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

Easton R. White

Center for Population Biology, University of California, Davis, California 95616 USA

eawhite@ucdavis.edu

Abstract

Long-term time series are necessary to better understand population dynamics, assess species' conservation status, and make management decisions. However, population data are often expensive, requiring a lot of time and resources. What is the minimum population time series length required to detect significant trends in abundance? I first present an overview of the theory and past work that has tried to address this question. As a test of these approaches, I then examine 822 populations of vertebrate species. I show that 72% of time series required at least 10 years of continuous monitoring in order to achieve a high level of statistical power. However, the large variability between populations casts doubt on commonly used simple rules of thumb, like those employed by the IUCN Red List. I argue that statistical power needs to be considered more often in monitoring programs. Short time series are likely under-powered and potentially misleading.

Keywords: ecological time series, experimental design, population monitoring, statistical power, sampling design

Introduction

Observational studies and population time series have become a cornerstone of modern ecological research and conservation biology (Magurran *et al.* 2010, Peters 2010, Hughes *et al.* 2017). Long-term data are necessary to both understand population dynamics and to assess species extinction risk. Some time series may now be considered "long-term" (e.g. continuous plankton recorder, Giron-Nava *et al.* (2017)), but most are still short. Time series are typically short due to short funding cycles and typical experimental time-frames (Field *et al.* 2007, Hughes *et al.* 2017).

How long of a time series is actually necessary? This question has important implications for both research and management (Nichols & Williams 2006). A short time series may lead to wrong conclusions given large natural year-to-year variability (McCain *et al.* 2016). Managers need to know when action is needed for a population. Therefore, managers must understand when a population trend over time is actually meaningful. In addition, sampling is typically expensive, therefore, we also do not want to sample for longer than is necessary. For example, Gerber *et al.* (1999) investigated the minimum time series required to estimate population growth of the endangered, but recovering, eastern North Pacific gray whale (*Eschrichtius robustus*). They used a long-term census to retroactively determine the minimum time series required to assess threat status. They found that only 11 years were needed, eight years before the delisting decision was made. This highlights the importance of estimating the minimum time series required as an earlier decision would have saved time and money (Gerber *et al.* 1999). Further, waiting too long to decide an action can imperil a species where management action could have been taken earlier (Martin *et al.* 2012, 2017). Specific guidelines are therefore needed to determine when a time series is adequate. For example, the International Union for Conservation of Nature (IUCN) Red List Categories and Criteria suggest, under Criterion A2, a species qualifies as vulnerable if it has experienced a 30% decline over 10 years, or three generations (IUCN 2012).

Past work has investigated questions related to the minimum time series required to estimate trends in population size over time (Wagner *et al.* 2009, Giron-Nava *et al.* 2017). For example, Rhodes & Jonzen (2011) examined the optimal allocation of effort between spatial and temporal replicates. Using simple populations models, they found that the allocation of effort depends on environmental variation, spatial and temporal autocorrelation, and observer error. Rueda-Cediel *et al.* (2015) also used a modeling approach, but parameterized a model specific for a threatened snail, *Tasmaphena lamproides*. They found that for this short-lived organism, 15 years was adequate to assess long-term trends in abundance. However, these studies, and other past work, have typically been only on theoretical aspects of monitoring design or focused on a few species.

Statistical power is not a new tool (Cohen 1992, Thomas 1997, Thomas & Krebs 1997, Gibbs *et al.* 1998), but it is still under-appreciated in ecological research (Legg & Nagy 2006). Therefore, I begin by reviewing key concepts of power analyses in relation to time series analysis. I then explain how simulation approaches have been used to estimate the minimum time required to estimate long-term population trends. Lastly, I take an empirical approach to estimate the minimum time required for 822 animal populations.

Statistical power in time series analyses

For any particular experiment four quantities are related to one another: significance level (α), statistical power, effect size, and sample size (Thomas 1997, Legg & Nagy 2006). The exact relationship between these quantities depends on the specific statistical test. If a time series was assessed as significantly increasing or decreasing—when there was no true significant trend—this would be a false positive. The false positive rate, or significance level (α), is often set at 0.05; although this is purely historical (Mapstone 1995). A type II error (β) is a failure to detect a true trend in abundance over time. Statistical power ($1-\beta$) is then one minus the probability of a type II error (β). Informally, statistical power is simply the probability of detecting a trend given it actually exists. The effect size is a estimate of the strength of a particular phenomenon.

Prior to an experiment, one could set appropriate levels of power, significance level, and the effect size to estimate the sample size required for the experiment. This approach, however, is not straight-forward for a time series, or more complicated scenarios (Johnson *et al.* 2015), as data are clearly non-independent.

In the context of time series data, sample size can be the number of study sites surveyed, frequency of surveys per year, and the number of years surveyed. For example, Gibbs *et al.* (1998) examined how many times within a year a population needs to be sampled to ensure high statistical power. They found that the sampling intensity within a year differed greatly depending on species, because of differences in population variability. I use a similar approach, but instead focus on the number of years required to estimate trends in abundance. In line with Gibbs *et al.* (1998), I would expect these results to be strongly dependent on population variability. Unlike Gibbs *et al.* (1998), I do not lump species together, and instead study the differences between, and within, species.

Simulation approach

One approach to determining the minimum time series length needed is through repetitive simulations of a population model (Gerrodette 1987, Gibbs *et al.* 1998). This is the same approach one might use in sample size calculations for any experimental design too complicated for simple power analyses (Bolker 2008, Johnson *et al.* 2015). Essentially, a population model, with a selected set of parameter values, is simulated repetitively for a number of years. As an example, we can take the following population model for population size N at time t:

$$N(t+1) = N(t) + r(t) + \epsilon \text{ with } \epsilon \sim N(\mu, \sigma) \tag{1}$$

where ϵ is a normally-distributed random noise term with mean μ and standard deviation σ. The rate of growth (r) is the trend strength of the increase or decrease (i.e. the estimated slope from linear regression). Although there are many approaches to studying populations trends (Thomas 1996), linear regression is the simplest and most commonly applied.

With regard to detecting time series trends, statistical power is the proportion of simulations where the slope parameter from linear regression is significant at the 0.05 threshold. Statistical power of 0.8 would indicate that, if there was a true trend in abundance, there would be a 0.8 probability of detecting the trend. Values of 0.05 for the significance level and 0.8 statistical power are purely historical (Cohen 1992). Thus, it is important to also examine the effect of changing these values (Fig. A4). Predictably, as the significance level decreases or the power required increases, more years of sampling are required (Fig. A4).

I set the significance level at 0.05 and then simulated the model in equation 1 (Fig. 1a). Statistical power increases with increases in the length of time sampled (Fig. 1b). Where power is greater than 0.8 (the dotted line), that is the minimum time required (T_{min}) to be confident in the detection of a long-term trend in abundance. As shown previously (Rhodes & Jonzen 2011, Rueda-Cediel *et al.* 2015), statistical power increases with larger trend strength and lower population variability (Fig. 1c,d). Simulation approaches can be useful before designing a monitoring program or when a realistic model exists for the population in question.

It is important to note that any population model could be used here (see example in Fig. A6). Ideally, the specific model choice should be tailored to the population of interest. As an example, I determined the minimum time required to estimate long-term population trends using a stochastic, age-structured model of lemon shark population dynamics in the Bahamas

(White *et al.* 2014). I found that over 27 years of continuous monitoring were needed in this particular scenario (Fig. A7). Not surprisingly, the minimum time required for the lemon shark population was strongly dependent on model parameters (see Fig. A7). Similarly, Rueda-Cediel *et al.* (2015) used a matrix model parameterized for a particular snail species. They used the model to argue that only 10-15 years were needed to accurately assess trends in abundance.

Empirical approach

As an empirical test of these ideas, I used a database of 2444 population time series compiled by Keith *et al.* (2015). The data are originally from the Global Population Dynamics Database (NERC Centre for Population Biology Imperial College 2010) and several other sources (Keith *et al.* 2015). I filtered out short time series (less than 35 years), and those with missing data, leaving 822 time series. The data includes information on 477 vertebrate species of birds ($n = 747$), mammals ($n = 7$), sharks ($n = 2$), and bony fish ($n = 66$).

I assumed that each time series was long enough to include all necessary information (e.g. variability) about the population. In other words, each time series was a representative sample. I then performed a type of 'retrospective' power analysis; termed retrospective because the data has already been collected (Thomas 1997). I first took all possible contiguous subsamples of each time series. For example, a time series of 35 years had 34 possible contiguous subsamples of length 2, 34 possible contiguous subsamples of length 3, and continuing until 1 possible contiguous subsample of length 35 (Gerber *et al.* 1999, Brashares & Sam 2005, Giron-Nava *et al.* 2017). The subsampling avoids some of the common pitfalls of retrospective power analyses (Thomas 1997, Thomas & Krebs 1997). In line with the simulation approach, I determined the proportion of subsamples of a particular length that had estimated slope coefficients statistically different from zero. This proportion is a measure

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3168v4 | CC BY 4.0 Open Access | rec: 29 Jul 2018, publ: 29 Jul 2018

of statistical power. Lastly, I determined which subsample length is required to achieve a certain threshold of statistical power (0.8, Cohen (1992)). The minimum subsampled length that met these criteria was the minimum time series length required (T_{min}). All analyses were conducted in R (R Core Team 2017).

Estimates of the minimum time required

Across all the populations, I found an average minimum time series length required (T_{min}) of 15.9 (SD=8.3), with a wide distribution (Fig. 2b). Approximately, 72% of populations required at least 10 years of monitoring. Estimates of T_{min} varied between biological class (Fig. 2a). Ray-finned fish (class Actinopterygii) typically had estimates of T_{min} over 20 years. Birds (class Aves) had a much wider distribution of T_{min}, but usually required less years of sampling. Differences between these classes can be explained by differences in variability in population size and strength of trends in abundance (Fig. A3).

This time-frame is in line with past work on a short-lived snail species (Rueda-Cediel *et al.* 2015) and a long-lived whale species (Gerber *et al.* 1999). Hatch (2003) used seabird monitoring data to estimate minimum sampling requirements. He found that the time required ranged from 11 to 69 years depending on species, trend strength, and study design. All of this past work has been limited to a small number or species. This manuscript is the first attempt to document the minimum sampling requirements for such a wide diversity and number of species.

Correlates for minimum time required

The minimum time series length required was strongly correlated with trend strength (i.e. estimated slope coefficient from linear regression), coefficient of variation in population size, and autocorrelation in population size (Figs. 3a-c). All three of these explanatory variables were significant and had large effect sizes (see Table A1). Combined, trend strength, coefficient of variation in population, and autocorrelation account for 75.1% of the explained deviance (Zuur *et al.* 2009) in minimum time series length required. Thus, by knowing these three aspects of a time series, a reasonable estimate for the minimum time series length required can be made.

There was life-history information available (Myhrvold *et al.* 2015) for 547 populations representing 315 different species, all of which were birds (Aves class). Some life-history traits were significant predictors for the minimum time required (Figs. 3d-h, Tables A2,A3). However, even combined, all five of the life-history traits accounted for only 5.99% of the explained deviance in minimum time series length required. In addition, when accounting for trend strength, coefficient of variation, and autocorrelation, no life-history traits were significant predictors of the minimum time required (Table A3). These results were not altered when taking phylogeny into account (Fig. A10).

I initially hypothesized that species with longer lifespans or generation times may require a longer sampling period. This result could have been a result of at least two factors. First, the data I used may not include a diverse enough set of species with different life-history traits. Second, the question I posed, whether a population is increasing or decreasing, was specifically concerned with population trends over time. Therefore, life-history characteristics may be more important for other questions more closely tied to species biology. For example, Blanchard *et al.* (2007) used detailed simulations of spatially-distributed fisheries to compare survey designs. They found that statistical power depended on temperature preferences and the degree of population patchiness, presumably because the survey designs included a spatial component.

Evaluating the IUCN criteria

I examined a subset of populations with observed declines of 30% or greater over 10 years, qualifying all of them as vulnerable under IUCN Criterion A2 (IUCN 2012). This resulted in n = 162 populations. I then compared the minimum time required to achieve 0.8 statistical power (T_{min}) to the minimum time required under the IUCN criteria (Fig. 4). For populations below the identity line in figure 4, IUCN criteria would require more sampling compared to estimates for T_{min}. Further, populations above the identity line are cases where the IUCN criteria would classify a population as vulnerable despite not having sampled enough years to achieve high statistical power (Fig. 4). The silhouettes on figure 4 highlight that species with long generation times had larger discrepancies between T_{min} and the minimum time required for IUCN assessments (Fig. A5).

For the populations I examined, the IUCN criteria may be overly simplistic (Fig. 4). For many populations, the IUCN criteria suggest more years than necessary are required to assess a population as vulnerable. Conversely, for other populations the IUCN criteria suggest sampling times that are less than the minimum time required for statistical power. This suggests that the IUCN criteria are probably too simplistic as the minimum time required does not correlate with generation time or other biological covariates (Fig. 3d-h). Instead, assessments of long-term trends could rely on one of two approaches. First, a specific model could be built and simulated for the species of interest. An estimate of the minimum number of years for a particular threshold of statistical power could then be determined (see example in Fig. A7). Alternatively, if one had estimates of the population trend, the population variability, and the autocorrelation then it would be possible to estimate the minimum number of years required using the regression model provided in Table A1.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3168v4 | CC BY 4.0 Open Access | rec: 29 Jul 2018, publ: 29 Jul 2018

Related questions

Keith *et al.* (2015) studied the same dataset to determine how predictive a current population trend was of future trends. They showed that for many species (except birds), past declines were actually more predictive of subsequent increases. This shows the non-linear nature of many time series. They do not explicitly determine the minimum time required for a population trend to predict a longer term term. Therefore, this manuscript adds to their work by determining the minimum number of years required to determine a population trend. Many populations require far greater than the IUCN rule of 10 years (or 3 generations) employed by Keith *et al.* (2015).

An important related idea is the optimal allocation of sampling effort in space versus time. In a theoretical investigation of this question, Rhodes & Jonzen (2011) found that the optimal allocation of sampling depended strongly on temporal and spatial autocorrelation. If spatial population dynamics were highly correlated, then it was better to sample more temporally, and vice versa. The empirical data supports this idea as populations with strong temporal autocorrelation needed less years of sampling (Fig. 3). Morrison & Hik (2008) also studied the optimal allocation of sampling effort in space versus time, but used emprical data from a long-term survey of the collared pika (*Ochotona collaris*) in the Yukon. They found that surveys less than 5 years may be misleading and that extrapolating from one population to another, even when nearby geographically, may be untenable.

Seavy & Reynolds (2007) asked whether statistical power was even a useful framework for assessing long-term population trends. They examined 24 years of census data on Red-tailed Tropicbirds (*Phaethon rubricauda*) in Hawaii. They always had 0.8 statistical power to detect a 50% decline over ten years. Therefore, they cautioned against only using power analyses to design monitoring schemes and instead argued for metrics that would increase precision: improving randomization, reducing bias, and increasing detection probability. Power analyses should not be the only consideration when designing monitoring schemes. However, unlike

Seavy & Reynolds (2007), the results here show that at least 10 years of monitoring were required for 72% of populations. Further, 30.7% of populations required at least 20 years of monitoring.

Limitations

This paper has some limitations in determining the minimum time series length required. First, the minimum time estimated is particular to the specific question of interest. Specifically, here I examined the minimum time required to determine if a long-term linear population trend exists. The minimum time required would differ if one was interested in examining non-linear trends, assessing multiple populations, or answering a different question. The empirical approach presented here was also limited to only 477 populations of vertebrate species. An additional complication is that the subsampling of the full time series allows for estimates of power, but the individual subsamples are clearly not independent of one another. In an ideal setting, a specific population model would be parameterized for each population of interest. Then, model simulations could be used to estimate the minimum time series required to address each specific question of interest. Clearly, this is not always practical, especially if conducting analyses for a wide array of species.

Conclusions

Power analyses are not a novel aspect of ecological research (Thomas & Krebs 1997, Legg & Nagy 2006). However, power analyses are still underutilized, especially in the context of time series analyses. This is the first paper to study such a large number of populations (n=822) to determine the minimum time series length required to detect population trends. This goes beyond previous work that either focused on theoretical investigations or a limited number of species. I show that to identify long-term changes in abundance, on average 15.91 years of continuous monitoring are often required (Fig. 2). More importantly, however, there is wide distribution of estimated minimum times. Therefore, it is not wise to use a simple threshold number of years in monitoring design. Further, contrary to my initial hypotheses, minimum time required did not correlate with generation time or any other life-history traits (Fig. 3d-h). These results argue against overly simplified measures of minimum sampling time based on generation length or other life-history traits, like those of the IUCN criteria (Fig. 4). Instead, simulation models or power analyses should be tailored to particular populations.

The design of monitoring programs should include calculations of statistical power, the allocation of sampling in space versus time (Larsen *et al.* 2001, Rhodes & Jonzen 2011), and metrics to increase precision (Seavy & Reynolds 2007). Ideally, a formal decision analysis to evaluate these different factors would be conducted to design or assess any monitoring program (Hauser *et al.* 2006, McDonald-Madden *et al.* 2010). This type of formal decision analysis would also include information on the costs of monitoring. These costs include the actual costs of sampling (Brashares & Sam 2005) and the ecological costs of inaction (Thompson *et al.* 2000).

For many populations, short time series are probably not reliable for detecting population trends. This result highlights the importance of long-term monitoring programs. From both a scientific and management perspective, estimates of the minimum time required are important. If a time series is too short, we lack the statistical power to reliably detect population trends. In addition, a time series that is too long may be a poor use of already limited funds (Gerber *et al.* 1999). Further, more data is not always best in situations where management actions need to be taken (Martin *et al.* 2012, 2017). When a population trend is detected, it may be too late for management action. In these situations, the precautionary principle may be more appropriate (Thompson *et al.* 2000). Future work should examine other species, with a wider range of life-history characteristics. In addition, similar approaches can be used to determine

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3168v4 | CC BY 4.0 Open Access | rec: 29 Jul 2018, publ: 29 Jul 2018

the minimum time series length required to address additional questions of interest.

Supporting Information

In the supporting material, I provide an expanded methods sections, additional figures, minimum time calculations for determining exponential growth, simulations with a more complicated population model, the use of generalized additive models to identify population trends, and regressions that correct for phylogenetic relationships. All code and data can be found at https://github.com/erwhite1/time-series-project

Acknowledgements

ERW was partially supported by a National Science Foundation Graduate Fellowship. I would like to thank members of the Ecological Theory group at the University of California, Davis for their insight. I would also like to thank T. Dallas, E. Malcolm-White, and two anonymous reviewers for their insights and helpful criticisms.

References

Blanchard, J.L., Maxwell, D.L. & Jennings, S. (2007). Power of monitoring surveys to detect abundance trends in depleted populations: the effects of density-dependent habitat use, patchiness, and climate change. *ICES Journal of Marine Science*, 65, 111–120.

Bolker, B.M. (2008). *Ecological Models and Data in R*. 1st edn. Princeton University Press, Princeton, New Jersey.

Brashares, J.S. & Sam, M.K. (2005). How much is enough? Estimating the minimum sampling required for effective monitoring of African reserves. *Biodiversity and Conservation*, 14, 2709–2722.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155–159.

Field, S.A., Connor, P.J.O., Tyre, A.J. & Possingham, H.P. (2007). Making monitoring meaningful. *Austral Ecology*, 32, 485–491.

Gerber, L.R., DeMaster, D.P. & Kareiva, P.M. (1999). Gray whales and the value of monitering data in implementing the U.S. endangered species act. *Conservation Biology*, 13, 1215–1219.

Gerrodette, T. (1987). A power analysis for detecting trends. *Ecology*, 68, 1364–1372.

Gibbs, J.P., Droege, S. & Eagle, P. (1998). Monitoring populations of plants and animals. *BioScience*, 48, 935–940.

Giron-Nava, A., James, C.C., Johnson, A.F., Dannecker, D., Kolody, B. & Lee, A. *et al.* (2017). Quantitative argument for long-term ecological monitoring. *Marine Ecology Progress Series*, 572, 269–274.

Hatch, S.A. (2003). Statistical power for detecting trends with applications to seabirds monitoring. *Biological Conservation*, 111, 317–329.

Hauser, C.E., Pople, A.R. & Possingham, H.P. (2006). Should managed populations be monitored every year? *Ecological Applications*, 16, 807–819.

Hughes, B.B., Beas-luna, R., Barner, A.K., Brewitt, K., Brumbaugh, D.R. & Cerny-Chipman, E.B. *et al.* (2017). Long-term studies contribute disproportionately to ecology and policy. *BioScience*, 67, 271–281.

IUCN. (2012). IUCN Red List Categories and Criteria: Version 3.1.

Johnson, P.C., Barry, S.J., Ferguson, H.M. & Müller, P. (2015). Power analysis for generalized linear mixed models in ecology and evolution. *Methods in Ecology and Evolution*, 6, 133–142.

Keith, D., Akçakaya, H.R., Butchart, S.H., Collen, B., Dulvy, N.K. & Holmes, E.E. *et al.* (2015). Temporal correlations in population trends: Conservation implications from time-series analysis of diverse animal taxa. *Biological Conservation*, 192, 247–257.

Larsen, D.P., Kincaid, T.M., Jacobs, S.E. & Urquhart, N.S. (2001). Designs for Evaluating Local and Regional Scale Trends. *BioScience*, 51, 1069.

Legg, C.J. & Nagy, L. (2006). Why most conservation monitoring is, but need not be, a waste of time. *Journal of Environmental Management*, 78, 194–199.

Magurran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A. & Scott, E.M. *et al.* (2010). Long-term datasets in biodiversity research and monitoring : assessing change in ecological communities through time. *Trends in Ecology and Evolution*, 25, 574–582.

Mapstone, B.D. (1995). Scalable decision rules for environmental impact studies : effect Size , type I , and type II errors. *Ecological Applications*, 5, 401–410.

Martin, T.G., Camaclang, A.E., Possingham, H.P., Maguire, L.A. & Chadès, I. (2017). Timing of Protection of Critical Habitat Matters. *Conservation Letters*, 10, 308–316.

Martin, T.G., Nally, S., Burbidge, A.A., Arnall, S., Garnett, S.T. & Hayward, M.W. *et al.* (2012). Acting fast helps avoid extinction. *Conservation Letters*, 5, 274–280.

McCain, C.M., Szewczyk, T. & Knight, K.B. (2016). Population variability complicates the accurate detection of climate change responses. *Global Change Biology*, 22, 2081–2093.

McDonald-Madden, E., Baxter, P.W.J., Fuller, R.A., Martin, T.G., Game, E.T. & Montambault, J. *et al.* (2010). Monitoring does not always count. *Trends in Ecology and Evolution*, 25, 547–550.

Morrison, S. & Hik, D.S. (2008). When? Where? And for how long? Census design considerations for an Alpine Lagomorph, the Collared pika. In: *Lagomorph biology*. Springer

Berlin Heidelberg, pp. 103–113.

Myhrvold, N.P., Baldridge, E., Chan, B., Sivam, D., Freeman, D.L. & Ernest, S.M. (2015). An amniote life-history database to perform comparative analyses with birds, mammals, and reptiles. *Ecology*, 96, 3109.

NERC Centre for Population Biology Imperial College. (2010). The Global Population Dynamics Database Version 2. http://www3.imperial.ac.uk/cpb/databases/gpdd.

Nichols, J.D. & Williams, B.K. (2006). Monitoring for conservation. *Trends in Ecology and Evolution*, 21, 668–673.

Peters, D.P. (2010). Accessible ecology: Synthesis of the long, deep, and broad. *Trends in Ecology and Evolution*, 25, 592–601.

R Core Team. (2017). R: A language and environment for statistical computing.

Rhodes, J.R. & Jonzen, N. (2011). Monitoring temporal trends in spatially structured populations: how should sampling effort be allocated between space and time? *Ecography*, 34, 1040–1048.

Rueda-Cediel, P., Anderson, K.E., Regan, T.J., Franklin, J. & Regan, M. (2015). Combined influences of model choice, data quality, and data quantity when estimating population trends. *PLoSONE*, 10, e0132255.

Seavy, N.E. & Reynolds, M.H. (2007). Is statistical power to detect trends a good assessment of population monitoring? *Biological Conservation*, 140, 187–191.

Thomas, L. (1996). Monitoring long-term population change: why are there so many analysis methods? *Ecology*, 77, 49–58.

Thomas, L. (1997). Retrospective Power Analysis. *Conservation Biology*, 11, 276–280.

Thomas, L. & Krebs, C. (1997). A review of statistical power analysis software. *Bulletin of* *the Ecological Society of America*, 78, 128–139.

Thompson, P.M., Wilson, B., Grellier, K. & Hammond, P.S. (2000). Combining power analysis and population viability analysis to compare traditional and precautionary approaches to conservation of coastal cetaceans. *Conservation Biology*, 14, 1253–1263.

Wagner, T., Vandergoot, C.S. & Tyson, J. (2009). Evaluating the power to detect temporal trends in fishery-independent surveys - A case study based on gill nets set in the Ohio waters of Lake Erie for walleyes. *North American Journal of Fisheries Management*, 29, 805–816.

White, E.R., Nagy, J.D. & Gruber, S.H. (2014). Modeling the population dynamics of lemon sharks. *Biology Direct*, 9, 1–18.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A. & Smith, G.M. (2009). *Mixed Effects Models and Extensions in Ecology with R*. Springer, New York.

Figure captions

Figure 1: (a) Example of a simulated time series for 40 time periods. (b) Statistical power versus the simulated time series length. The horizontal, dashed line is the desired statistical power of 0.8. The vertical, dashed line is the minimum time required to achieve the desired statistical power. (c) Minimum time required (T_{min}) for simulations with different values of the trend strength (r) and $\sigma = 5.0$. (d) Minimum time required for different levels of population variability (σ) and $r = 1.5$. In each case, the minimum time required is the minimum number of years to achieve 0.8 statistical power given a significance level of 0.05..

Figure 2: (a) Distributions of the minimum time required for populations from four different biological classes. (b) Distribution of minimum time required for all populations regardless of biological class. The minimum time required calculation corresponds to a significance level of 0.05 and statistical power of 0.8.

Figure 3: Minimum time required to estimate change in abundance correlated with (a) trend strength (absolute value of slope coefficient estimated from linear regression), (b) coefficient of variation in interannual population size, (c) temporal lag-1 autocorrelation, (d) generation length (years), (e) litter size (n), (f) log adult body mass (grams), (g) maximum longevity (years), and (h) incubation (days). The lines in each plot represent the best fit line from linear regression.

Figure 4: 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. The silhouettes highlight that species with longer generation times typically have larger discrepancies between T_{min} and the minimum time required for IUCN assessments.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3168v4 | CC BY 4.0 Open Access | rec: 29 Jul 2018, publ: 29 Jul 2018

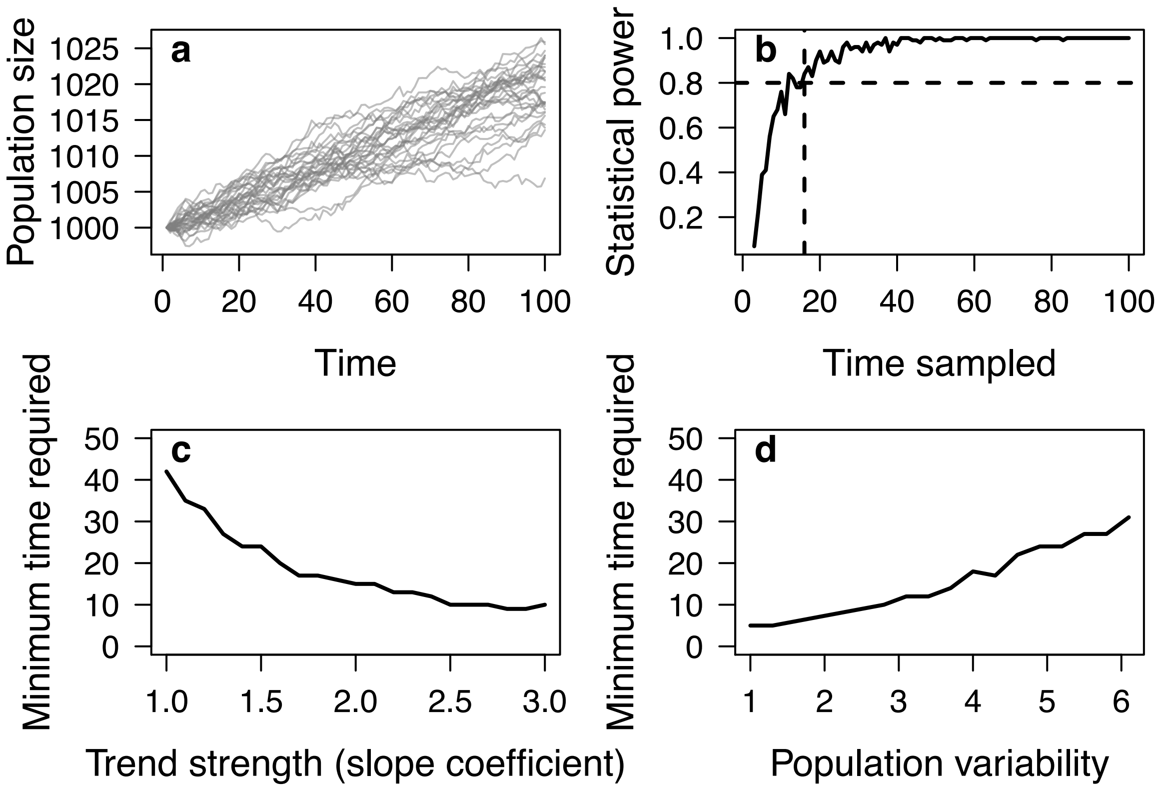

Figure 1: (a) Example of a simulated time series for 40 time periods. (b) Statistical power versus the simulated time series length. The horizontal, dashed line is the desired statistical power of 0.8. The vertical, dashed line is the minimum time required to achieve the desired statistical power. (c) Minimum time required (T_{min}) for simulations with different values of the trend strength (r) and $\sigma = 5.0$. (d) Minimum time required for different levels of population variability (σ) and $r = 1.5$. In each case, the minimum time required is the minimum number of years to achieve 0.8 statistical power given a significance level of 0.05.

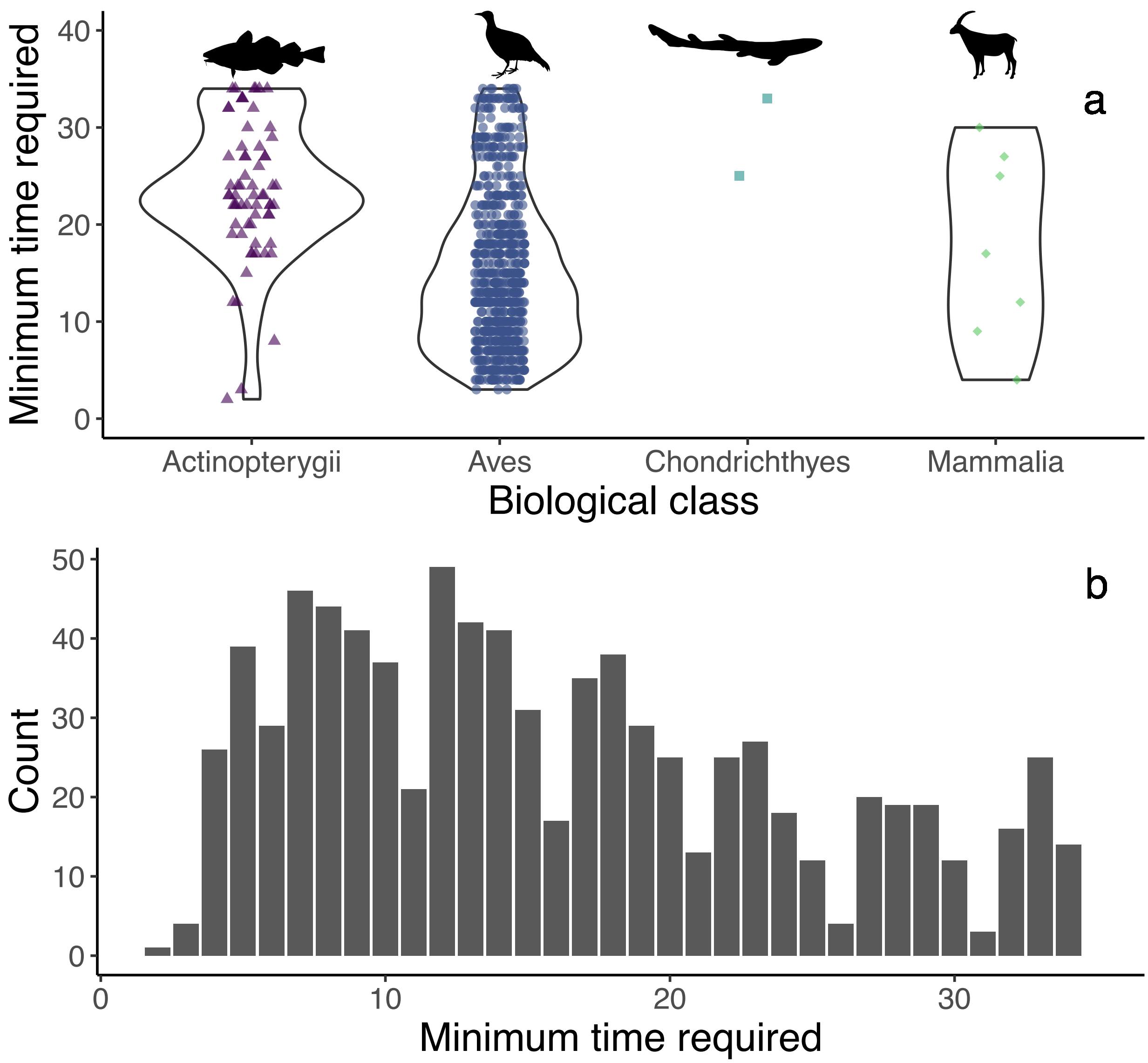

Figure 2: (a) Distributions of the minimum time required for populations from four different biological classes. (b) Distribution of minimum time required for all populations regardless of biological class. The minimum time required calculation corresponds to a significance level of 0.05 and statistical power of 0.8.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.3168v4 | CC BY 4.0 Open Access | rec: 29 Jul 2018, publ: 29 Jul 2018

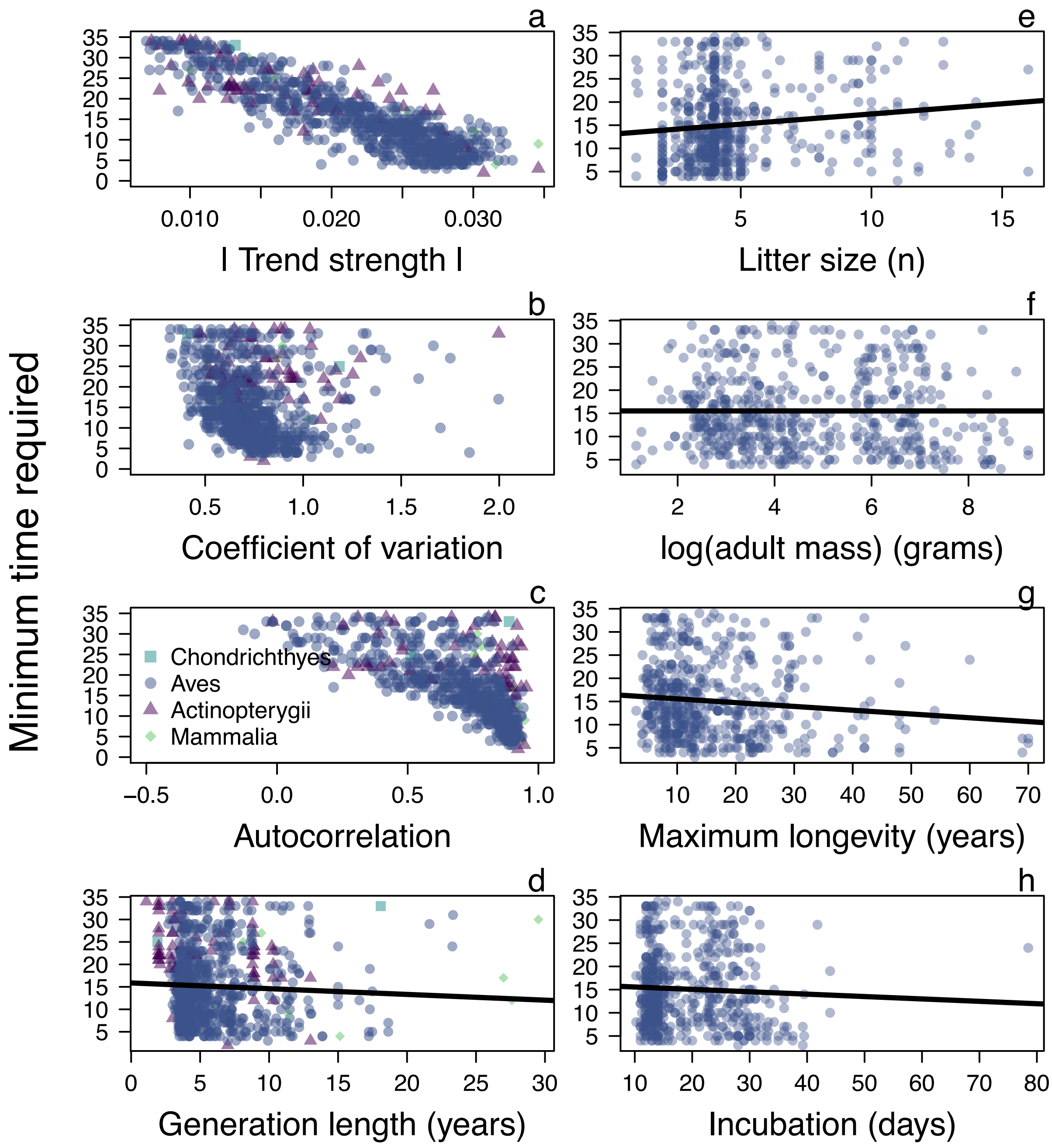

Figure 3: Minimum time required to estimate change in abundance correlated with (a) trend strength (absolute value of slope coefficient estimated from linear regression), (b) coefficient of variation in interannual population size, (c) temporal lag-1 autocorrelation, (d) generation length (years), (e) litter size (n), (f) log adult body mass (grams), (g) maximum longevity (years), and (h) incubation (days). The lines in each plot represent the best fit line from linear regression.

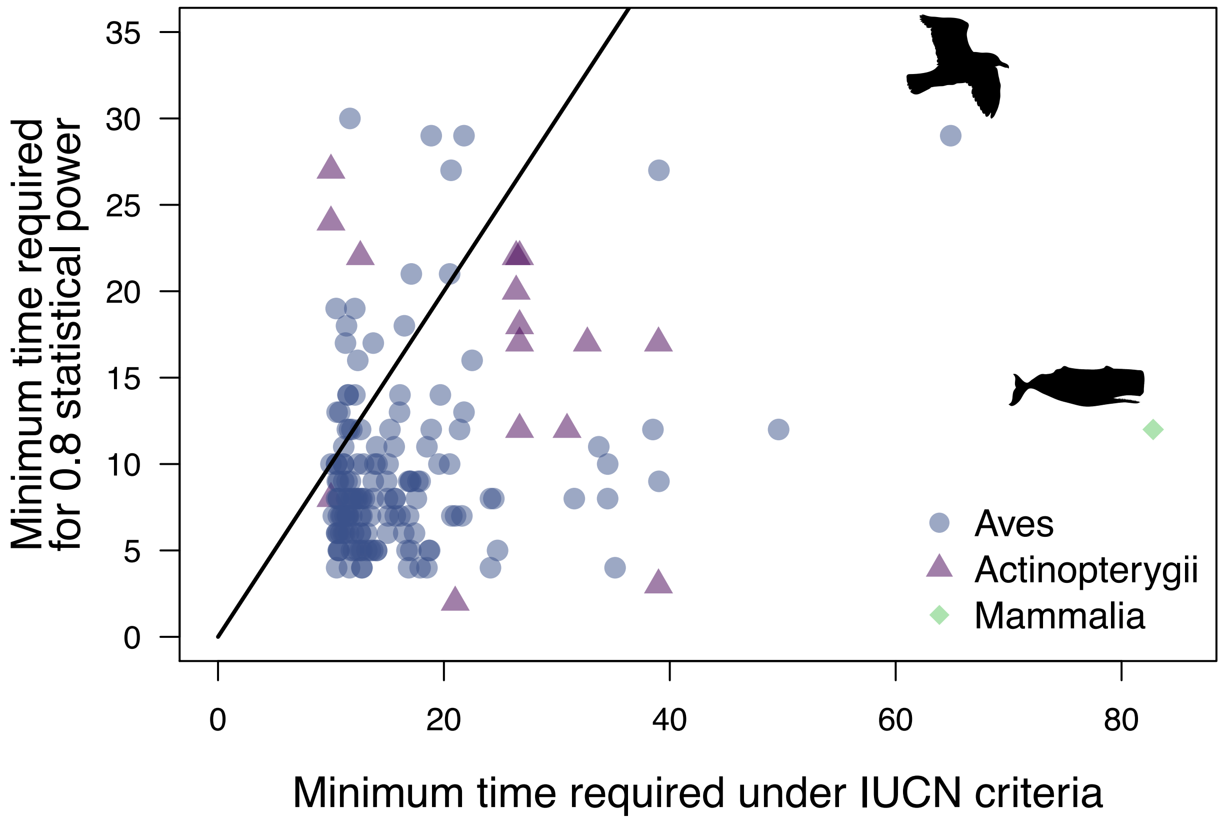

Figure 4: Minimum time required to achieve 0.8 statistical power versus the minimum time required under IUCN criteria A2 to classify a species as vulnerable. Each point represents a single population, all of which saw declines of 30% or greater over a 10 year period. The silhouettes highlight that species with longer generation times typically have larger discrepancies between T_{min} and the minimum time required for IUCN assessments.