of the variation in the minimum time required (Table A1).

For a subset of populations, I examined life-history correlates as potential explanatory variables for the minimum time required. Table A2 and A3 summarize results for a model with only life-history traits and an additional model with life-history traits and time series characteristics combined, respectively. The model with all life-history traits explained only 5.99% of the variation in the minimum time required (Table A2). The model with all life-history traits and time series characteristics explained 77.26% of the variation in the minimum time required (Table A3). When accounting for characteristics of the time series, no life-history variables were significant.

S2.2 Minimum time series and biological class

Minimum time required differed for different biological classes. For instance, more time is required for species within the Actinopterygii class compared to other species (Fig. A3a). These differences between biological classes can be explained by differences in population variability, with species in the Actinopterygii class experiencing larger inter-annual variability in population size (Fig. A3c). In addition, populations in the Actinopterygii class generally had weaker trends in abundance (i.e. smaller slope coefficients) compared to the Aves and Mammalia populations.
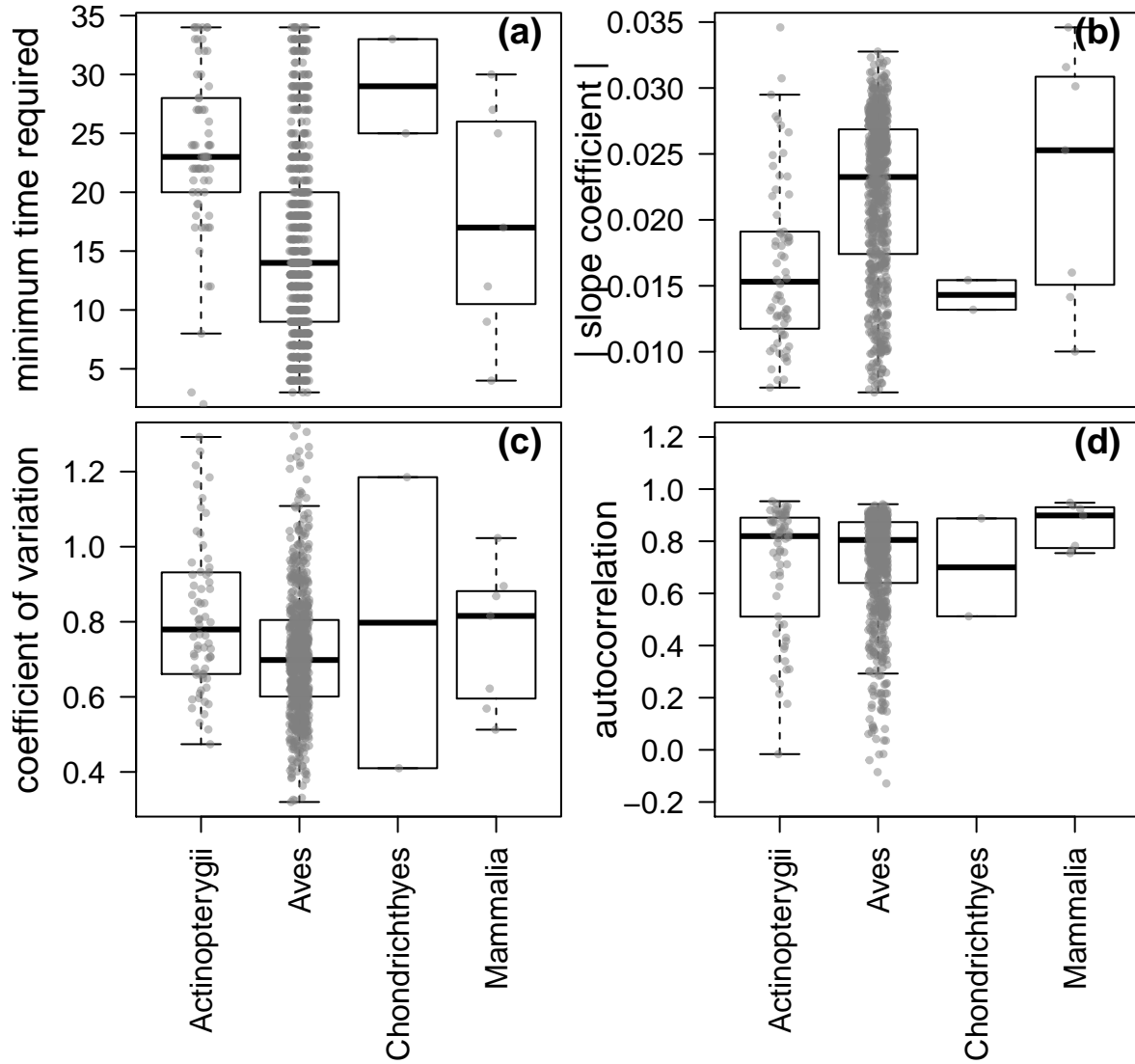

Figure A3: (a) Minimum time required to estimate change in abundance by biological class, (b) long-term trend (estimated slope coefficient) by class, (c) coefficient of variation in population size by class, and (d) temporal autocorrelation by class.

S2.3 Sensitivity analysis of significance level and power

Estimates of T_{min} depend strongly on the values used for the statistical significance level (α) and the probability of type II error (β), both of which are set by the practitioner. Again, statistical power is $1 - \beta$. Here we used simulations of the model described in the main manuscript. The model simulates linear trends in population abundance. We explored how estimates of T_{min} are affected by changes in each of these parameters. We see that the minimum time required increases with increases in statistical power or decreases with increases in the threshold for statistical significance (Fig A4).
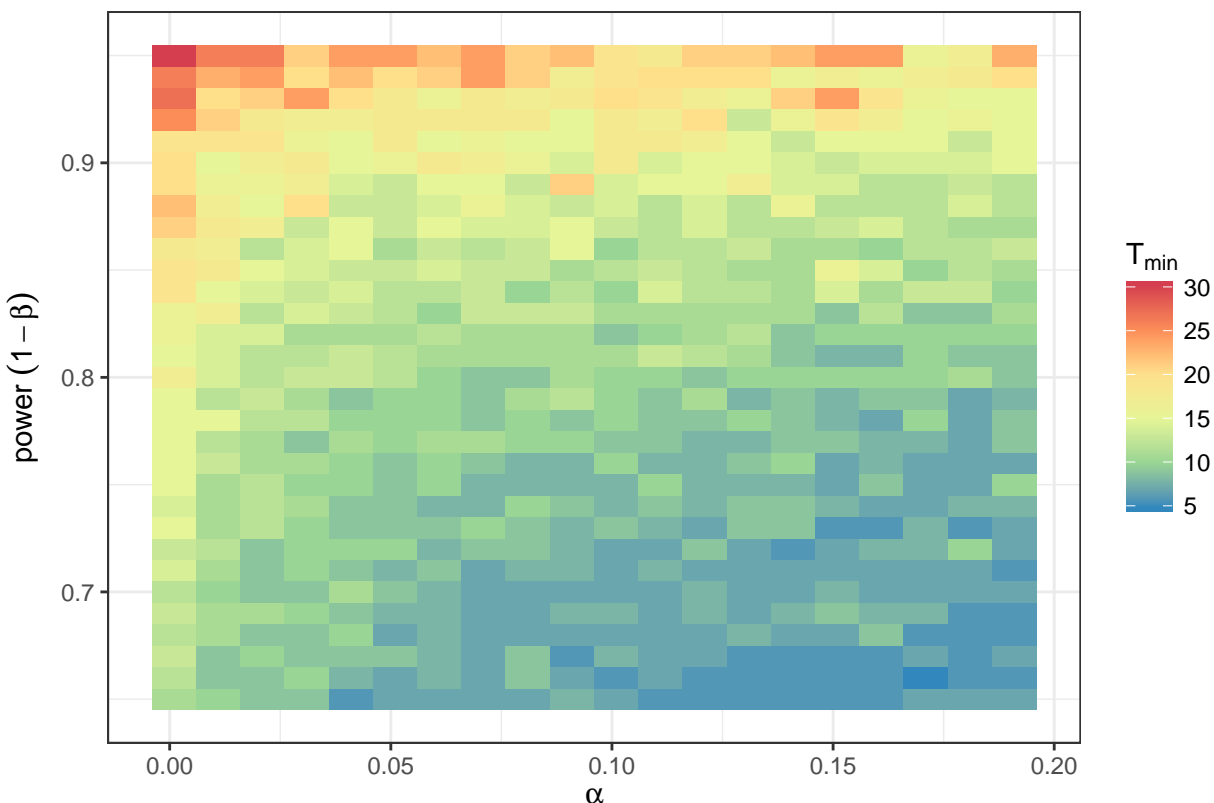

Figure A4: Minimum time required to assess long-term trends in abundance for values of statistical significance (α) and power ($1 - \beta$).

S2.4 Additional results for IUCN analysis

In the main text, I presented results detailing the difference between the minimum time required to achieve a high level of statistical power and the number of years suggested by IUCN based on generation length to assess a species as vulnerable. There were deviations between the two approaches. These deviations were mostly explained by the differences in inter-annual population variability (coefficient of variation) and generation time (Fig. A5).
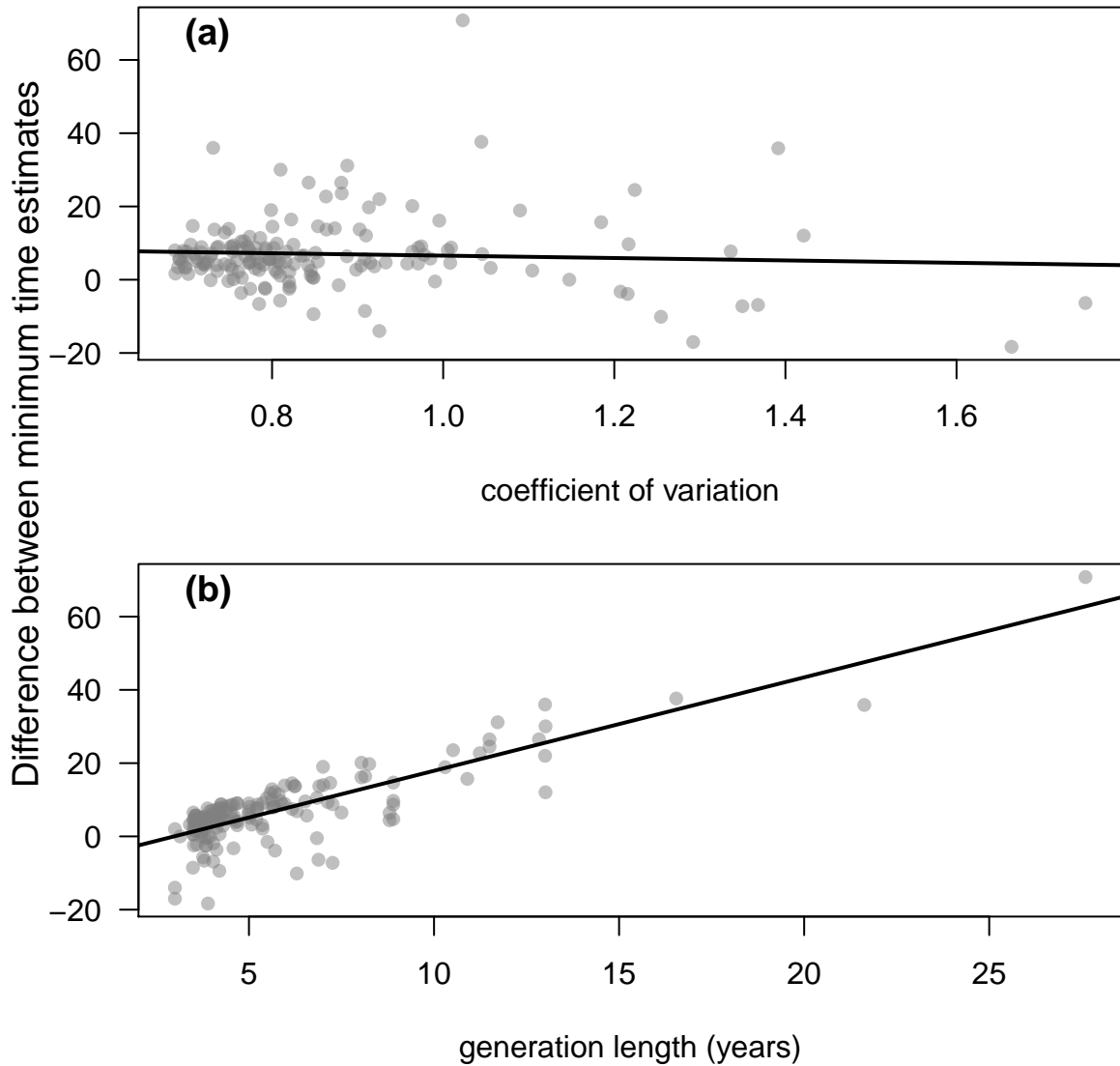
Figure A5: The difference between minimum time estimates is the minimum time required to achieve 0.8 statistical power versus the minimum time required under IUCN criteria A2 to classify a species as vulnerable. Each point represents a single population, all of which saw declines of 30% or greater over a 10 year period. (a) Difference between minimum time estimates versus the coefficient of variation in population size. (b) Difference between minimum time estimates versus the generation length in years.

S3   Minimum time calculations testing exponential growth

In the main text, I evaluated the minimum time required to determine long-term trends in abundance via linear regression. This process examined linear trends in abundance over time. Here I examine the minimum time required to estimate long-term trends that are either exponential growth or decay. I use the same methods as described in the previous sections, but we take the *log* of population density, or abundance.
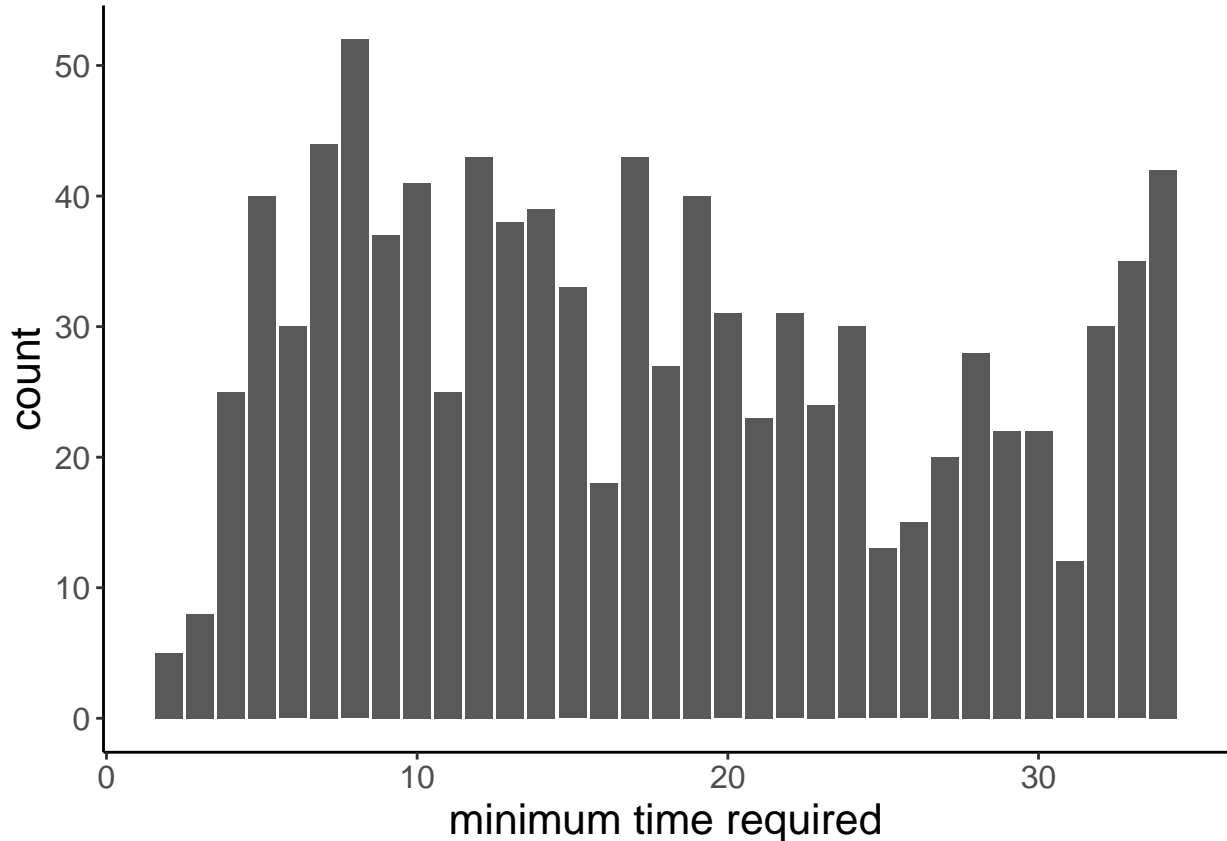

Figure A6: Distribution of the minimum time required in order to detect a significant trend (at the 0.05 level) in log(abundance) given power of 0.8.

We see that the distribution of T_{min} is almost identical to that in the main manuscript (Fig. A6). This is perhaps not surprising as most time series that would significantly increase or decrease linearly would probably also significantly increase or decrease at an exponential rate. Further, the calculations here and in the main manuscript both use linear regression. Therefore both calculations estimate the same number of parameters.

S4 Simulations with more complicated population model

In the main text, I showed how a simple population model could be simulated repeatedly to estimate the power obtained with time series of increasing length (Bolker 2008; Johnson et al. 2015). The model in the main text simulated linear population growth with only a slope coefficient, y-intercept, and noise parameter required. This model is purely phenomenological and does not include any species life-history information. Here, I use the same routine as the main text, but simulate from a more biologically-realistic population model. I use the model described in White, Nagy, and Gruber (2014). The model is a stochastic, age-structured population model that includes density-dependence for lemon sharks (*Negaprion brevirostris*) in Bimini, Bahamas.
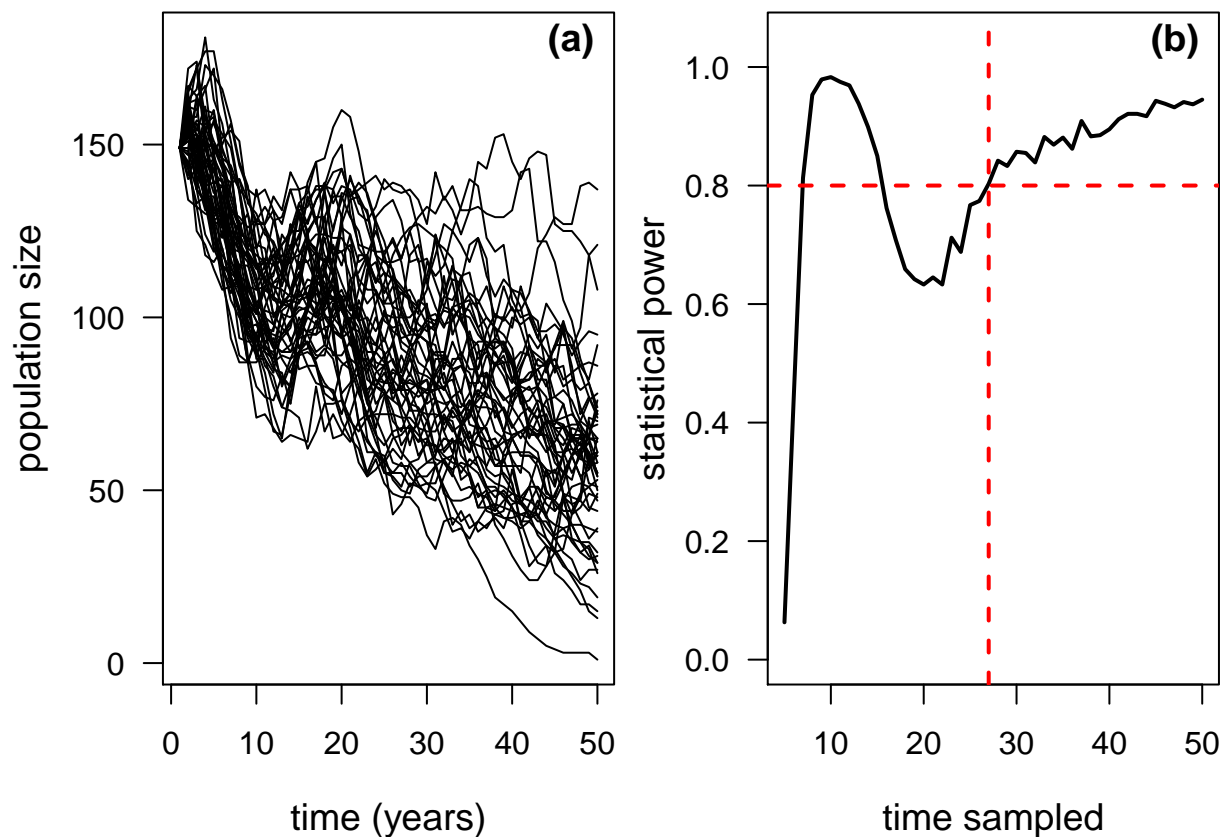

Figure A7: Statistical power for different length of time series simulations for a lemon shark population in Bimini, Bahamas.

Here, I parameterize the model for a situation where adult mortality rate is high enough to cause a population decline. As in the main text, I simulated this more biologically-realistic model for different lengths of time. For each length of time, I calculated the statistical power. Similar to results with the simpler model, statistical power generally increases with longer sampling time (Fig. A7). In this example, the minimum time required (T_{min}) t0 obtain at least 0.8 statistical power, given a significance level of 0.05, is 27 years.

S5   Using Generalized additive model to identify significant trends

In the main text, I examined the minimum time required to identify a trend in abundance via linear regression. This approach allowed us to identify increases or decreases, but a linear model may not always be a good fit. Generalized additive models (GAMs) are more general than general linear models and allow more flexibility (Wood 2006). GAMs are models where a response variable depends on unknown smooth functions of explanatory variables. GAMs, therefore, can identify relationships between response variables and explanatory variables that are non-linear and perhaps more complicated. The downside of GAMs is they typically require more data and are also prone to overfitting.

Here, I conduct the same analyses in the main text, but instead calculate the minimum time series required to detect trends over time according to a GAM model. I hypothesized that GAMs should require less time to detect a trend as they are more flexible than linear regression. I provide an example that shows statistical power increases with more time sampled (Fig. A8).
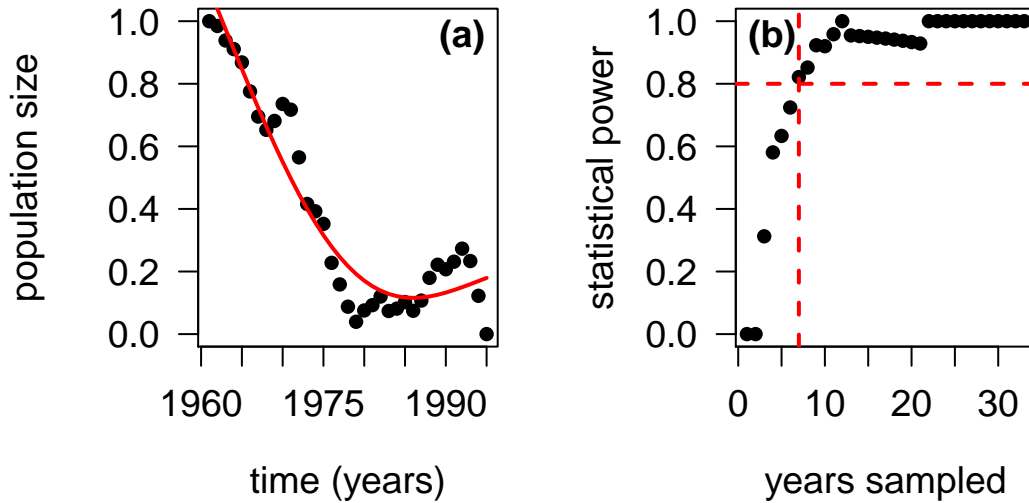

Figure A8: (a) Time series for Bigeye tuna (*Thunnus obesus*) with corresponding fitted GAM model in red (with a smoothing parameter of 3) and (b) statistical power as a function of the number of years sampled. The horizontal ine at 0.8 indicates the minimum threshold for statistical power and the vertical line denotes the minimum time required to achieve 0.8 statistical power.

We then fit GAM models for 799 populations. We found a similar distribution of minimum time required as in the main text for linear regression (Fig. A9). However, in line with our hypothesis, the GAM models did result in a lower mean minimum time required of 14.12 years compared to the results from the main text of 15.91 years.
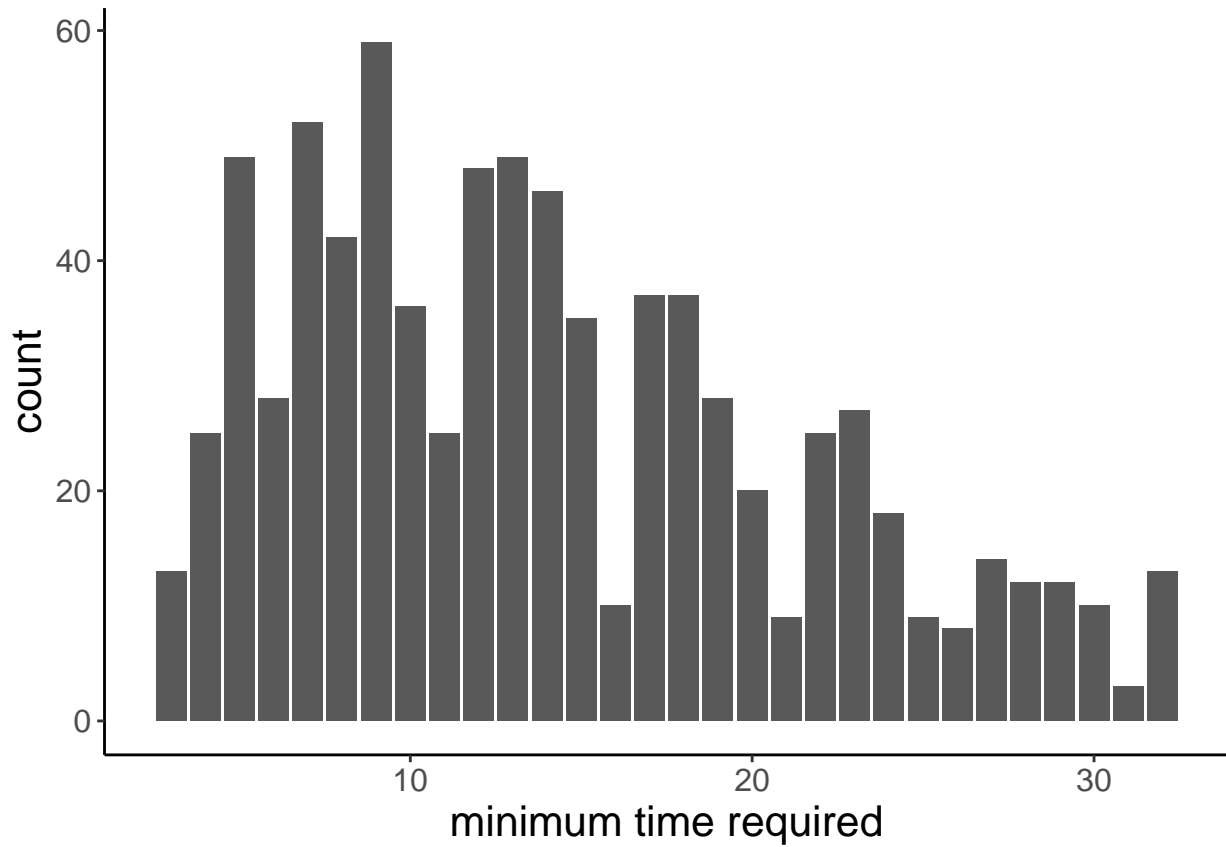
Figure A9: Distribution of the minimum time required in order to detect a significant trend (at the 0.05 level) in abundance according to a GAM model given statistical power of 0.8. The smoothing parameter was set to 3 for each population.

References

Bolker, Benjamin M. 2008. *Ecological Models and Data in R*. 1st ed. Princeton, New Jersey: Princeton University Press.

Gerber, L R, D P DeMaster, and P M Kareiva. 1999. "Gray whales and the value of monitering data in implementing the U.S. endangered species act." *Conservation Biology* 13 (5): 1215–9.

Johnson, Paul CD, Sarah JE Barry, Heather M Ferguson, and Pie Müller. 2015. "Power analysis for generalized linear mixed models in ecology and evolution." *Methods in Ecology and Evolution* 6 (2): 133–42. doi:10.1111/2041-210X.12306.

NERC Centre for Population Biology, Imperial College. 2010. "The Global Population Dynamics Database Version 2."

R Core Team. 2016. "R: A language and environment for statistical computing." Vienna, Austria: R Foundation for Statistical Computing. https://www.r-project.org/.

White, Easton R, John D Nagy, and Samuel H Gruber. 2014. "Modeling the population dynamics of lemon sharks." *Biology Direct* 9 (23): 1–18.

Wood, S.N. 2006. *Generalized Additive Models: An Introduction with R*. Chapman Hall/CRC.