Ecological Applications, 0(0), 2018, pp. 1–12

A field ecologist's adventures in the virtual world: using simulations to design data collection for complex models

Freya M. Thomas,[1] Peter A. Vesk, and Cindy E. Hauser

School of BioSciences, ARC Centre of Excellence for Environmental Decisions, The University of Melbourne, Melbourne, Victoria 3010 Australia

Abstract. Field data collection can be expensive, time consuming, and difficult; insightful research requires statistical analyses supported by sufficient data. Pilot studies and power analysis provide guidance on sampling design but can be challenging to perform, as ecologists increasingly collect multiple types of data over different scales. Despite a growing simulation literature, it remains unclear how to appropriately design data collection for many complex projects. Approaches that seek to achieve realism in decision-making contexts, such as management strategy evaluation and virtual ecologist simulations, can help. For a relatively complex analysis, we develop and demonstrate a flexible simulation approach that informs what data are needed and how long those data will take to collect, under realistic fieldwork constraints. We simulated data collection and analysis under different constraint scenarios that varied in deterministic (field trip length, travel, and measurement times) and stochastic (species detection and occupancy rates and inclement weather) features. In our case study, we fit plant height data to a multispecies, three-parameter, nonlinear growth model. We tested how the simulated data sets, based on the varying constraint scenarios, affected the model fit (parameter bias, uncertainty, and capture rate). Species prevalence in the field exerted a stronger influence on the data sets and downstream model performance than deterministic aspects such as travel times. When species detection and occupancy were not considered, the field time needed to collect an adequate data set was underestimated by 40%. Simulations can assist in refining fieldwork design, estimating field costs, and incorporating uncertainties into project planning. We argue that combining data collection, analysis, and decision-making processes in a flexible virtual setting can help address many of the decisions that field ecologists face when designing field-based research.

Key words: experimental design; fieldwork; hierarchical models; management strategy evaluation; sampling; simulation; virtual ecology.

Introduction

Field data collection is expensive, time consuming, and difficult. Difficulties can arise when ecologists ask relatively simple questions over complicated environmental gradients or ask questions involving complicated causal pathways to understand processes that underlie ecological theories. These questions require multivariate data collection and complex analyses such as hierarchical or mixed effects modeling (Zuur et al. 2007). In addition to collecting data over complex environmental space, the ecological relationships we care about are often nonlinear in nature, for example, plant species growth (Thomas and Vesk 2017*b*), species area curves (Scheiner 2003), species abundance or dispersal curves (Tjørve 2003), or long-term changes in population structure (Seavy and Reynolds 2007). The complexity of natural systems can place substantial demands on the volume, type, and sampling design of data required for ecological research.

The value of statistical inference depends on the underlying model assumptions and the data collected (Nicholls 1989, Di Stefano 2003). Power analysis is frequently recommended as a key element of study design (Day and Quinn 1989, Johnson et al. 2015) to assess this relationship between data and inference. However, while many ecologists are familiar with the concept of power analysis, it is commonly misapplied due to, among other things, research lacking clear objectives (Yoccoz et al. 2001), arbitrarily defined statistical thresholds that lack biological meaning (Di Stefano 2003) and an under-appreciation of the relative costs of different statistical errors (Di Stefano 2003). Power analyses are often not conducted effectively in ecological studies or monitoring programs (Peterman 1990, Fairweather 1991, Vos et al. 2000, Legg and Nagy 2006, Low-Décarie et al. 2014, Johnson et al. 2015). This is partly due to a culture in ecology that does not ask for them, but also reflects a lack of available methods for contemporary analysis (Johnson et al. 2015).

Designing robust, field-based research is not only about having adequate sample sizes, and achieving the nominal sample size does not ensure a robust study design (van de Pol and van de Pol 2011). A target sample size may not be helpful when a study is constrained by time or budget (Di Stefano 2003). Field ecologists often face missing data and unbalanced designs (Kain et al. 2015), because field data are not always as available as we expect. Methods to appropriately design field campaigns are increasingly becoming more realistic. For instance, occupancy studies have been informed by search effort optimization (although they do not typically quantify time and cost; MacKenzie and Royle 2005), and other survey designs have focused on visits to a single site (Moore and McCarthy 2016) and including optimized

Manuscript received 18 January 2018; revised 1 July 2018; accepted 20 August 2018. Corresponding Editor: David T. Barnett.
[1] E-mail: freyamthomas@gmail.com

searches across multiple sites with varying detectability (Hauser and McCarthy 2009, Moore and McCarthy 2016).

Methods to adequately design ecological studies for complex analysis can be broadly typified as "virtual ecologist" approaches (Tyre et al. 2001, Meyer et al. 2009, Zurell et al. 2010), mathematical optimization methods (Moore and McCarthy 2016) and software packages (Johnson et al. 2015). The "virtual ecologist" approach uses simulation to mimic real-world conditions to virtually construct and analyze data. It has been used to explore the effects of survey design under economic constraints (Field et al. 2007), test analytical tools (Wunder et al. 2007), mimic experimental design in greenhouse conditions (Meyer et al. 2009), determine the largest sources of variability in analysis of species distribution models (Tessarolo et al. 2014) and to monitor dynamic spatiotemporal ecological processes (Williams et al. 2018). Optimization methods differ from "the virtual ecologist" approach. Optimizations target specific study objectives, describe processes and constraints with mathematical equations, and solve them. Optimization has been used to prioritize survey effort for eradication programs (Hauser et al. 2016) and to optimize survey effort over space and time incorporating species detection rates and budget (Moore and McCarthy 2016). Numerous examples exist of software such as R packages aimed at making power analysis easier for ecologists, including odprism, for optimal design of mixed models (van de Pol and van de Pol 2011), simm.glmm (Johnson et al. 2015), and SIMR (Green and MacLeod 2016), which use simulations to conduct power analysis for generalized linear mixed models.

An important question is whether simulations will help researchers design better studies (Peck 2004, Meyer et al. 2009, Johnson et al. 2015). One area where real-world complexities have been successfully included in virtual settings is in Management Strategy Evaluation (MSE; Bunnefeld et al. 2011). This approach uses simulation models in an adaptive framework to test a range of management approaches that meet objectives given a complex decision environment (Bunnefeld et al. 2011). It combines data collection with specific analysis and management implications based on those analyses in a virtual setting (Punt et al. 2016). Some of the greatest advantages of MSE are that it allows experimentation under a range of circumstances, various forms of uncertainties can be included and it encourages prospective rather than retrospective evaluation of actions (Bunnefeld et al. 2011). Management strategy evaluation has been championed in fisheries science and used to identify "realizable" management performance given uncertainties (Punt et al. 2016) and it is beginning to be implemented in complex terrestrial conservation problems (Milner-Gulland et al. 2010, Bunnefeld et al. 2011).

We argue that a similar mentality of combining data collection, analysis, and decision-making processes in a flexible, virtual setting can be applied for the decisions that an ecologist faces when designing field research. Despite the vital role that field studies play in ecology, rigorous design and testing of field studies is rarely conducted (or reported) in the ecological literature. To address this gap, we demonstrate a flexible simulation approach designed to identify how much and what kinds of data are needed for a relatively complex analysis (e.g., trade-offs between number of sites and numbers of individuals sampled at each site), and how long those data will take to collect. Our approach differs from others described above in that we incorporate realistic fieldwork constraints including travel times, poor weather, and varying species detection rates.

We use a case study of height–growth data collected in a semiarid system, where the data were used to develop a multispecies nonlinear growth model to predict heights of multiple plant species given time since fire at a site (Thomas and Vesk 2017*b*). The study area is vast and heterogeneous; not all species are in the same place and some species are harder to find than others. The simulation structure incorporates data analyses, providing a direct causal chain between from design to sampling choices in the field, resultant data sets and the precision and accuracy of the chosen model.

Methods

The aim of this paper was to demonstrate a "context-specific" simulation approach that can be used for designing robust field-based studies. This approach differs from traditional power analyses, which typically do not incorporate realistic fieldwork constraints. Although actual field collection methods vary considerably, there are some common factors that need to be considered when designing robust field studies. These factors, or the "context," involve both environmental and logistical constraints. In the following sections, we present questions that help to structure a fieldwork simulation and can be used to identify factors that need to be captured by the simulation including variables that we are concerned about, variables that we are capable to making decisions about, and constraints. Of course, the first step is always to clearly define the research objective, identify the purpose of the monitoring program or clearly articulate the reasons for why data are being collected (Fig. 1).

Ecological variables

Once the purpose of the field study has been determined, the next step is to determine what types of data will fulfill this purpose. The ecological variables reflect the "true" state in the simulation and the "real" data, as we perceive them in the world. Specifying ecological variables concisely reflects what variables will help address the purpose of the field study. Important questions to consider here are as follows: What are the data? Do they vary across space or time? How are they measured? When is it best to measure them? Ecological variables also can include "nuisance variables"; these are measurable variables that do not directly serve the study's purpose but are correlated with the response variables. Including them in the analysis is expected to improve inference, even though they do not directly address the study's purpose. In our case study, the values of the ecological variables inform data collection decisions such as how many individuals should be collected of a species at each site (Fig. 1).

Field conditions

Field conditions represent logistics and include information such as where the sites are, how many sites there are, how the sites differ, whether species occur at every site, how

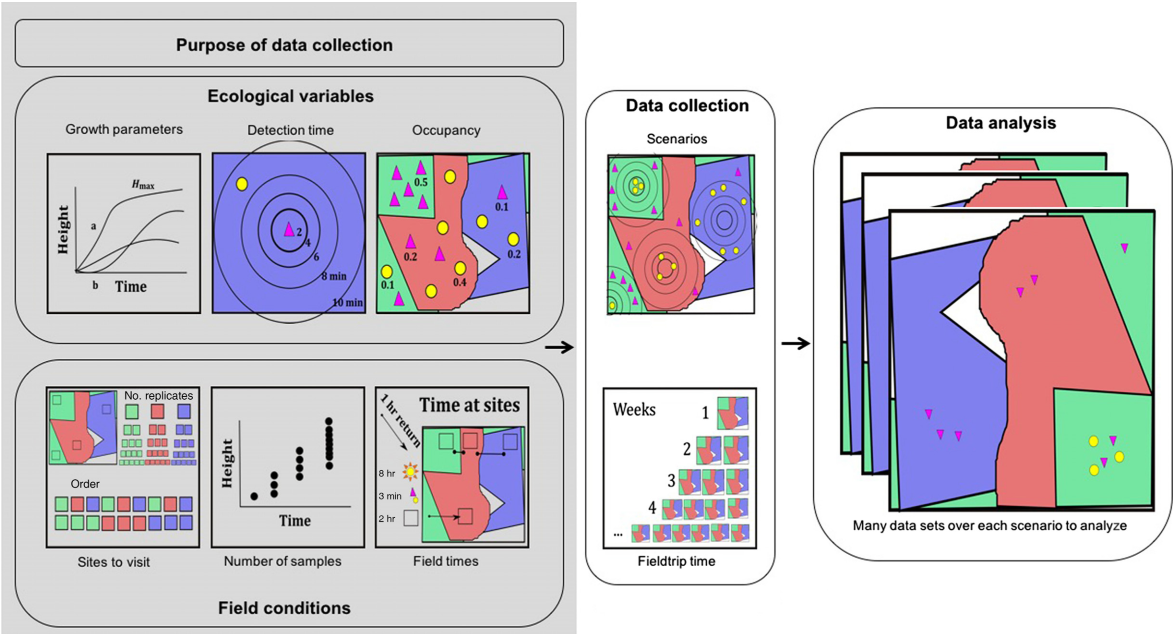

FIG. 1. The overarching structure of the simulation, with specific components taken from the mallee case study. Ecological variables are parameterized growth models of the target species, detection times for individual plants, and each species' probability of occupying each site. Field conditions include the number of sites to visit, the number of individuals per species to collect, and fixed parameters describing how long field days are and travel and setup times associated with fieldwork. Ecological variables and field conditions inform the data collection process, which simulates how much data are collected based on specified constraint scenarios. Data analysis takes many data sets collected over multiple scenarios to analyze and test model performance under different constraint scenarios.

long it takes to move between sites, how long it takes to travel to sites from field accommodation, and how long it takes to set up sites.

Data collection

The data collection process captures both the variable combinations to be replicated and which of these variables most affect the sample sizes. This can be a point to incorporate uncontrollable constraints (e.g., the chance of no data collection on some days due to bad weather) and fieldwork time constraints (e.g., a maximum of 10 working hours per day). Attributes that reflect survey design can be included, for example, the minimum or maximum time allocated to one site. Simulated data are collected, and the field time required to collect each data point is recorded, so that the amount of data collected under different time constraints, given ecological variables and field conditions, can be compared.

Data analysis

The fieldwork simulation imitates the data set that might be produced from a single field trip of a given duration. This single field trip can be replicated many times to account for natural variation between each visit into the field. The accumulated data from multiple field trip replications can then be analyzed to determine under which conditions the data collected suits the analysis envisioned (Fig. 1). This is a key point with which to interrogate the analysis methods, as well as the overall purpose of the study. Are the right data being collected? Is it possible to collect enough data for the analysis? How much extra time will it take to collect an adequate amount of data? Is this feasible? Can the purpose of the study be achieved?

CASE STUDY

This case study uses data on the height of multiple plant species across a chronosequence (a space-for-time substitution [Walker et al. 2010]) of time-since-fire sites in Murray Sunset National Park, Victoria (34.7683° S, 141.8542° E), a large conservation area within the semi-arid Murray Mallee region of southeastern Australia (for detailed methods, see Thomas and Vesk [2017*a*]). The vegetation and fire histories of the Murray Mallee region have previously been comprehensively mapped (Haslem et al. 2012, Avitabile et al. 2013). Our field sites were 11 areas varying in time since fire (1, 2, 4, 8, 13, 15, 26, 33, 36, 46, and 86 yr) that were visited as our space-for-time sequences to measure the increase in height of post-fire recruiting re-seeding plants through time. The purpose of collecting these data was to build multi-species models of plant growth to assess how plant functional traits may be used to predict plant species' growth trajectories both within an ecosystem (Thomas and Vesk 2017*b*) and between ecosystems (Thomas and Vesk 2017*a*).

With the aim of interrogating the efficiency of this fieldwork, the authors timed all aspects of their data collection. Times noted on a typical day included the time of leaving accommodation, the time of arrival at national park boundaries, the time of arrival at the first field site, the time taken to set up field gear, the time taken to find the first plant to measure, the time taken to measure the plant, the time taken to find the next plant, the entire time spent at one site, and the time spent driving to another site. Times measured were based on two field ecologists measuring plant heights (one researcher and one volunteer). This account of time for field-based data collection was used to inform the simulation of the field data collection process. Additionally, we used modeled estimates of the height data collected for each of the plant species to set up the simulated "true growth parameter states" (Appendix S1: Table S1).

Ecological variables

We begin the simulation by setting up "true" states, in the form of three matrices: one holds the mean height for each species at each observed time since last fire in the chronosequence, the second holds the species-specific mean time required to find each individual at a site (Appendix S1: Table S1), and the third holds the occurrence probability of each species (Appendix S1: Table S1). By changing each of these matrices, we can explore different scenarios of on-ground constraints (Fig. 1).

The "true" mean plant heights are generated using the Hillslope equation (Tjørve 2003), which is a re-parameterization of a logistic equation. Its three growth parameters are biologically interpretable and relate to the maximum height achieved (parameter $H_{\max}$), the maximum relative growth rate (maximum RGR, parameter a in cm·cm^{-1}·yr^{-1} (Atwell 1999)), and the time at which maximum growth occurs (parameter b, in years)

$$H_{i,j} = \frac{\mathrm{Hmax}_{i,j}}{\left(1 + \exp\left[-a_{i,j}\left(T_i - b_{i,j}\right)\right]\right)} \tag{1.1}$$

where $H_{i,j}$ is the observed height of individual plant i of species j (cm) and T_i is the time since fire (yr) at which individual i is observed.

Number of individuals collected.—We specify the minimum and maximum number of individuals we want to collect per species per replicate site with a given time since fire. This may be a single value (e.g., exactly five individuals per species) or a range (e.g., we can collect at least three but no more than 10 individuals of each species within each site).

Field conditions

Field site selection.—For each time-since-fire value, there may be multiple replicate sites that can be visited. Sampling across a broad range of time since fire values will enable effective estimation of many species' growth curves. However, the species available to sample may vary among replicate sites with the same time-since-fire value. Thus, we specify the number of replicates to visit, and the order of time-since-fire values? in which the sampling takes place. For example, we may wish to visit early, mid, and late time-since-fire sites once before getting replicates of each, or alternatively we might wish to visit multiple replicate sites of a given time since fire to collect a broader range of species, before moving to a site with a different time-since-fire value.

Travel and measurement times.—Additional fixed parameters are set that relate to the total number of hours available to work per day, the available time-since-fire chronosequence and replicate sites available, the home-to-site travel time, the time it takes to measure individuals, the travel time between sites, the setup time required at each site, and the maximum time allowed to spend at a site (Appendix S1: Table S2). We set up different available field times in order to explore field trip durations from 7 d in weekly increments up to 200 d of total time spent during a field trip.

Data collection

The process of field data collection is simulated in an R function (see *Data Availability*; Fig. 2). In its first phase, it

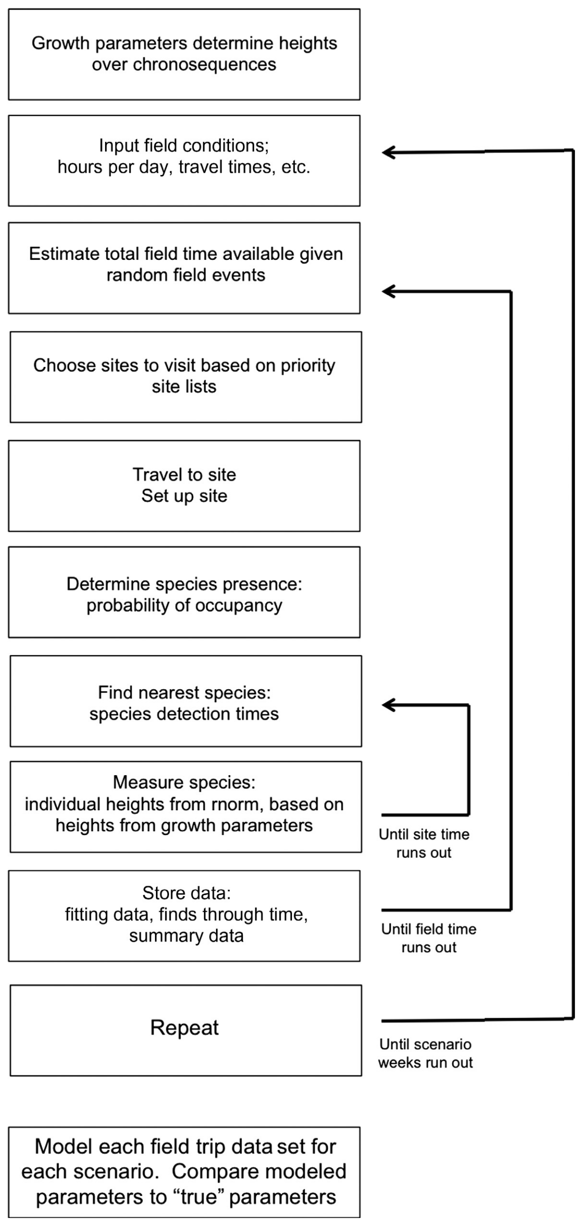

FIG. 2. The data collection process generates multiple data sets across many field trip scenarios. Individual heights are randomly drawn based on mean heights through time generated from modeled heights based on "true" growth parameters.

identifies the time available for on-ground data collection. We use a custom "problem" probability distribution to describe the probability of half days of fieldwork being removed from the overall fieldwork time due to unforeseen circumstances, such as flat tires or extreme weather (using our field-based data to inform this). The number of realized fieldwork days available is drawn from this distribution. From the number of realized fieldwork days, we subtract the time required to travel to and from the field sites each day. The time available for finding and measuring individuals at sites accounts for the travel time between sites and the setup time required at each new site.

We use three arrays to store the simulated data. (1) "Fitting Data" represents a regular field data sheet, storing each species ID, individual height and the time-since-fire value and replicate it was collected in. (2) "Finds through time" records the order of species measured, the time elapsed, and the height of the measured species. (3) "Summary data" records the species identity, occupancy, the number of individuals collected for each species, a flag indicating whether we have collected the required number of individuals, and mean species height based on the site's time since fire.

With the data collection time calculated and storage arrays defined, the simulation of data collection can proceed. Given the total number of sites, the algorithm chooses which site and which replicate to visit based on a "priority list" (an ordered list of sites and replicates within sites to visit based on the researcher's field design) designated in the field site selection code. Each species' presence at a site is drawn from the occupancy probability distribution, and it is possible to search for this species if it is present at a site and the maximum number of individuals has not yet been collected. A setup time elapses, and then individuals are sought, with the time to detection drawn from a species-specific exponential distribution. Time elapses as the individual is measured. Height values are drawn from a lognormal distribution with parameters specified by the height-growth model (with variation specified in Appendix S1: Table S2). The algorithm continues searching, measuring and recording randomly drawn individuals until the pre-defined number of individuals of all required species have been found or until the total time allowed at that site elapses. The algorithm then moves to the next site based on the field design (priority list) until the total field trip time elapses. All data are stored in arrays (Fig. 2).

Data analysis

The field trip function collects data from each simulated field trip (see *Data Availability*; Fig. 2) and uses the Bayesian modeling package jags (Plummer 2013) via the statistical software environment R version 2.5.2 (R Core Team 2015) with the package R2jags (Su and Yajima 2015) to fit a multispecies, hierarchical, three-parameter, nonlinear, growth model (Eq. 1.1; Thomas and Vesk 2017*a,b*).

Estimating power and choosing sites

Number of individuals.—We wanted to compare the simulated results to a benchmark with sufficient individuals per species, replication, and diversity of time-since-fire sites for an acceptable performance of parameter estimates in the model. To develop these benchmark values, we freed the simulation from time and species-specific constraints, and controlled the number individual plants collected (ranging from 1 per species up to 20 individuals per species) over all time-since-fire sites. We fit these data to the same height-growth models and compared model fit between this approach and the context-specific simulations. These simulations reveal the number of individuals per species needed to generate stable parameter estimates (i.e., collecting more data would not reduce the variation around the estimates significantly).

Nature of chronosequences.—In reality, the number of sites that can feasibly be visited may be restricted. We simulated a range of possible constraint scenarios capturing various numbers and configurations of time-since-fire values; we also explored low and high goals for the number of individuals collected per species (5 and 20, respectively).

A key design question when undertaking our case study was whether to visit all time-since-fire values equally, to subset sites and focus on capturing more data in early, mid, or late growth patterns, or to sample sparsely across the range of time-since-fire values available. We used our simulation to test some different strategies, named "early" (time-since-fire ages 1, 2, 4, 6, 8 yr), "late" (time-since-fire ages 28, 33, 36, 41, 86 yr), "middle" (time-since-fire ages 8, 13, 15, 26, 28 yr), and "sparse" (time-since-fire ages 1, 13, 26, 36, 86 yr) strategies, comparing these to the "all time-since-fire set" (time-since-fire ages 1, 2, 3, 4, 6, 8, 13, 15, 26, 28, 33, 36, 41, 86 yr). This all sampling represents the fieldwork undertaken in the Murray Sunset National Park. We simulated the collection of five individuals per species per time-since-fire over each of these chronosequences (we also replicated this using 20 individuals per species, see Appendix S1: Fig. S1). We compared performance of the statistical models fit to these data using precision and capture rate (see Incorporating realistic field-based constraints for details).

Incorporating realistic field-based constraints

The aim of these case study simulations was to determine how differing quantities and kinds of data affect the modeled estimates of plant height growth. Additionally, we aimed to assess how realistic fieldwork constraints based on time impact the ability to collect an adequate amount of data. We achieved this by simulating a number of different data collection scenarios with differing numbers of individuals collected per species and differing spread of chronosequence sites. This allowed us to assess how sample sizes and site type affected the parameter estimates of our growth model (Eq. 1.1).

Time is then incorporated, to simulate a realistic field-sampling situation. Time elapses during travel from the researcher's home base to the field accommodation, travel between field accommodation and field sites, between field sites, during site setup, and when finding and measuring species. We compare four broad "constraint scenarios": a baseline or "naive" scenario where time minimally constrains data collection, a scenario adding travel times; scenarios involving travel times and species measurement time; and scenarios involving travel times, species detection time, and variable occupancy of species across sites, which we refer to as the “Mallee” scenario (Appendix S1: Table S2). We use time as a surrogate for cost, but replacing it or including monetary cost as well as time into the simulation would be straightforward. Each scenario was replicated 20 times (e.g., representing 20 field trips under the same conditions), producing many arrays of data across multiple dimensions (i.e., estimates for each of the three parameters for different numbers of individuals collected for each species for each field trip under each broad scenario and with differing field trip duration). For this reason, the data we present are the mean parameter estimates averaged across all species for each replicated field trip, with the mean and standard deviation across all field trips represented. We use three different evaluation metrics to compare model performance across scenarios. First, bias measures how far away and in which direction parameter estimates from simulated data sets are from true height-growth parameters. Second, uncertainty was measured as the width of the 95% confidence intervals around parameter estimates from simulated data sets. Third, capture rate measured the number of times that the true growth parameters were contained within the 95% credible intervals of modeled parameter estimates from different scenario data sets.

Code Availability

For the original code for this simulation, as described in this paper see *Data Availability*. Our current and future research aims are to make this code more general for multiple types of field trips and models.

Results

Our suite of simulations revealed practical information, such as the minimum number of individuals per species to adequately characterize all three of the growth parameters, but it also helped to diagnose model behavior. Sample size affected the estimates of the three model parameters differently (Fig. 3). As the number of individuals sampled for each species increased, the capture rate for parameter H_{max} (average top height) decreased, while parameters a (maximum relative growth rate) and b (age at maximum relative growth rate) have a relatively consistent capture rate across sample size. Parameter H_{max} was consistently negatively biased; on average underestimated by 10 cm across a range of sample sizes (1–20 individuals per site). Bias decreased in parameters a and b as sample size increased. Model uncertainty decreased for all parameters with increasing sample sizes. Estimates of H_{max} were increasingly overconfident as sample size increased, which explains the decrease in capture rate for this parameter: it is the outcome of bias (10 cm) and narrowing of the confidence intervals that lead to the decline in the capture rate with increasing sample size (Fig. 3). Based on this information, between three and five individuals for each species would be adequate to provide robust parameter estimates (the confidence intervals stabilize at these sizes), and collecting twenty individuals per species may not provide enough extra model accuracy to offset the extra fieldwork time needed.

Chronosequence types

The spread of time-since-fire values available for the chronosequence approach influenced model performance (Fig. 4). H_{max} was relatively easy to estimate across a range of chronosequence types, but a and b were difficult to characterize with a sparse set of times since fire. Early and all strategies were good candidates for minimizing uncertainty and bias in parameter estimates, although parameter b in the all strategy had a very low capture rate as a consequence of having some positive bias and narrow credible intervals (Fig. 4).

Effect of field constraints

Given negligible constraints on handling and travel time, species detectability, and assuming all species occupy all sites (the baseline constraint scenario), one week of data collection for 20 species did not produce an adequate data set to parameterize this model, as parameters a and b did not converge and, as such, were very uncertain (Fig. 5, Appendix S1: Fig. S2). Capture rate appeared to be high for one week’s fieldwork (i.e., close to 20; Fig. 5), but this was an artefact of large credible interval width (large uncertainty) in parameter estimates leading to a large “capture area” of underlying data points. One week was insufficient for model parameterization under any scenario of field constraints.

Increasing the time available in the field improved model fitting, but not uniformly across the constraint scenarios. The least constrained scenarios improved generally except in the bias and thus capture rate in H_{max}, as seen above when the number of individuals per species was varied (Fig. 3). Yet, as more constraints were added, the benefits of additional available field time were frequently reduced.

The full Mallee scenario included the most realistic constraints, considering deterministic travel times, as well as stochastic detection and occupancy of species. This stochasticity yielded more variation between field trips in estimated parameters as time in the field increased. This variation was likely due to a trade-off between sampling fewer species but more individuals per species in short field trips (i.e., sampling common species well) vs. more species but less sampling within all species at longer times (spending time finding rare species). This trade-off arose due to the defined searching strategy of always measuring the closest species until the maximum allowable number of each species was reached. While seven weeks may seem adequate for estimating the model parameters, harder-to-find species were not always represented in these data sets (Fig. 6).

Simulated data sets

Simulations can be used to estimate the amount of time it will take to collect a sufficient data set for analysis. Incorporating the constraints on field data collection influenced the number of individuals and species sampled in this case (Fig. 6). When species occupancy and detection were not included, time allocated to effective data collection was underestimated. There was a 40% reduction in the average number of individuals collected through time under the

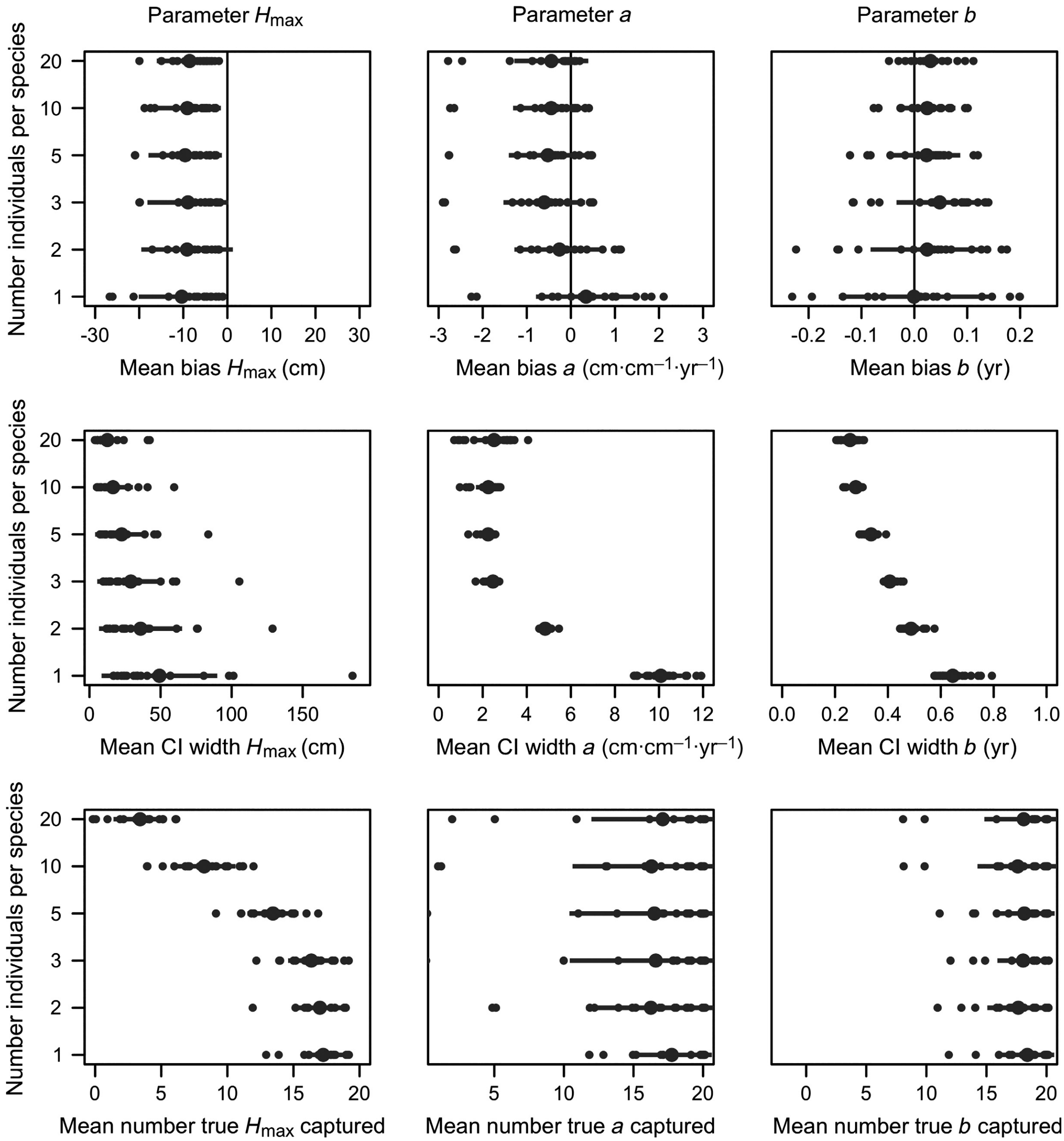

FIG. 3. Sample size influences parameter estimates. Bias, uncertainty, and capture rate for the three estimated growth parameters, as a function of the number of individuals collected across all species. Individual small black points are the outcome of each of the 20 replicate field trips averaged across all species, with the larger black dot the average across replicate field trips, and the black lines represent 95% confidence intervals across each replicate field trip.

most constrained, the Mallee scenario, and the naive baseline (Fig. 6). All species were found for any length of field trip under every scenario except the Mallee scenario, for which there was a chance of collecting all 20 species only after four weeks in the field but where even 150 d in the field did not guarantee finding all species (Fig. 6).

DISCUSSION

This work has demonstrated how the interactions of sampling design, sampling process, and field trip constraints affect parameter estimation in a multispecies height-growth model. It contributes specifically to the aim of this case study, which was understanding how mallee plant heights could be sampled and modeled. More generally, this study articulates the steps and elements of field study simulation and evaluation. These are expanded here.

Our simulation approach starts from the perspective of a field ecologist departing for a field trip to stand at a site looking for their first "data point." This is not in fact just a "data point," and it is an individual organism that requires time to find and time to measure. This approach is most valuable in cases where data are hard to collect, large sample sizes are required, and travel times are long. Fieldwork takes

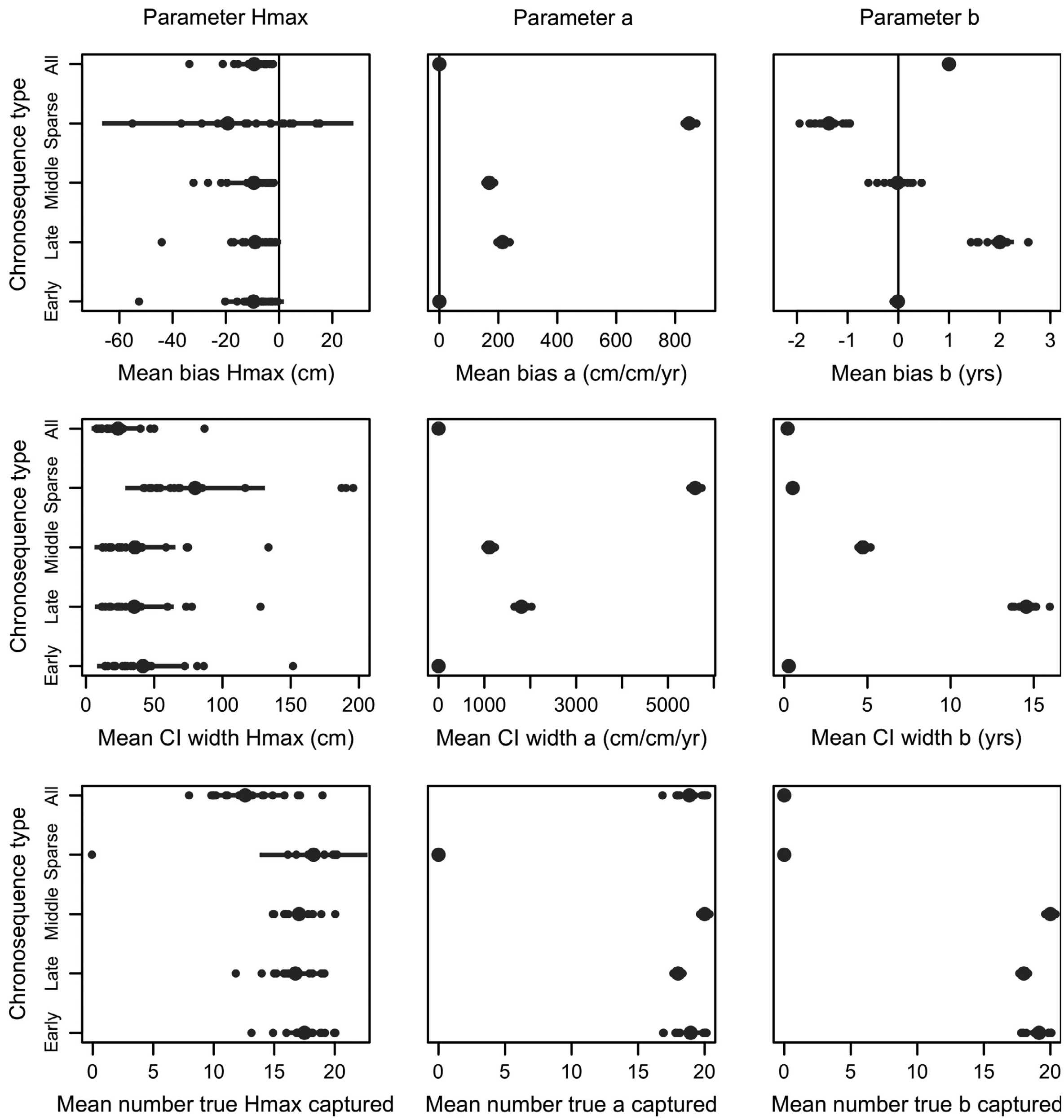

Fig. 4. The arrangement and times available in a chronosequence field design affect model parameter estimates. Bias, uncertainty, and capture rate for the three estimated growth parameters, as a function of the type of time-since-fire sites visited. Individual small black points are the outcome of each of the 20 replicate field trips averaged across all species, with the larger black dot the average across replicate field trips, and the black lines represent 95% credible intervals across each replicate field trip.

time and money, and natural variability generates additional time-consuming constraints, which are hard to predict. We aimed to design a simulation to mimic realistic field conditions. Natural variability was propagated throughout each simulated field trip in multiple ways, including incorporating the random removal of half days from the overall fieldwork time allocated and generating random height values. After we simulated multiple field trips for each scenario we wanted to test, we then analyzed the simulated data to evaluate how the realistically collected sample sizes influenced the parameter estimates within the chosen model. Finally, we explored various field times to account for realistic time and budget constraints, and tested how the available time influenced the analysis and the ability of the field plan to achieve the research goals. This approach to simulation is structurally different to other simulation approaches, which typically start from the analysis and back transform to estimate an optimal data set. This simulation approach is a context-specific approach to simulation rather than a context-free approach.

Our suite of simulations allowed us to generate information analogous to a traditional power analysis but for a more complex nonlinear hierarchical model; it revealed the minimum number of individuals (between three and five) per species to adequately characterize each of the three growth parameters in the nonlinear growth model. We were

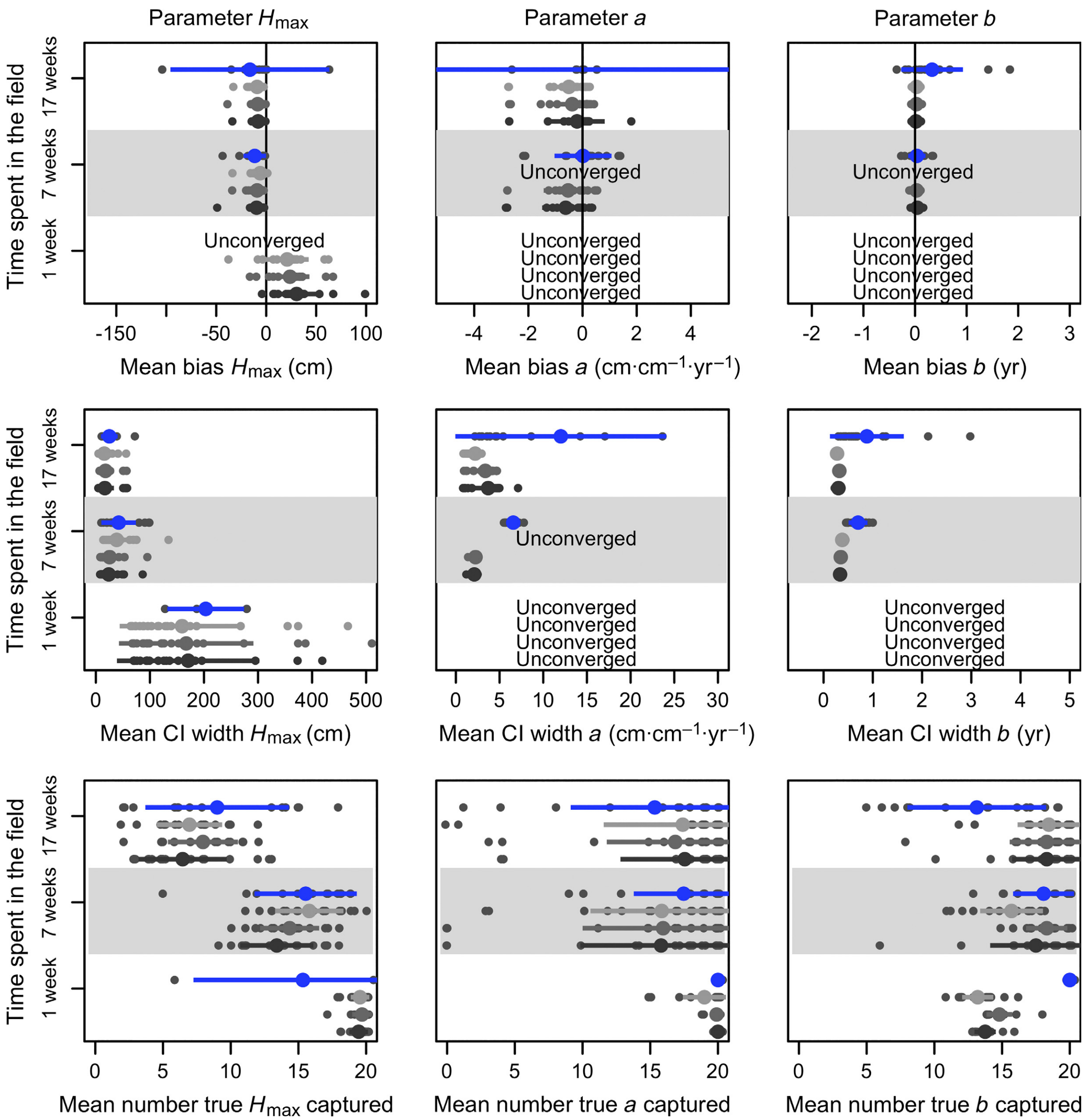

Fig. 5. Parameter estimation performance under different scenarios, each tested over different lengths of field trip and averaged over all species (see Appendix S1: Fig. S2 for results displayed over all weeks). Each of the three parameters is compared across bias, uncertainty, and capture rate. The black lines are baseline naive data sets, which do not include travel and measurement times or species detection and occupancy, dark gray lines are the scenario where only travel time is included, light gray is scenario where travel time and measurement times are included, and blue is the full Mallee scenario, which includes all time costs as well as occupancy and detection of species. Unconverged refers to model runs that did not successfully converge.

able to test the model across varying sample sizes, which was useful for diagnosing model behavior, such as consistent under-prediction of top height (H_{max}) by about 10 cm. Additionally, we were able to test the sensitivity of the model to the ages of chronosequence sites, which provided us with confidence that the data were fit for purpose but also provided valuable information for guiding future sampling efforts. Even intensive data collection was unlikely to yield robust parameter estimates if we were limited to sampling sites sparsely distributed across time-since-fire values. This would ideally prompt us to re-think the study design. The above two examples, selecting numbers of species and testing the sensitivity to numbers of sites, could conceivably be achieved through power analysis, and this is often performed using a variety of packages (see *Introduction*). However, we note that our approach is not bedded in significance testing, though it could easily be accommodated.

Designing robust field-based research is not only about adequate sample sizes. Many common approaches to simulating data sets for complex analysis do not account for realistic field-based time constraints nor species detection rates (but see optimizations such as MacKenzie and Royle 2005, Guillera-Arroita et al. 2010, Moore and McCarthy 2016). While we cannot possibly anticipate all the fieldwork constraints we may encounter, we can account for many field-based parameters and incorporating practical time

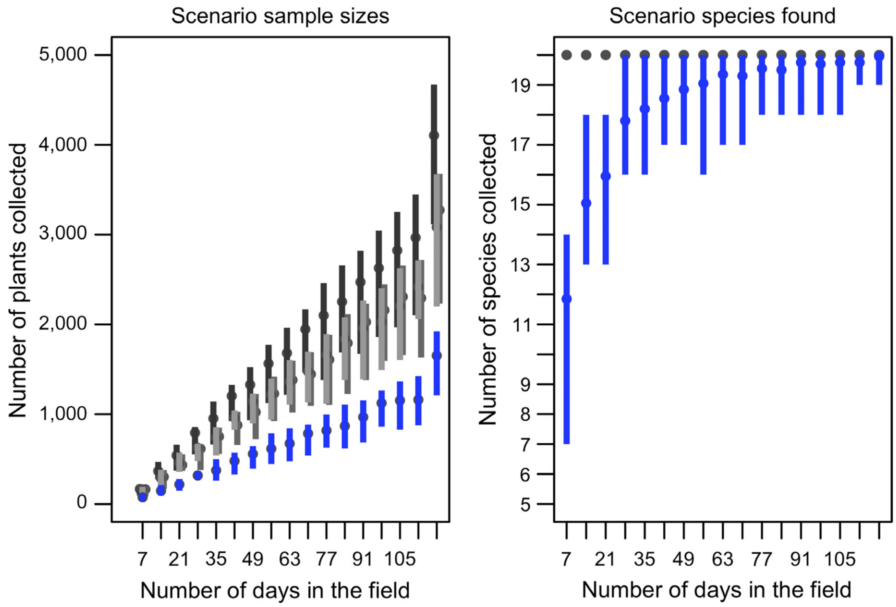

FIG. 6. The number of individual plants collected for each scenario, and the number of species found for each scenario. Black is naive baseline, dark gray is travel, light gray is measurement and travel, and blue is the full Mallee scenario. All scenarios except the Mallee (blue) find at least one individual of all species under all field trip lengths.

constraints into the simulation allows us to test the feasibility of studies. We can bound expectations of data between best-case and worst-case scenarios including natural variation arising from multiple sites and species. For example, incorporating measurement time and travel time is important for a field plan, particularly if there are large travel costs, which is information that is often readily available or easily estimated. Our results demonstrate that when designing field studies it is crucial to account for species occupancy and detection, because ignoring them causes the time required for effective data collection to be vastly underestimated. Our results suggest that deterministic features of fieldwork (travel times) have relatively predictable effects on resultant data sets, but stochastic features of fieldwork (species occupancy) lead to considerable variation between replicate field trips. This is likely to be particularly important if the analysis planned is sensitive to the number of species, for example, hierarchically structured models that may be differentially sensitive to sample sizes at different levels of the data set (Paccagnella 2011). Finding every species may be important, and a simulation approach could be used to allow the field ecologist to explore trade-offs in spending less time at a site but visiting more replicate sites or spending more time searching fewer sites in order to find rarer as well as common species.

Conducting realistic field-based simulations before fieldwork that includes specific and targeted data collection methods forces the field ecologist to confront the constraints of their plan early. It helps to reveal aspects of the field plan or analysis plan that are flexible and can be controlled, as well as aspects that are impossible or harder to change (Peck 2004). This allows forethought into trade-offs in sampling design. For example, studies of chronosequences are commonly limited by availability of site "ages," and this might mean an ideal analysis is not feasible. In our case study, the spread of time-since-fire sites available mattered and sampling some chronosequence combinations (sparse or late) were inappropriate for the three-parameter model. We also now know that if time was limiting, we would focus on collecting early time-since-fire sites rather than capturing a broad spread of sites. Knowing these field design sensitivities provides an opportunity to change direction in either field approach or analytical approach before a lot of time and money is spent.

Johnson et al. (2015) suggested that for simulations to be useful for designing better studies, they must give substantially more accurate estimates of sample sizes than conventional power analysis and be reasonably straightforward to use so as to justify the extra time and effort required for the simulation. A counter-argument is that ecologists commonly spend a lot of time and effort improving the analysis of messy, suboptimal data. Less effort is directed to improving data collection, despite emerging tools to do so. Creating a simulation, particularly one such as this, is time-consuming and complex. However, the use of simulations is likely to become easier in the future thanks to increasing efforts to create packages and functions that encourage researchers to use simulations to aid complex field design and analytical approaches. In agreement with Peck (2004), we argue that the process of creating a simulation has value beyond optimal sample size estimation. For example, simulation of field data collection and analysis of simulated data is a very effective way to make the best use of pilot and pre-existing data. The design of simulations is the design of an experimental system (Peck 2004) and requires forethought and planning of data collection and analysis, which may be daunting for those not conversant with complex analysis techniques. However, undertaking a simulation provides an excellent incentive to learn and interrogate data and data analysis approaches, and to ensure that both data and data analysis are fit for purpose. Wider adoption of this approach may help bring together ecologists with expertise in field collection methods with experts in data analysis in the planning stages of projects, which would be to the benefit of all.

Beyond providing guidance on fieldwork design and analysis, our simulated approach and its realistic time constraints have value in providing transparent and robust information for budgets. This helps the ecologist to assess the consequences of design decisions, such as whether more field days (and costs related to those) are justified by the likely increase in the data set, or whether staying in cheaper accommodations further away from field sites is economical overall. This is relevant to academic researchers who may have a limited budget, but also to organizations responsible for designing field programs with a commitment to transparent budgets. Examples of this might be publicly funded monitoring projects assessing environmental impacts (Langford et al. 2011) or studies targeting rare or threatened species (Rueda-Cediel et al. 2015). Efficiencies in field design are likely to save significant time and money for large-scale programs. For this reason, a logical and useful extension to a realistic simulation such as demonstrated here would be to include a budget of the time and equipment needed for entering, storing, and analyzing data, because budgets for field programs often ignore these components (Jones 2013).

Our simulation approach is flexible but currently specific to our case study, making it hard for others to immediately implement. We anticipate several research directions that will improve the usability of our approach. First and ideally, our approach could be tested by implementing it in new and independent fieldwork studies. Second, we are currently focusing on making our code for this simulation more general and efficient. We aim to make our simulation approach more general so that it can be used across other types of fieldwork beyond our chronosequence-based fieldwork for the purpose of growth modeling. Generalizing the code in this way involves conceptually mapping out some common approaches to fieldwork and analysis, and isolating common influential variables and workflows.

Chapman (2000) suggests that experiments are often designed to minimize or ignore natural variation rather than to measure it. Kain et al. (2015) suggest that introducing new strategies for analyzing variance and designing field studies may lead to inspiring new experimental designs in ecology. Simulation approaches have enormous practical potential. Using simulations that connect the purpose of fieldwork and outcomes of research with data analysis and real estimates of field-based effort can help maximize the effectiveness of projects given limited funding. We see an exciting future blending simulation with robust field design to address a range of ecological fieldwork purposes.

Acknowledgments

We thank Daniel Falster for early discussions on simulation structure; we thank Natalie Briscoe for a friendly and constructive review, and two anonymous reviewers for further recommendations and improvements. We thank Rafael Schouten for current and ongoing development of simulation code to increase efficiency and generalizability. F. M. Thomas is supported by an Australian Postgraduate Award, Holsworth Wildlife Research Scholarship and a top up from Australian Research Council Centre of Excellence in Environmental Decisions (CEED). P. A. Vesk is supported by CEED. C. E. Hauser was funded by the NERP Environmental Decisions Hub.

Literature Cited

Atwell, B. 1999. Growth analysis: a quantitative approach. Pages 188 *in* B. Atwell, P. Kriedemann, and C. Turnbull, editors. Plants in action. Macmillan Education, Melbourne, Australia.

Avitabile, S. C., et al. 2013. Systematic fire mapping is critical for fire ecology, planning and management: A case study in the semi-arid Murray Mallee, south-eastern Australia. Landscape and Urban Planning 117:81–91.

Bunnefeld, N., E. Hoshino, and E. J. Milner-Gulland. 2011. Management strategy evaluation: a powerful tool for conservation? Trends in Ecology and Evolution 26:441–447.

Chapman, M. 2000. Poor design of behavioural experiments gets poor results: examples from intertidal habitats. Journal of Experimental Marine Biology and Ecology 250:77–95.

Day, R. W., and G. P. Quinn. 1989. Comparisons of treatments after an analysis of variance in ecology. Ecological Monographs 59:433–463.

Di Stefano, J. 2003. How much power is enough? Against the development of an arbitrary convention for statistical power calculations. Functional Ecology 17:707–709.

Fairweather, P. G. 1991. Statistical power and design requirements for environmental monitoring. Australian Journal of Marine and Freshwater Research 42:555–567.

Field, S. A., P. J. O'Connor, A. J. Tyre, and H. P. Possingham. 2007. Making monitoring meaningful. Austral Ecology 32:485–491.

Green, P., and C. J. MacLeod. 2016. SIMR: an R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution 7:493–498.

Guillera-Arroita, G., M. S. Ridout, and B. J. Morgan. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1:131–139.

Haslem, A., et al. 2012. Time-since-fire and inter-fire interval influence hollow availability for fauna in a fire-prone system. Biological Conservation 152(C):212–221.

Hauser, C. E., and M. A. McCarthy. 2009. Streamlining 'search and destroy': cost-effective surveillance for invasive species management. Ecology Letters 12:683–692.

Hauser, C. E., K. M. Giljohann, M. Rigby, K. Herbert, I. Curran, C. Pascoe, N. S. Williams, R. D. Cousens, and J. L. Moore. 2016. Practicable methods for delimiting a plant invasion. Diversity and Distributions 22:136–147.

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

Jones, J. P. G. 2013. Monitoring in the real world. Pages 335–347 *in* B. Collen, N. Pettorelli, J. Baillie, and S. Durant, editors. Biodiversity monitoring and conservation: bridging the gap between global commitment and local action. First edition. John Wiley and Sons, London, UK.

Kain, M. P., B. M. Bolker, and M. W. McCoy. 2015. A practical guide and power analysis for GLMMs: detecting among treatment variation in random effects. PeerJ 3:e1226.

Langford, W. T., A. Gordon, L. Bastin, S. A. Bekessy, M. D. White, and G. Newell. 2011. Raising the bar for systematic conservation planning. Trends in Ecology and Evolution 26:634–640.

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

Low-Décarie, E., C. Chivers, and M. Granados. 2014. Rising complexity and falling explanatory power in ecology. Frontiers in Ecology and the Environment 12:412–418.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105–1114.

Meyer, K. M., W. M. Mooij, M. Vos, W. G. Hol, and W. H. van der Putten. 2009. The power of simulating experiments. Ecological Modelling 220:2594–2597.

Milner-Gulland, E. J., et al. 2010. New directions in management strategy evaluation through cross-fertilization between fisheries science and terrestrial conservation. Biology Letters 6:719–722.

Moore, A. L., and M. A. McCarthy. 2016. Optimizing ecological survey effort over space and time. Methods in Ecology and Evolution 7:891–899.

Nicholls, A. O. 1989. How to make biological surveys go further with generalised linear models. Biological Conservation 50:51–75.

Paccagnella, O. 2011. Sample size and accuracy of estimates in multilevel models. Methodology 7:111–120.

Peck, S. L. 2004. Simulation as experiment: a philosophical reassessment for biological modeling. Trends in Ecology and Evolution 19:530–534.

Peterman, R. M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences 47:2–15.

Plummer, M. 2013. JAGS Version 3.4.0 user manual. http://www.stats.ox.ac.uk/~nicholls/MScMCMC14/jags_user_manual.pdf

Punt, A. E., D. S. Butterworth, C. L. Moor, J. A. De Oliveira, and M. Haddon. 2016. Management strategy evaluation: best practices. Fish and Fisheries 17:303–334.

R Core Team. 2015. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. www.r-project.org

Rueda-Cediel, P., K. E. Anderson, T. J. Regan, J. Franklin, and H. M. Regan. 2015. Combined influences of model choice, data quality, and data quantity when estimating population trends. PLoS ONE 10:e0132255.

Scheiner, S. M. 2003. Six types of species-area curves. Global Ecology and Biogeography 12:441–447.

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

Su, Y. S., and M. Yajima. 2015. R2jags: a package for running jags from R. R package version 0.5-7. //CRAN.R-project.org/package=R2jags

Tessarolo, G., T. F. Rangel, M. B. Araújo, and J. Hortal. 2014. Uncertainty associated with survey design in species distribution models. Diversity and Distributions 20:1258–1269.

Thomas, F. M., and P. A. Vesk. 2017*a*. Are trait-growth models transferable? Predicting multi-species growth trajectories between ecosystems using plant functional traits. PLoS ONE 12: e0176959.

Thomas, F. M., and P. A. Vesk. 2017*b*. Growth races in The Mallee: height growth in woody plants examined with a trait-based model. Austral Ecology 42:790–800 early view.

Tjørve, E. 2003. Shapes and functions of species–area curves: a review of possible models. Journal of Biogeography 30:827–835.

Tyre, A. J., H. P. Possingham, and D. B. Lindenmayer. 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters? Ecological Applications 11:1722–1737.

van de Pol, M., and van de Pol M. M. 2011. Package odprism. https://github.com/cran/odprism

Vos, P., E. Meelis, and W. J. T. Keurs. 2000. A framework for the design of ecological monitoring programs as a tool for environmental and nature management. Environmental Monitoring and Assessment 61:317–344.

Walker, L. R., D. A. Wardle, R. D. Bardett, and B. D. Clarkson. 2010. The use of chronosequences in studies of ecological succession and soil development. Journal of Ecology 98:725–736.

Williams, P. J., M. B. Hooten, J. N. Womble, G. G. Esslinger, and M. R. Bower. 2018. Monitoring dynamic spatio-temporal ecological processes optimally. Ecology 99:524–535.

Wunder, J., B. Reineking, C. Bigler, and H. Bugmann. 2007. Predicting tree mortality from growth data: how virtual ecologists can help real ecologists. Journal of Ecology 96:174–187.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16:446–453.

Zurell, D., et al. 2010. The virtual ecologist approach: simulating data and observers. Oikos 119:622–635.

Zuur, A., E. N. Ieno, and G. M. Smith. 2007. Analyzing ecological data. Springer-Verlag, New York, New York, USA.

Supporting Information

Additional supporting information may be found online at: http://onlinelibrary.wiley.com/doi/10.1002/eap.1801/full

Data Availability

The original code for this simulation is available from GitHub at https://doi.org/10.5281/zenodo.1400861.