

# Experimenting with the past to improve environmental monitoring programs

EASTON R. WHITE<sup>1\*</sup> AND CHRISTIE A. BAHLAI<sup>2</sup>

<sup>1</sup>Department of Biology, University of Vermont, VT 05405, USA,

<sup>2</sup>Department of Biological Sciences, Kent State University, OH 44242, USA

## Abstract

1. Long-term monitoring programs are a fundamental part of both understanding system dynamics and informing management decisions. However, monitoring programs not always designed to consider statistical power, site selection, or the full costs and benefits of monitoring. Further, data from monitoring programs with different goals and protocols are now being combined for comparative analyses.
2. Key considerations can be incorporated into the optimal design of a management program with simulations and experiments. Here, we advocate for the expanded use of a third approach: non-random sampling of previously-collected data. This approach conducts experiments with available data to understand the consequences of different monitoring approaches.
3. We first illustrate non-random sampling in the context of monitoring programs to assess species trends. We then apply the approach to a pair of additional, more general case studies to show the versatility of conducting experiments with previously-collected data.
4. Non-random sampling of previously-collected data is underutilized, but has the potential to improve monitoring programs. We show that this approach is useful in monitoring species trends, understanding fisheries and agriculture, as well as other areas. When combined with data on the cost of monitoring, this approach can also be used to assess the value of information gained from monitoring.

Keywords: statistical power, population trends, data-poor fisheries, species monitoring

## Long-term environmental monitoring

Long-term monitoring programs are an essential piece of modern ecological research and conservation science (Hughes *et al.* 2017). Numerous studies have demonstrated that long-term monitoring can have disproportionately large contributions in terms of advancing scientific understanding and policy (Giron-Nava *et al.* 2017). Environmental monitoring programs, like the USA-based Long Term Ecological Research (LTER) Network, as well as compilations of time series, like the Living Planet Index, show the scope of long-term datasets now

---

\*Easton.White@uvm.edu

available (Maguran *et al.* 2010, Foundation 2016). Furthermore, with the advent of infrastructure that connects and stores data collected by a wide variety of professional and amateur naturalists, monitoring should continue to become more feasible and cost-effective. For example, large-scale citizen science programs, like iNaturalist (<https://www.inaturalist.org/>) and eBird (<https://ebird.org/home>), allow for increased collection of data documenting species occurrence, extent and relative population size, as well as providing platforms to support data use and reuse across a variety of applications (Sullivan *et al.* 2009, Joppa 2017). Similarly, numerous new technologies, including eDNA and drones, will bring down the cost of monitoring through automation and increasing the sheer taxonomic, temporal, and spatial resolution of observations (Bohmann *et al.* 2014, Hodgson *et al.* 2018). All of these efforts will lead to increases in the number of species monitored as well as the quantity and quality of the data collected, to previously unimaginable levels.

## Challenges of monitoring programs

Despite the recognized importance of long-term monitoring programs, key questions remain. Are long-term monitoring programs designed in a way to address key questions of interest (Legg & Nagy 2006, Nichols & Williams 2006, Field *et al.* 2007, McDonald-Madden *et al.* 2010, Lindenmayer *et al.* 2012; 2020)? For instance, suppose that a monitoring site was selected to monitor dynamics in a bird population near a university. A long-term study could certainly reveal the population trend for that specific location. However, the site may have been chosen specifically because it was at high abundance at the beginning of the study—this causes a site-selection bias (Fournier *et al.* 2019). Populations naturally vary, both in time and in space, so the very act of initially selecting a site to monitor with particular population attributes can potentially confound the very patterns they seek to monitor. Birds in this hypothetical population may undergo a cyclic dynamic related to resource exploitation, or rotate between different patches for nesting from year to year. Thus, when we ask new questions of long-term monitoring data, we have to think carefully about how the monitoring program was originally designed and whether or not we have adequate statistical power (Lindenmayer & Likens 2010, White 2019) as well as the risks of making type I versus type II errors (Mapstone 1995). These considerations, amongst others, are especially relevant when data from different sources are combined for comparisons—which is increasingly performed (Maguran *et al.* 2010, Keith *et al.* 2015, Giron-Nava *et al.* 2017, White 2019). Lastly, the tradeoff between information gained from monitoring and the cost of monitoring has to be considered (Bennett *et al.* 2018).

## Designing monitoring programs

Principles from experimental design, including randomization and replication, are key components in designing any monitoring program (Seavy & Reynolds 2007). However, many tools from experimental design are inadequate for monitoring programs. For example, by their very nature, data from monitoring programs will be confounded by spatial and temporal auto-correlation between sampling points. As an alternative, a lot of work on optimizing monitoring programs has its roots in decision science (McDonald-Madden *et al.* 2010). Decision theory allows one to build a structured process to decide between alternative solutions

while accounting for costs and benefits (Raiffa 1968, McDonald-Madden *et al.* 2010, Conroy & Peterson 2012). In addition, decision theory allows for the incorporation of uncertainties (McCarthy 2014). In the context of environmental monitoring, decision theory can help formalize the process of selecting sites and the specific survey design (Gerber *et al.* 2005, Chades *et al.* 2011, Tulloch *et al.* 2013). As an example, Hauser *et al.* (2006) explored how frequently a managed population of red kangaroo (*Macropus rufus*) in Australia should be monitored. They used a simulation approach to determine how frequently monitoring should occur given tradeoffs between the costs of monitoring and the potential insights for management. They found that an adaptive monitoring program outperformed the standard fixed-interval monitoring.

This leads to one of the most important contributions of decision theory to environmental monitoring—the value of information (Maxwell *et al.* 2015, Canessa *et al.* 2015, Bennett *et al.* 2018). Value of information theory explicitly accounts for the information gained from performing some action, and the costs associated with doing so, when designing a management plan. For example, Bennett *et al.* (2018) used information theory in the context of threatened plant management in Southern Ontario. They simulate a situation where a conservation agency is willing to pay landowners to protect part of their land. However, species occurrence is uncertain on each plot of land. Thus, it is necessary assess the value of the information gained from monitoring versus its costs. They were able to demonstrate how the information gained from monitoring increased with high levels of species detectability and low costs of monitoring.

## Tools for designing monitoring programs

### 1. Simulations

To address the challenges associated with designing monitoring programs, there are three classes of tools available to design and evaluate monitoring programs. First, the most commonly used approach are simulation models (Gerrodette 1987, Rhodes & Jonzen 2011). Using prior knowledge about the system under question, a virtual ecologist (Zurell *et al.* 2010) approach would use simulation models constructed to incorporate key factors that affect species dynamics. With an appropriate model, simulations can then be run for a variety of scenarios, including changing the number of samples taken per year, altering the number of sites sampled, and sampling for different lengths of time (Rhodes & Jonzen 2011, Barry *et al.* 2017, White 2019, Christie *et al.* 2019, Weiser *et al.* 2019) Simulations can also be useful in deciding which streams of data to use (Weiser *et al.* 2020) or the effect of changing sampling methodology during the course of a study (Southwell *et al.* 2019). A lot of prior work has also used simulations to better understand optimal sampling schemes for invasive species (Chades *et al.* 2011, Holden & Ellner 2016). Although powerful, this approach is limited to systems in which many aspects of the biology are already known to some extent.

### 2. Experimental & Comparative

Second, experiments can also be used to test the effect of different sampling protocols. As in the case of simulation models, experiments with different levels of monitoring, or different monitoring approaches, can be used. Experiments provide replication, which is important to understand the probability that monitoring is likely to achieve the desired goals given inherent variability in the system (White *et al.* 2019). A related approach would

simply be to compare different sampling regimes across systems to evaluate which are the most successful at realizing monitoring goals. Indeed, integrated population modelling was developed as an analytical approach to identify and address data discrepancies between data taken by differing methodologies or at differing times in a species’s life history (Saunders *et al.* 2019). This method has been applied with great success to advance understanding of the trajectories of populations of well-monitored taxa such as waterfowl (Arnold *et al.* 2018). However, the key disadvantage of this approach is, like simulation models, integrative modelling approaches are reliant on the availability of large amounts of data, documenting multiple facets of a species’ biology. Of course, these types of experiments providing multi-faceted data are often infeasible or impossible for many systems.

### 3. Non-random sampling

Here, we advocate for the expanded use of a third approach: **non-random sampling of previously-collected monitoring data**. This concept leverages existing information by starting with long-term monitoring data already collected for a system. The data can then be subsampled, or divided, in various ways depending on the question of interest (Fig. 1a). Then a metric (for example, a mean or a slope) could be calculated for each subsample. Each subsample metric would then be compared to the metric for complete data (all the data combined). The complete data acts as a “true value” for comparison. This is analogous to simulation studies where the true parameters are known (Bolker 2008). We learn about the elements of a good monitoring program by examining which subsamples of the data are most influential and the number of subsamples needed to have a high probability of detecting the true value of the metric. Although this approach has been used previously (Grantham *et al.* 2008, Bennett *et al.* 2016, White 2019, Wauchope *et al.* 2019), its adoption has been largely informal and not specifically stated.

This approach is best described with a simple example (Fig. 1b). White (2019) studied how many years of monitoring were required to detect population trends. For a given time series, White (2019) examined all possible subsamples of different lengths of time. For instance, a 10-year time series consists of two 9-year subsamples, three 8-year subsamples, and so forth. (White 2019) then calculated the population trend (i.e. the slope from linear regression) for each subsample (Fig. 1b). The fraction of subsamples of a particular length, that had the same overall trend as the complete time series (i.e. the “true trend”), is the statistical power. Thus, the minimum time series required was the time series length that met a high enough threshold of statistical power. (White 2019) applied this approach to 822 population time series, allowing for comparison across species and systems. Using a similar approach, (Wauchope *et al.* 2019) examined both the minimum time and frequency of sampling required to be confident in determining species trends. Using resampling of the breeding bird survey, they found that sampling for a short period, or infrequently, was adequate to determine the species trend direction, i.e. positive or negative. However, more frequent and longer monitoring was required to estimate the percent changes over time.

Monitoring programs need to be designed to both adequately address a question of interest and to be cost effective (Grantham *et al.* 2008, Rout *et al.* 2014, Maxwell *et al.* 2015, Bennett *et al.* 2018). In this context, it is essential to study the trade-off between the information gained from monitoring and the cost. Again, non-random sampling of past data can be useful here (Grantham *et al.* 2008, Bennett *et al.* 2016). For example, Bennett *et al.*

(2016) used resampling approaches to study monitoring requirements for diatoms in lake samples. They examined how the number of lakes sampled and observer effort (number of diatoms counted in each sample) affected various common environmental analyses. They found that in several cases much lower levels of sampling were required to ensure accuracy. Importantly, they also weighted this accuracy with the cost of data collection. This translated into potentially millions of dollars in savings (Bennett *et al.* 2016).

Non-random sampling of past data also helps address some common issues with other approaches to designing monitoring programs. For instance, simulation models require at least some knowledge of the basic species biology in order to construct the model (Zurell *et al.* 2010). In addition, although possible with simulations, resampling approaches already explicitly account for the inherent temporal and spatial autocorrelation of monitoring data. Resampling approaches are also particularly useful in situations where experiments are expensive or impractical. Along with these advantages, resampling of past data has two major limitations. First, monitoring data already has to be available for the system of interest or a related one. In addition, the previously-collected data needs to be a good representation of the system dynamics. This is due to the full data acting as the “true representation” of this system. If however, the full data is not representative of the true dynamics, then any resampling approaches will likely be unhelpful.

## Non-random sampling in other contexts

### Fisheries management

Non-random sampling of past data can be applied to a variety of contexts beyond estimating long-term population trends. For example, consider the importance of studying data-poor fisheries (Dowling *et al.* 2015 and Table 1 of Chrysafi & Kuparinen 2015). Using mostly simulation models, a lot of work has been done to develop tools when observational data is limited. To study data-poor fisheries using non-random sampling, one should instead study data-rich fisheries. The goal would be to artificially degrade the data-rich examples until the point that the fishery would be considered data-poor (Fig. 2). We can then see how various methods of data-poor fisheries perform given that we have the full data set to act as a true comparison. The idea is the same as in simulation studies where we know the parameter values of simulated data exactly. As an example, we took data on darkblotched rockfish (*Sebastes crameri*) from the U.S. West Coast Groundfish Bottom Trawl Survey data (Keller *et al.* 2017, Stock *et al.* 2019). We then conducted two “experiments” with the data. First, we suppose we only had access to shallow or deep data because of technology limitations. We show that regression estimates of parameters differs based on which depths were included (Fig. 2b). We also examine the effect of degrading the data to only a fraction of the total records we have available. We see that model estimates for the effect of being within a rockfish conservation area are not accurate until a large fraction of the original data is included (Fig. 2c).

### Agricultural practices

A practical example of the application of non-random sampling applied to management decision-making can be found in Cusser *et al.* (2019). Agricultural management recommendations are often based on conclusions from short to medium-term field trials (ca. 1-5 years), and it is common to observe contradictory findings between trials. In this study, the

authors applied a non-random sequential sampling algorithm (Bahlai et al. 2019; similar to that described in Fig. 1b) to long-term data examining the effects of tillage practice on productivity and return-on investment. They found that, because of high natural variability in the system, 10 years of data was required to observe the “true” pattern of difference between treatments, and that more than a third of the sampled sequences shorter than 10 years led to outright misleading results (i.e. statistically significant trends which showed the opposite relationship between treatments). Whereas it is unlikely that practitioners making management decisions can consistently rely on a decade of data to guide them, the results of the non-random sequential sampling of long-term data provide guidance on reconciling apparently differing trends between trials.

## Conclusions

Data from long-term monitoring programs are used in assessing trends in environmental observations, understanding system dynamics, and making management decisions. It is critical that these monitoring programs be designed in order to address our questions of interest (Field *et al.* 2007, McDonald-Madden *et al.* 2010, Lindenmayer *et al.* 2020). This is particularly relevant when new questions are asked of monitoring data or data from disparate monitoring programs are combined. We show that non-random sampling of past monitoring programs can be used to understand sampling requirements and the consequences of bias (Figs. 1,2). This approach can be applied to a variety of systems and questions (Table 1) beyond environmental monitoring. Combined with simulations and experimental approaches, we argue that non-random sampling of past data should be used more widely to study questions related to sampling design. Combined with information on the cost of monitoring, this approach also helps identify when ecological monitoring is a good investment or when it may be a waste of effort that does not answer to a program’s aims and objectives. Continued research in this area will allow scientists and managers to better evaluate past efforts and to design new monitoring programs using evidence-based approaches.

## References

- Arnold, T.W., Clark, R.G., Koons, D.N. & Schaub, M. (2018) Integrated population models facilitate ecological understanding and improved management decisions. *Journal of Wildlife Management*, **82**, 266–274.
- Bahlai, C. (2019) The bad breakup Algorithm. *Zenodo*, pp. Version v1.0–Version v1.0.
- Barry, J., Maxwell, D., Jennings, S., Walker, D. & Murray, J. (2017) Emon: An R-package to support the design of marine ecological and environmental studies, surveys and monitoring programmes. *Methods in Ecology and Evolution*, **8**, 1342–1346.
- Bennett, J.R., Maxwell, S.L., Martin, A.E., Chadès, I., Fahrig, L. & Gilbert, B. (2018) When to monitor and when to act: Value of information theory for multiple management units and limited budgets. *Journal of Applied Ecology*, **55**, 2102–2113.
- Bennett, J.R., Rühland, K.M. & Smol, J.P. (2016) No magic number: Determining cost-effective sample size and enumeration effort for diatom-based environmental assessment analyses. *Canadian Journal of Fisheries and Aquatic Sciences*, **74**, 208–215.
- Bohmann, K., Evans, A., Gilbert, M.T.P., Carvalho, G.R., Creer, S., Knapp, M., Yu, D.W. & de Bruyn, M. (2014) Environmental DNA for wildlife biology and biodiversity monitoring. *Trends in Ecology and Evolution*, **29**, 358–367.
- Bolker, B.M. (2008) *Ecological Models and Data in R*. Princeton University Press, Princeton, New Jersey, 1 edition.
- Canessa, S., Guillera-Aroita, G., Lahoz-Monfort, J.J., Southwell, D.M., Armstrong, D.P., Chadès, I., Lacy, R.C. & Converse, S.J. (2015) When do we need more data? A primer on calculating the value of information for applied ecologists. *Methods in Ecology and Evolution*, **6**, 1219–1228.
- Chades, I., Martin, T.G., Nicol, S., Burgman, M.A., Possingham, H.P. & Buckley, Y.M. (2011) General rules for managing and surveying networks of pests, diseases, and endangered species. *Proceedings of the National Academy of Sciences*, **108**, 8323–8328.
- Christie, A.P., Amano, T., Martin, P.A., Shackelford, G.E., Simmons, B.I. & Sutherland, W.J. (2019) Simple study designs in ecology produce inaccurate estimates of biodiversity responses. *Journal of Applied Ecology*, pp. 1–13.
- Chrysafi, A. & Kuparinen, A. (2015) Assessing abundance of populations with limited data: Lessons learned from data-poor fisheries stock assessment. *Environmental Reviews*, **24**, 25–38.
- Conroy, M.J. & Peterson, J.T. (2012) *Decision Making in Natural Resource Management: A Structured, Adaptive Approach*. John Wiley and Sons, West Sussex, UK.
- Cusser, S., Bahlai, C., Swinton, S.M., Robertson, G.P. & Haddad, N.M. (2019) Long-term research needed to avoid spurious and misleading trends in sustainability attributes of no-till. *bioRxiv*, pp. 1–31.

- Dowling, N.A., Dichmont, C.M., Haddon, M., Smith, D.C., Smith, A.D. & Sainsbury, K. (2015) Empirical harvest strategies for data-poor fisheries: A review of the literature. *Fisheries Research*, **171**, 141–153.
- Field, S.A., Connor, P.J.O., Tyre, A.J. & Possingham, H.P. (2007) Making monitoring meaningful. *Austral Ecology*, **32**, 485–491.
- Foundation, W.W. (2016) Living Planet Report 2016 - Risk and resilience in a new era. Technical Report 9782940529407.
- Fournier, A.M., White, E.R. & Heard, S.B. (2019) Site-selection bias and apparent population declines in long-term studies. *Conservation Biology*, **00**, 1–10.
- Gerber, L.R., Beger, M., McCarthy, M.A. & Possingham, H.P. (2005) A theory for optimal monitoring of marine reserves. *Ecology Letters*, **8**, 829–837.
- Gerrodette, T. (1987) A power analysis for detecting trends. *Ecology*, **68**, 1364–1372.
- Giron-Nava, A., James, C.C., Johnson, A.F., Dannecker, D., Kolody, B., Lee, A., Nagarkar, M., Pao, G.M., Ye, H., Johns, D.G. & Sugihara, G. (2017) Quantitative argument for long-term ecological monitoring. *Marine Ecology Progress Series*, **572**, 269–274.
- Grantham, H.S., Moilanen, A., Wilson, K.A., Pressey, R.L., Rebelo, T.G. & Possingham, H.P. (2008) Diminishing return on investment for biodiversity data in conservation planning. *Conservation Letters*, **1**, 190–198.
- Hauser, C.E., Pople, A.R. & Possingham, H.P. (2006) Should managed populations be monitored every year? *Ecological Applications*, **16**, 807–819.
- Hodgson, J.C., Mott, R., Baylis, S.M., Pham, T.T., Wotherspoon, S., Kilpatrick, A.D., Raja Segaran, R., Reid, I., Terauds, A. & Koh, L.P. (2018) Drones count wildlife more accurately and precisely than humans. *Methods in Ecology and Evolution*, **9**, 1160–1167.
- Holden, M.H. & Ellner, S.P. (2016) Human judgment vs. quantitative models for the management of ecological resources. *Ecological Applications*, **26**, 1553–1565.
- Hughes, B.B., Beas-luna, R., Barner, A.K., Brewitt, K., Brumbaugh, D.R., Cerny-Chipman, E.B., Close, S.L., Coblenz, K.E., Nesnera, K.L.D., Drobnych, S.T., Figurski, J.D., Focht, B., Friedman, M., Freiwald, J., Heady, K.K., Heady, W.N., Hettinger, A., Johnson, A., Karr, K.A., Mahoney, B., Moritsch, M.M., Osterback, A.M.K., Reimer, J., Robinson, J., Rohrer, T., Rose, J.M., Sabal, M., Segui, L.M., Shen, C., Sullivan, J., Zuercher, R., Raimondi, P.T., Menge, B.A., Grorud-Colvert, K., Novak, M. & Carr, M.H. (2017) Long-term studies contribute disproportionately to ecology and policy. *BioScience*, **67**, 271–281.
- Joppa, L.N. (2017) The case for technology investments in the environment. *Nature*, **552**, 325–328.



- Keith, D., Akçakaya, H.R., Butchart, S.H., Collen, B., Dulvy, N.K., Holmes, E.E., Hutchings, J.A., Keinath, D., Schwartz, M.K., Shelton, A.O. & Waples, R.S. (2015) Temporal correlations in population trends: Conservation implications from time-series analysis of diverse animal taxa. *Biological Conservation*, **192**, 247–257.
- Keller, A.A., Wallace, J.R. & Methot, R.D. (2017) The Northwest Fisheries Science Center’s West Coast Groundfish Bottom Trawl Survey: History, Design, and Description. Technical report.
- 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.
- Lindenmayer, D., Woinarski, J., Legge, S., Southwell, D., Lavery, T., Robinson, N., Scheele, B. & Wintle, B. (2020) A checklist of attributes for effective monitoring of threatened species and threatened ecosystems. *Journal of Environmental Management*, **262**, 110312.
- Lindenmayer, D.B., Gibbons, P., Bourke, M., Burgman, M., Dickman, C.R., Ferrier, S., Fitzsimons, J., Freudenberger, D., Garnett, S.T., Groves, C., Hobbs, R.J., Kingsford, R.T., Krebs, C., Legge, S., Lowe, A.J., Mclean, R., Montambault, J., Possingham, H., Radford, J., Robinson, D., Smallbone, L., Thomas, D., Varcoe, T., Vardon, M., Wardle, G., Woinarski, J. & Zerger, A. (2012) Improving biodiversity monitoring. *Austral Ecology*, **37**, 285–294.
- Lindenmayer, D.B. & Likens, G.E. (2010) The science and application of ecological monitoring. *Biological Conservation*, **143**, 1317–1328.
- Maguran, A.E., Baillie, S.R., Buckland, S.T., Dick, J.M., Elston, D.A., Scott, E.M., Smith, R.I., Somerfield, P. & Watt, A.D. (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.
- Maxwell, S.L., Rhodes, J.R., Runge, M.C., Possingham, H.P., Ng, C.F. & McDonald-Madden, E. (2015) How much is new information worth? Evaluating the financial benefit of resolving management uncertainty. *Journal of Applied Ecology*, **52**, 12–20.
- McCarthy, M.A. (2014) Contending with uncertainty in conservation management decisions. *Annals of the New York Academy of Sciences*, **1322**, 77–91.
- McDonald-Madden, E., Baxter, P.W.J., Fuller, R.A., Martin, T.G., Game, E.T., Montambault, J. & Possingham, H.P. (2010) Monitoring does not always count. *Trends in Ecology and Evolution*, **25**, 547–550.
- Nichols, J.D. & Williams, B.K. (2006) Monitoring for conservation. *Trends in Ecology and Evolution*, **21**, 668–673.

- Raiffa, H. (1968) *Decision Analysis: Introductory Lectures on Choices Under Uncertainty*. Addison-Wesley, Reading, MA.
- 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.
- Rout, T.M., Moore, J.L. & McCarthy, M.A. (2014) Prevent, search or destroy? A partially observable model for invasive species management. *Journal of Applied Ecology*, **51**, 804–813.
- Saunders, S.P., Farr, M.T., Wright, A.D., Bahlai, C.A., Ribeiro, J.W., Rossman, S., Sussman, A.L., Arnold, T.W. & Zipkin, E.F. (2019) Disentangling data discrepancies with integrated population models. *Ecology*, **100**, 1–14.
- Seavy, N.E. & Reynolds, M.H. (2007) Is statistical power to detect trends a good assessment of population monitoring? *Biological Conservation*, **140**, 187–191.
- Southwell, D.M., Einoder, L.D., Lahoz-Monfort, J.J., Fisher, A., Gillespie, G.R. & Wintle, B.A. (2019) Spatially explicit power analysis for detecting occupancy trends for multiple species. *Ecological Applications*, **29**.
- Stock, B.C., Ward, E.J., Eguchi, T., Jannot, J.E., Thorson, J.T., Feist, B.E. & Semmens, B.X. (2019) Comparing predictions of fisheries bycatch using multiple spatiotemporal species distribution model framework. *Canadian Journal of Fisheries and Aquatic Sciences*, pp. 1–66.
- Sullivan, B.L., Wood, C.L., Iliff, M.J., Bonney, R.E., Fink, D. & Kelling, S. (2009) eBird: A citizen-based bird observation network in the biological sciences. *Biological Conservation*, **142**, 2282–2292.
- Tulloch, A.I.T., Chadès, I. & Possingham, H.P. (2013) Accounting for Complementarity to Maximize Monitoring Power for Species Management: Complementary Monitoring Indicators. *Conservation Biology*, **27**, 988–999.
- Wauchope, H.S., Johnston, A., Amano, T. & Sutherland, W.J. (2019) When can we trust population trends? A method for quantifying the effects of sampling interval and duration. *Methods in Ecology and Evolution*, pp. 498170–498170.
- Weiser, E.L., Diffendorfer, J.E., Grundel, R., López-Hoffman, L., Pecoraro, S., Semmens, D. & Thogmartin, W.E. (2019) Balancing sampling intensity against spatial coverage for a community science monitoring programme. *Journal of Applied Ecology*, **56**, 2252–2263.
- Weiser, E.L., Diffendorfer, J.E., Lopez-Hoffman, L., Semmens, D. & Thogmartin, W.E. (2020) Challenges for leveraging citizen science to support statistically robust monitoring programs. *Biological Conservation*, **242**, 108411.
- White, E.R. (2019) Minimum time required to detect population trends: The need for long-term monitoring programs. *BioScience*, **69**, 40–46.

- White, E.R., Cox, K., Melbourne, B.A. & Hastings, A. (2019) Success and failure of ecological management is highly variable in an experimental test. *Proceedings of the National Academy of Sciences*, **116**, 23169–23173.
- Zurell, D., Berger, U., Cabral, J.S., Jeltsch, F., Meynard, C.N., Münkemüller, T., Nehrbass, N., Pagel, J., Reineking, B., Schröder, B. & Grimm, V. (2010) The virtual ecologist approach: Simulating data and observers. *Oikos*, **119**, 622–635.

**Table 1: Example questions that could be addressed using non-random sampling**

Question	Non-random sampling approach
<p>How many test water wells should be drilled to understand subsurface water flow?</p>	<p>We would start with an example system where a large number of test wells produced accurate dynamics. Then, we artificially degrade this data using less test wells. Last, we would examine when the predicted dynamics change as a result of less test wells.</p>
<p>What is the effect of not being able to identify microorganisms to the species level?</p>	<p>We first select data from a well-resolved tree that does identify organisms to the species level. Then, we artificially degrade the data in a way where we pretend a tree is only resolved to the genus or family level. We could then study the effect of not identifying organisms to the species level.</p>
<p>What is the effect of scuba diving depth limitations on estimates of biological diversity?</p>	<p>We would use high-quality diver survey data that was collected along a gradient of depths. We would then artificially degrade the data by removing deeper dives. We could then compare the diversity metrics when all the data is included versus only shallow dives.</p>

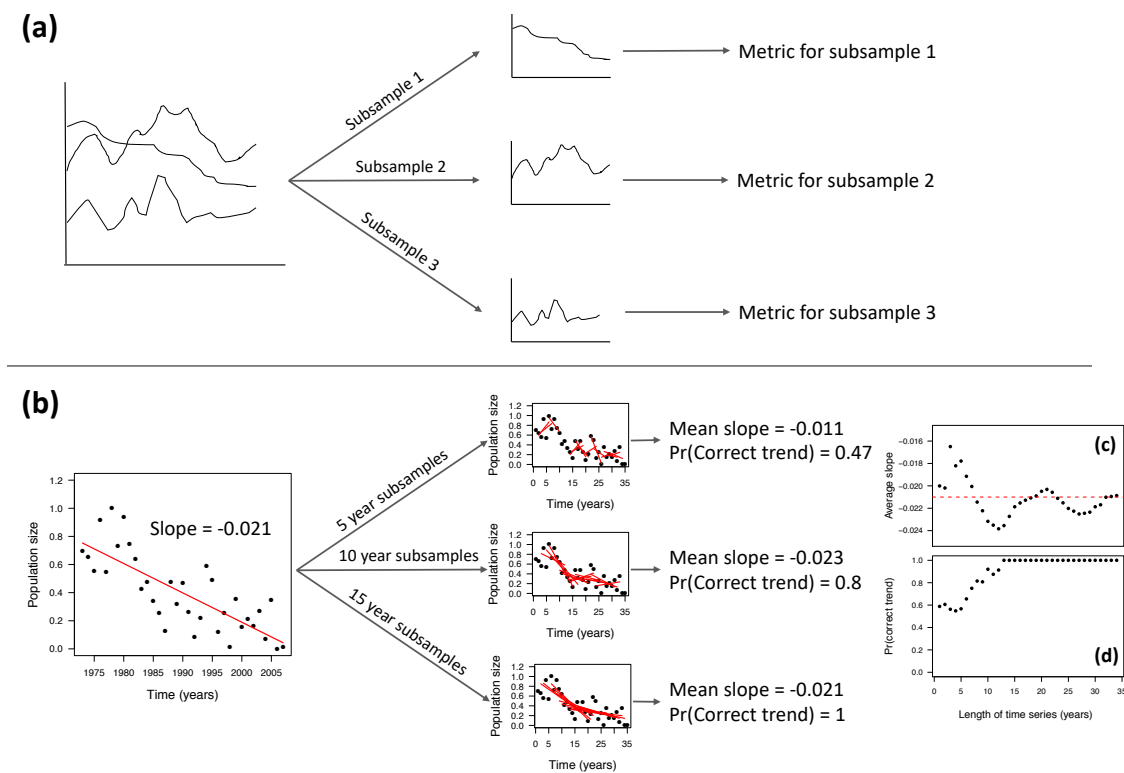


Figure 1: (a) The general process of non-random sampling of past data from left to right (i.e. sequentially, starting with data from farthest into the past) includes: dividing data into subsamples, calculating metrics on those subsamples, and comparing the subsample metrics to the combined (i.e. “true metric”) dataset, (b) same process as Table (a), but for specific example of examining the minimum number of years required to detect long-term population trends (White 2019). The pair of figures on the bottom right show how the (c) average slope and the (d) probability of correctly identifying a trend change with the number of years monitored.

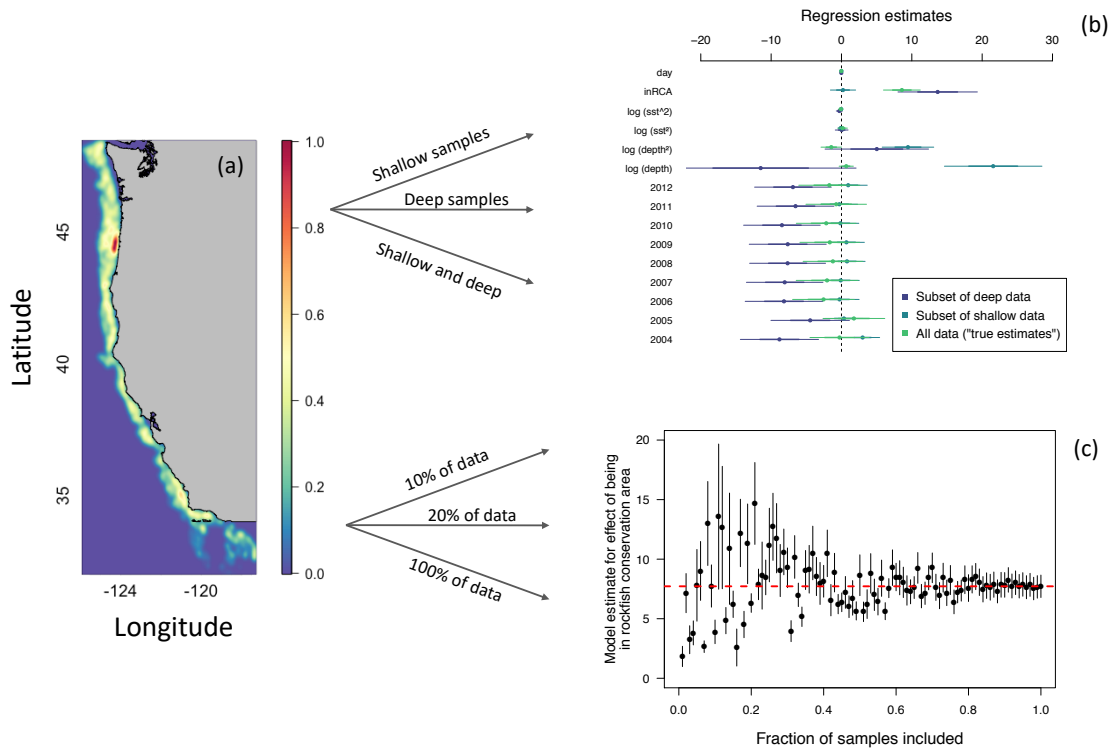


Figure 2: Subsampling process for darkblotched rockfish (*Sebastes crameri*) catch in kg in U.S. West Coast Groundfish Bottom Trawl Survey data from 2003-2012. (a) Bivariate kernel density estimate showing smoothed density of fishing effort (7,161 haul locations) (Keller *et al.* 2017, Stock *et al.* 2019) (b) Parameter estimations for linear regression of three subsamples of data: deep trawls, shallow trawls, and all combined data (Note: the number of records was kept consistent for the three groups). (c) Estimate for the effect of being in a rockfish conservation area (inRCA) on catch for different amounts of data included. The horizontal, dashed line is the “true” estimate which is the estimate when 100% of the data is included.