A comparison of models for interval-censored plant cover data, with applications to monitoring schemes

Oliver L. Pescott[1*], Mark Jitlal[1], Simon M. Smart[2], Kevin J. Walker[3], David B. Roy[1] & Stephen N. Freeman[1]

1. Centre for Ecology & Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK

2. Centre for Ecology & Hydrology, Library Avenue, Bailrigg, Lancaster LA1 4AP, UK

3. Botanical Society of Britain and Ireland, Suite 14, Bridge House, 1-2 Station Bridge, Harrogate HG1 1SS, UK

* Corresponding author. E-mail: olipes@ceh.ac.uk; telephone: +44 (0)1491 692 215.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

Abstract

1. Plant cover data collected by monitoring schemes are often expressed on interval-censored scales to reduce field effort. Existing statistical approaches to such data may not make full use of available information, or may both induce bias and assume more precision than may be warranted, e.g. by analysing mid-points and disregarding the spread of observations within a class.
2. We compare four approaches to modelling such data: two established methods (the proportional odds model and generalised linear mixed models) and two novel methods that explicitly accommodate the interval-censored nature of much data on plant cover. Of the latter, the first is a maximum likelihood (ML) approach that incorporates knowledge of the metric interval in which each datum lies. The second uses a Bayesian approach to incorporate interval-censoring and random effects to account for variation in annual changes between sites. All four methods are compared using data simulated with parameter values derived from analysis of a long-term monitoring dataset.
3. We demonstrate that model choice can influence the quality of statistical inference, particularly between models that make simplifications for convenience of fitting, and those which combine realistic distributional assumptions with accommodation of imprecise observations. A comparison of three of the methods demonstrated that all provide good accuracy and increasing precision over time. A comparison of power across the three frequentist approaches showed higher power for the novel ML approach. This is likely to be due to this non-hierarchical method underestimating residual variance. The Bayesian model is not directly comparable, but the measure of belief in a negative trend considered here was generally high, providing gradual increases in the believability of a decline with increasing time, number of sites, initial abundance, and larger effect sizes.
4. Our results suggest that the use of hierarchical models for plant monitoring schemes, conveniently applied in a Bayesian context, will help to bring greater realism and sensitivity to assessments of population change, and allow the use of more of the underlying information contained within cover data. Interval-censored methods will also allow for the integration of long-term plant datasets collected according to different cover scales, as well as presence/absence data.

Tweetable abstract

Plant cover data may currently be modelled inefficiently. We explore two new approaches to getting the most out of interval-censored data.

Key-words: Hierarchical Bayesian models, participatory monitoring, abundance data, time series, generalized linear mixed model, random effects, sample size, variance, long-term monitoring, citizen science

Introduction

Monitoring schemes are a key part of national and global initiatives to gather evidence on changes to biodiversity (Collen *et al.* 2013). Such schemes are often long-term, with periodic analyses of data expected to provide evidence for shifts or stability in the abundances or distributions of species (Magurran *et al.* 2010; Dornelas *et al.* 2013). Robust statistical design is therefore an essential part of any such scheme if the resulting models are to be widely accepted as credible indicators of biodiversity change by policy-makers, managers and other scientists (Lindenmayer & Likens 2010; Gitzen *et al.* 2012). An important part of such work is to determine (i) the level of any error or bias in the models adopted and (ii) what intensity of monitoring activity is sufficient to detect a statistically significant change in abundance if one has actually occurred; that is, suitable power analyses or equivalent are recommended as best practice (Jones 2013).

Conducting an appropriate power analysis for a monitoring scheme involves deciding on a set of relevant scenarios to investigate, covering a range thought plausible once the proposed scheme is established. Important variables affecting the quality of inference include those that represent the underlying structure of the data, e.g. the number of years of monitoring, the number of sites monitored or the arrangement of repeated site visits in time and space (Urquhart 2012), and those that represent the hypothetical effect that the monitoring is intended to capture, e.g. changes in species' abundances or distributions within a specified time frame, which may be a constant change of a fixed number of organisms or area of cover per year, or a proportional change in such a measure. Temporal trends, of course, may also vary across sites. Simple, mathematically-explicit estimates of power are not available in such multi-faceted studies, but, in a classical framework, simulation-based approaches to power analysis (Gelman & Hill 2007; Bolker 2008) have meant that ecologists have increasingly had a greater ability to capture the complex generating processes that often characterise data collected by monitoring schemes. These include the possibility of modelling variation in trends over time at different sites through the use of mixed models (Gelman & Hill 2007; Miller & Mitchell 2014; Johnson *et al.* 2015). These approaches should help to ensure that the results obtained embody a greater realism; this may be particularly important for monitoring schemes, which often cover large geographic areas across which the drivers of change for particular species or habitats may vary. The inclusion of greater flexibility in the modelling of spatially-varying structures in power analyses is therefore likely to ensure that decisions made regarding the inauguration and funding of particular monitoring schemes are better informed (Miller &

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

Mitchell 2014), and may also help to avoid unrealistic expectations. A key challenge for such an approach is the derivation of realistic estimates of variance parameters to use in prospective analyses. Ideally pilot studies would always be conducted, although existing datasets collected using similar species and methodologies to those of the proposed schemes are also often used for convenience (e.g. Irvine & Rodhouse 2010; Lebuhn *et al.* 2013).

Particularly where plants are concerned, an additional challenge may exist: plant species' abundances are often recorded according to class-based scales, scales which typically attempt to discretise the visual cover assessments of surveyors. Interval-based cover, or 'cover-abundance', scales typically encountered in vegetation science include those initiated by Daubenmire, Domin and Braun-Blanquet (Damgaard 2009, 2014; Kent 2012; Peet & Roberts 2013). Here, the abundances of species within monitored areas or plots in a given year are only known each to fall within one of a certain number of exhaustive and mutually exclusive classes, ignoring potential errors of classification. To our knowledge, only one previous study (Irvine & Rodhouse 2010) has attempted to incorporate plant cover scales into a general approach to modelling for plot-based plant monitoring activity. Irvine & Rodhouse (2010) approached the problem from the point of view that plant cover scales are best treated as ordinal classifications, and, correspondingly, used the proportional odds model (Agresti 2002) to investigate the power to detect change in plant cover data collected using such scales, providing a general framework for such analyses. Indeed, ordinal models have sometimes been recommended as the most suitable approach to plant cover data collected using cover scales (Guisan & Harrell 2000). This is due to the fact that some of the most frequently used scales are not purely based on metric intervals; for example, some variants of the Domin scale define their lowest cover classes as combined cover-frequency scores, e.g. points 1 and 2 on one frequently used version of the Domin scale are given as "< 4%, 1 individual" and "< 4%, several individuals" respectively (Rodwell 1991). Some authors have provided transformations of such scales in order to provide approximate metric equivalents for all categories (Mueller-Dombois & Ellenberg 1974; van der Maarel 1979, 2007; Currall 1987); such an approach is attractive, because it would provide a means to combine information collected according to different cover scales, whilst also allowing the derivation of intuitive measures of change on the percentage cover scale (Damgaard 2009, 2014). However, it should also be pointed out that some vegetation scientists have objected to similar approaches in the context of descriptive multivariate analyses (Podani 2006), suggesting that such operations as substituting ordinal class memberships with, for example, mean values of percentage cover classes would "increas[e] uncertainty in the data considerably" (Podani 2006); despite this, other workers have emphasised the progress that has been made in plant ecology by making such simplifying assumptions (van der Maarel 2007).

If the intervals of the cover scale used have clear percentage-cover equivalencies, or these can be estimated through field trials, or otherwise agreed or approximated (Currall 1987; van der Maarel 2007; Damgaard 2009; Irvine & Rodhouse 2010), then a range of additional modelling techniques become available. Specifically, we suggest that techniques for modelling censored data can be applied to plant cover data collected using many of the standard scales widely used in plant ecology today, such as the variant of the Domin scale used in Britain and Ireland since its application to the British National Vegetation Classification (Rodwell 1991), the Daubenmire scale often used in North America, or the Braun-Blanquet scale popular in continental Europe (Peet & Roberts 2013). A metric interval-based approach should offer at least two advantages over the proportional odds model (in addition to the fact that it may allow for the combination of disparate data sources): it avoids the potentially unrealistic assumption of equal transition probabilities between cover classes intrinsic to this latter model; and, it enables the use of linear mixed modelling techniques when cover data are logit-transformed, allowing real estimates of trend to be estimated and providing the option of estimating hierarchical variance structures (Johnson *et al.* 2015). The modelling techniques that can be used include those within a Bayesian framework, making hierarchical models that combine the use of metric interval-censored data with random effects relatively straightforward to apply. A metric interval-censored approach may also be of broader relevance to scientists working with volunteers or land managers in other areas of environmental science, where other types of observation may be made according to interval scales with different types of censoring. For example, in citizen science, where volunteers may be requested to report some feature of the environment according to a scale that simplifies an underlying metric reality (e.g. Pocock & Evans 2014).

The aim of this paper is to investigate different options for modelling interval-censored plant cover data, both in order to potentially increase the realism and information content of prospective power analyses for plant monitoring activity, and to broaden the toolbox of techniques available to vegetation scientists, whilst also highlighting potential trade-offs for error, bias and variance. We achieve this by comparing inferences resulting from: (1) a proportional odds model that treats cover data as ordinal classes (Irvine & Rodhouse 2010); (2) a hierarchical model with random effects using data representing overly precise observations of plant cover, i.e. approaches that transform interval memberships into point estimates of cover (Currall 1987; van der Maarel 2007); (3) a novel, non-hierarchical, frequentist interval-censored linear model; and (4) a novel hierarchical Bayesian interval-censored linear model.

We provide *R* and JAGS code for the last two models respectively as supplementary material. We also discuss the appropriate circumstances under which our interval-censored models might be used over the approaches previously described in the literature.

Methods

Statistical Models

Model numbers are as follows: Model 1: the proportional odds model, which assigns observed cover data to ordinal classes; Model 2: a generalised linear mixed model (GLMM) using data representing (overly) precise observations of plant cover; Model 3: a novel non-hierarchical frequentist interval-censored linear model; Model 4: a hierarchical Bayesian model with interval-censoring. All models fitted in this paper use a set of interval cover classes based on a commonly used variant of the Domin scale (Table 1).

Models 1 and 2

The proportional odds model and GLMMs have been described frequently in the literature (Agresti 2002; Gelman & Hill 2007; Irvine & Rodhouse 2010; Johnson *et al.* 2015). However, we reproduce them here for ease of comparison with Models 3 and 4.

Model 1: The proportional odds model for ordinal cover data, where *K* is the total number of intervals, states that

Eqn 1 $logit\left[P\left(C_{i,j} \leq k\right)\right] = log\left[\frac{P\left(C_{i,j} \leq k\right)}{P\left(C_{i,j} > k\right)}\right] = \alpha_k - \beta\ Year_i;$

$P\left(C_{i,j} \leq k\right)$ is the cumulative probability of an observation being in interval *k* or less (i.e. in any of the following intervals: 1, 2, …., *K*-1); $C_{i,j}$ is the observed interval for the percentage cover at year *i*, site *j*; α_k is the intercept for the k^{th} interval (k = 1, 2, ..., *K*-1); β is the slope for year. If the abundance increases over time ($\beta > 0$), such that a species moves up the category levels, then $P\left(C_{i,j} \leq k\right)$ for $k < K$ becomes smaller over time.

Model 2: Generalised linear mixed model for raw cover data – the logit normal model:

Eqn 2 $git(p_{i,j}) = \log\left(\frac{p_{i,j}}{1-p_{i,j}}\right) \sim N(\mu_{i,j},\ \sigma^2);$

$p_{i,j}$ describes the proportional cover of the given species at year i, site j; $\mu_{i,j} = \alpha_j + \beta_j Year_i$; $\alpha_j \sim N(\mu_\alpha, \sigma_\alpha^2)$ - random intercepts; $\beta_j \sim N(\mu_\beta, \sigma_\beta^2)$ - random slopes; μ_α and μ_β are the mean intercept and slope on the logit scale. Model 1 was fitted using the *R* package 'MASS' v. 7.3-43 (Venables & Ripley 2002), whilst Model 2 was fitted using the package 'lme4' v. 1.1-10 (Bates *et al.* 2015).

Model 3 – Censored data

A non-hierarchical linear model was proposed by Walker *et al.* (2010) for interval-censored cover data. We extend this model here to derive our Model 3. We assume that the unknown percentage cover at site *j* in year *i* is expressed as a proportion p_{ij} and is observed only to lie within the interval (l_{ij}, u_{ij}) where $l_{ij} \geq 0$ and $u_{ij} \leq 1$. We then assume the logit-transformed proportion is normally distributed:

Eqn 3 $$logit(p_{ij}) = log\left[\frac{p_{ij}}{1-p_{ij}}\right] \sim N(\mu_i, \sigma^2)$$

Then the probability of an observation lying within (l_{ij}, u_{ij}) is simply $\Phi(U_{ij}) - \Phi(L_{ij})$ where $\Phi(.)$ is the cumulative distribution function of a normal distribution and U_{ij} and L_{ij} are respectively equal to logit(u_{ij}) and logit(l_{ij}). Finally, to account for change over time, we define the expected coverage in year *i* via $\mu_i = \alpha + \beta Year_i$ with α and β additional parameters to be estimated.

The adoption of interval-censored data in this way means that the model cannot be fitted via standard subroutines which might be used to fit random intercepts and a constant slope, or a simple generalised linear model, to point data. It is, however, straightforward to programme the log-likelihood for a set of interval-censored data and optimise to obtain maximum likelihood estimates of all parameters. *R* code for fitting Model 3 using the optimiser 'optim' (part of the base package 'stats' in *R*) is provided as supplementary material.

We note also the flexibility of this approach: percentages recorded exactly (i.e. $L_{ij} = U_{ij}$) are readily accommodated should they be available, as are simple records of presence/absence, which can be considered to lie in the intervals (ϵ, 1) and (0, ϵ) respectively, where ϵ is some arbitrarily small value – nor is it necessary for all observations used to be recorded using a consistent set of category limits, thus allowing for the combined modelling of datasets collected using different cover scales.

Model 4 – A hierarchical model for censored data

Model 3 assumes a constant rate of change in the odds of plant cover (*p*/1-*p*) at all sites. It is appealing, and more realistic, to consider this slope as a random effect, varying spatially; site-dependent intercepts can be similarly treated. The censoring of the data, however, make this difficult in a frequentist framework. Such a hierarchical model is more readily fitted via Bayesian methods, which we introduce here as our Model 4; Model 4 therefore combines the treatment of random effects (Model 2) with interval-censoring (Model 3). The hierarchical Bayesian model for continuous interval-censored cover data has the same form as Equation 3, but with a modification to the specification of the expected proportion, so that $\mu_{ij} = \alpha_j + \beta_j Year_i$ where α_j is the intercept at site j and β_j is the slope, also at site j, with α_j and β_j taking normal distributions. There is one slope and intercept per site for this model.

All Bayesian models were run in JAGS v. 3.4.0 (Plummer 2013) using the package 'R2jags' v. 0.5-7 (Yu-Sung Su & Yajima 2015) to call the program from *R*. For this, the response can be re-expressed as logit(p_{ij}) ~ $N(\mu_{ij}, \tau)$, where τ is the precision ($1/\sigma^2$). Vague normal priors were used for regression coefficients and uninformative gamma priors used for variance components, following standard advice (Gelman & Hill 2007). The total number of iterations for each of three chains was 50,000, with the first 10,000 values of each chain discarded as a burn-in; every fifth value in a chain was kept, resulting in 24,000 iterations being used for inference regarding posterior distributions. Values of the Brooks-Gelman statistic $\hat{R}$ were checked for evidence of satisfactory convergence for all parameters within individual model runs before running models in a loop for the calculation of summary statistics (Brooks & Gelman 1998).

Data simulation and scenarios

Plant cover data with a hierarchical variance structure were simulated by adapting methods presented by Gelman & Hill (2007 pp. 449-454) and Bolker (2008 pp. 156-161). We simulated response data directly on the logit scale, assuming normal distributions. Data were simulated to match the hierarchical variance structure of Model 4, that is, the most complex and realistic of the models under consideration. We also investigated the effects on error, precision and bias induced by Models 2 and 3, in which more restrictive assumptions about the data structure are imposed.

Countryside Survey data

To ensure the greater realism of this exercise, the mean values for all standard deviation hyper-parameters using for the simulation of response data were estimated from an existing UK long-term plant monitoring programme, the Countryside Survey (CS; Carey *et al.* 2008). These surveys have been carried out in 1978, 1990, 1998/'99 and 2007, and involve visiting the same fixed vegetation quadrats nested within a sample of 1 km squares dispersed across Great Britain. Both quadrat locations and 1 km squares were selected according to a stratified random sampling design. CS plant cover data used here were from 4 m^2 (i.e. 2 × 2 m) plots nested within larger plots (known as 'X'-plots within the CS; see Maskell *et al.* 2008 for a more detailed description of the data collection process). Within each quadrat, plant species presence is recorded and cover estimated to the nearest 5%, except in the range 1–10%, where unit cover estimates are made; species' presences with cover < 1% are recorded as 0.1%. Here, a small value, $\varepsilon = 0.0025$, was added to, or subtracted from, cover values equal to 0 and 1 respectively to avoid undefined values when fitting logistic GLMMs (Warton & Hui 2011). For each species with at least one (non-zero) observation at each time point in the CS dataset, a varying-intercept, varying-slope GLMM (Gelman & Hill 2007) was fitted to logit-transformed cover data from those quadrats surveyed in 1990, 1998/'99 and 2007 within the following widespread UK broad habitats: broad-leaved woodland; neutral grassland; calcareous grassland; and acid grassland. That is to say, a hierarchical version of Model 3 was used but data were considered as point estimates rather than as interval-censored observations; the 'sites' in the above model descriptions equal quadrats for the CS data, with no other spatial nesting considered. The *R* package 'lme4' (Bates *et al.* 2015) was used to fit all models to CS data. We also assessed the sensitivity of the estimation of the standard deviation hyper-parameters to the choice of species included in the model-fitting by restricting the data modelled to those species with at least two non-zero observations at all three time points (i.e. increasing the number of quadrats modelled for any one species, but reducing the number of species modelled overall).

Across the 653 species so modelled, the median values for the residual standard deviation (0.80) and the intercept (0.15) and slope (0.05) random effect standard deviations were calculated and rounded to the nearest 0.05; these were the standard deviation hyper-parameters used in the data simulation step. The sensitivity analysis (models fitted for 343 species) gave median values of 0.85, 0.30 and 0.05 for these parameters; given that the main focus here is on the relative performance of the different models, given a particular dataset, we used the initial set of standard deviations for all simulations (Table 2).

A simulation exercise

For all models, the investigated values of key parameters are given in Table 2, resulting in 48 different scenarios. These were chosen as realistic values linked to a new volunteer-based (i.e. 'citizen science') plant monitoring scheme recently launched in the UK (Walker *et al.* 2015; Pescott *et al.* 2015). We restricted our investigation to declines to limit the number of scenarios

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

investigated. Proportional declines were specified in terms of an annual trend in the log-odds of plant cover, i.e. a linear trend on the logit scale, but a proportional decline on the scale of the odds of plant cover (*p*/1-*p*; Irvine & Rodhouse 2010). Declines were simulated from a range of initial starting percentage cover values (Table 2). For all scenarios we recorded the power to detect trend after 3, 5 and 10 years of monitoring under Models 1-3, assuming one visit per site per year. A one-sided α level of 5% was used across all models run within the frequentist mode of inference.

Within a Bayesian framework power in the frequentist sense cannot be defined. Therefore, for Model 4, an indication of the strength of support for the inclusion of directional temporal change in the model was estimated by averaging the proportion of the posterior distribution for the slope coefficient (i.e. μ_β) that was < 0 across all simulations. This can be conceptualised as the average belief in a negative trend of any magnitude given the data observed. Whilst we have chosen this simple summary statistic for the current investigation, the Bayesian approach is highly flexible, and the posterior distribution of a parameter can be divided into different sections to estimate the relative beliefs in parameter values of different magnitudes (see Brooks *et al.* 2008 and King *et al.* 2008 for conservation-oriented examples). Results are reported for 100 simulations for all scenarios, irrespective of the mode of inference; the number of simulations was limited to 100 due to the computationally expensive nature of the Bayesian model.

Results

Statistical power, in the case of the frequentist models, increased with sample size, trend magnitude, time (i.e. duration of survey), and initial percentage cover (Figs 1, 2, 3). For Model 1 (the proportional odds model), Model 2 (the GLMM using cover-class interval midpoints), and Model 3 (the frequentist interval-censored linear model), power typically reached 80% within 5 years for large proportional declines (75% and 90%), irrespective of initial percentage cover (Figs 1, 2, 3). This was partly dependent on the number of sites monitored, in that where only 15 sites were simulated, power required 7-8 years to reach 80% for 75% declines at low initial percentage cover (Figs 1, 2, 3). Across all models, power typically remained low for the smallest proportional decline (30%), with scenarios with 30 or 50 sites and/or larger starting covers rising above 80% power within 8-9 years.

Within this general pattern, there were important differences in power between models (Fig. 4a, b, c), despite the fact that the underlying data were treated according to the same interval scale in each case. The GLMM fitted to the point estimates of cover derived from interval scale midpoints performed very similarly to the proportional odds model, with the 144 power estimates across all scenarios and time-points mapping closely to a 1:1 relationship between models, with no clear effect of trend magnitude or initial percentage cover on this relationship (Fig. 4a). The interval-censored linear model, however, exhibited higher power than both the GLMM and the proportional odds model for all scenarios where power had not reached 100% (Fig. 4b, c), particularly at intermediate values of the initial percentage cover and proportional decline parameters (Figs 1, 2, 3). The hierarchical Bayesian interval-censored model cannot be directly compared to the frequentist models, but the general trend in a belief in a negative slope of any magnitude matched the results from the frequentist models, in that a negative trend became a more believable feature of the underlying data with increasing sample size, trend magnitude, time and initial percentage cover (Fig. 5). One noteworthy feature of the Bayesian posterior summaries was the apparently reduced influence of all scenario variables (e.g. sample size etc.) on the strength of belief in a negative trend (Fig. 5), compared to the larger influence of these on frequentist power (Figs 1, 2, 3). For example, the Bayesian model indicated a greater than 75% belief in a decline after 6-7 years even with 30 sites and low covers; the frequentist models only approached higher levels of power for higher initial percentage covers, larger numbers of sites or longer periods of monitoring (Figs 1, 2, 3, 5). Similarly, the Bayesian posterior summaries exhibited relatively smooth increases in the believability of a negative trend over time, whereas the frequentist models tended to exhibit large jumps in power (Figs 1, 2, 3, 5). These differences are due to the fact that the summaries of the Bayesian models focus on belief in a negative trend of any size, whereas the power of frequentist models requires a significant trend to be detected at the $\alpha = 0.05$ level, creating a ‘stepped’ change in significance rather than a gradual increase in the believability of a phenomenon. Given that the declaration of significance in frequentist statistics depends on the degrees of freedom, and so the sample size, this difference in how the importance of results is judged is also likely to be behind the greater effect of the number of sites observed in the frequentist models (Figs 1, 2, 3). However, it should also be recalled that the Bayesian curves are bounded below by about 50%, given that an uninformative prior has 50% of its distribution below zero, meaning that the curves in Figure 5 are only able to vary over a smaller range.

For the models providing parameter estimates of the slope coefficient (Models 2, 3, 4), the clearest trend was the general reduction in error with the duration of monitoring, and the fact that the frequentist models gave lower error that the Bayesian model (Fig. 6, results presented for the 30 site scenario only). Bias did not show a clear pattern across models, and was generally low (Fig. 7, results presented for the 30 site scenario only). The precision of the slope coefficient estimates increased in accuracy with longer periods of monitoring (Fig. 7). The 95% confidence intervals of the slope coefficient estimates were smaller for Model 3 than Model 2 (Figs 7a, b). The Bayesian 95% credible intervals were wider than the frequentist confidence intervals for any particular scenario; the Bayesian credible intervals also indicated increasing precision with longer periods of monitoring, as would be expected (Fig. 7c).

Discussion

Analytical approaches to plant cover data often recommend the use of proportional odds models (Guisan & Harrell 2000; Irvine & Rodhouse 2010), although the practice of using estimated percentage cover equivalents for cover intervals is also frequent, both for community and single species analyses (van der Maarel 1979, 2007; Currall 1987; van der Maarel & Franklin 2012). A clear advantage of analysing interval-based cover data using explicit values on a true metric scale is the ability to create more intuitive measures of trends in plant cover, however, the choice of model may have other important consequences. We investigated whether the choice of model for plant cover data resulted in significant changes to prospective power analyses, and assessed their impacts on error, bias and precision. In particular, we sought to quantify how the use of a linear modelling framework utilising interval-censoring, with or without the inclusion of random effects, could affect modelling outcomes.

Unsurprisingly, the probability of detecting a true underlying trend increased as a function of sample size, trend magnitude and time; these are standard results from power analyses (e.g. Miller & Mitchell 2014; Johnson *et al.* 2015). Our results also showed that initial starting cover is an important determinant of power for a response variable on a proportional scale, irrespective of the model used. This was also demonstrated by Irvine & Rodhouse (2010) for plant cover data analysed using the proportional odds model. Whilst the results presented here have been used to inform the development of a new abundance-based plant monitoring scheme for the UK (Walker *et al.* 2015; Pescott *et al.* 2015), increasing the flexibility of modelling approaches for plant cover data could have important consequence for ecological synthesis in

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

general, because much historical information on plant communities has been collected using interval-based cover scales (e.g. Dengler *et al.* 2011). The availability of flexible frameworks under which to combine data collected according to different cover scales may outweigh the costs of occasionally approximating metric boundaries for certain interval categories (van der Maarel 2007). Additionally, the fact that hierarchical variance structures can be more easily estimated within a linear modelling framework means that data varying across space and time can be more accurately modelled (Johnson *et al.* 2015). Our interval-censored models also explicitly acknowledge and utilise the uncertainty underlying actual estimates of plant cover collected using standard cover scales, potentially mitigating the objections of some workers who have pointed to the uncertainty created by approximating metric equivalents of ordinal scale classes by explicitly accounting for it (Podani 2006). Indeed, lower cover classes that are defined in terms of frequency and cover (e.g. the first two classes of the Domin scale referred to above; Rodwell 1991), could be grouped under our scheme, ensuring that only easily defined percentage cover groupings are used for modelling, reducing the need for workers to decide upon metric equivalents for all classes. The option to include presence/absence data alongside traditional interval-censored cover data is another advantage of our method.

Our results suggest that the use of estimated interval midpoints in place of actual observations in a GLMM framework offers no clear advantages over the ordinal proportional odds model in terms of power; results from the maximum likelihood estimation based interval-censored model (Model 3) suggested more general increases in power. The power improvements seen with Model 3 were also mirrored in its slightly lower levels of error and increased precision. Whilst at first sight these facts might seem to favour Model 3 over the alternative modelling approaches, we should recall that the neglect of the hierarchical variance structure by Model 3 is likely to lead to smaller standard errors, and so higher power, a result also obtained by Johnson *et al.* (2015) in the context of binomial and Poisson GLMMs. Neglecting the hierarchical variance structure is also likely to be responsible for the lower average error, given that Model 3 is not attempting to estimate slope coefficients that vary by site. However, marginal increases in error could also be seen as an inevitable corollary of the hierarchical models explored here (Models 2 and 4), in that the explicit modelling of random effects attempts to account for underlying variation across monitored sites, rather than assuming that it does not exist. Whilst this has led to slight increases in error and reduced precision in the current simulation exercise, in the real world, where trends may be the outcome of numerous environmental drivers, parameter estimates and measures of variance are more likely to be of

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

use when presented with their attendant uncertainties, particularly if this results in more realistic estimates of power; low levels of error are likely to have less significance for ecological inference. The fact that our Bayesian interval-censored model (Model 4) shows slightly higher average error, and reduced precision, compared to the GLMM using interval mid-points (Model 2), may be a result of the fact that the Bayesian model accounts for the increased uncertainty of plant cover observations within categories, an additional complexity which better reflects the limited state of our knowledge concerning the variable being modelled. In addition, conservationists using our Bayesian model would be able to quantify the strength of their belief in a particular trend, rather than requiring a significant result before a trend was accepted. In this case, a small amount of error attached to a particular point estimate of the median value of a parameter and reduced precision may be less important than the fact that a strength of belief, and associated uncertainty, can be attached to a range of trend estimates (Wade 2000; Brooks *et al.* 2008).

We suggest that increased utilisation of the knowledge of interval boundaries associated with the popular 'cover-abundance' scales used in vegetation science is likely to be worthwhile. Whilst this may create challenges for scales where the lower categories are not explicitly defined in terms of percentage cover (Peet & Roberts 2013), it seems likely that standard approximations may often be able to be agreed (Mueller-Dombois & Ellenberg 1974; Currall 1987; van der Maarel 2007). Indeed, the availability of percentage cover approximations for all cover scales would greatly facilitate opportunities for combining datasets using our methods. The availability of Bayesian options also provides a challenge to those formulating prospective power analyses. The results will be less clear cut, in that workers will have to think harder about what constitutes a desirable level of posterior belief in a trend parameter being beyond a certain size (Morrison 2007), but this is little different to advice offered to modellers using classical frameworks concerning the choice of effect size to investigate (Seavy & Reynolds 2007; Johnson *et al.* 2015). Indeed, we consider that the flexibility of the Bayesian approach is likely to be an advantage in this respect, because 'out-of-the-box' levels for power (e.g. 80%) cannot be chosen without thought (Di Stefano 2003; Morrison 2007). The approach may also serve to further highlight the fact that all power analyses rest on strong assumptions (Morrison 2007; Johnson *et al.* 2015), and to encourage users to highlight uncertainty as well as summary measures of power (Olsen *et al.* 1997).

We suggest that the ability of workers to quantify beliefs in trends of different magnitudes is likely to be an advantage when communicating results or planning conservation actions

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

(Brooks *et al.* 2008). Here we have chosen the simplest Bayesian summary metric for illustration, but these can be tailored to the categories of most relevance to the species or system under study. Finally, we note that for monitoring programmes wishing to monitor multiple species, more sophisticated approaches to estimating power (or a Bayesian alternative) may be required; for example, any particular set of plots or larger sites will lead to variable occupancies amongst the target species of interest (Manley *et al.* 2004).

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

Acknowledgements

We would like to thank Pete Henrys for providing useful comments on the manuscript. This work was funded as part of the National Plant Monitoring Scheme contract NEC05294 funded by the Joint Nature Conservation Committee (Peterborough, UK) and the Centre for Ecology & Hydrology (national capability funding through NERC).

References

Agresti, A. (2002). *Categorical Data Analysis*, 2nd edn. Wiley-Blackwell, New York.

Bates, D., Mächler, M., Bolker, B. & Walker, S. (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, **67**, 1–48.

Bolker, B.M. (2008). *Ecological Models and Data in R*. Princeton University Press, USA.

Brooks, S.P., Freeman, S.N., Greenwood, J.J.D., King, R. & Mazzetta, C. (2008). Quantifying conservation concern–Bayesian statistics, birds and the red lists. *Biological Conservation*, **141**, 1436–1441.

Brooks, S.P. & Gelman, A. (1998). General Methods for Monitoring Convergence of Iterative Simulations. *Journal of Computational and Graphical Statistics*, **7**, 434–455.

Carey, P., Wallis, S., Chamberlain, P.M., Cooper, A., Emmett, B.A., Maskell, L.C., McCann, T., Murphy, J., Norton, L.R., Reynolds, B., Scott, A., Simpson, I.C., Smart, S.M. & Ullyett, J. (2008). Countryside Survey: UK Results from 2007. URL http://www.countrysidesurvey.org.uk/outputs/uk-results-2007 [accessed 28 April 2015]

Collen, B., Pettorelli, N., Baillie, J.E.M. & Durant, S.M. (Eds.). (2013). *Biodiversity Monitoring and Conservation: Bridging the Gap between Global Commitment and Local Action*. John Wiley & Sons, Cambridge, UK.

Currall, J.E.P. (1987). A transformation of the Domin scale. *Vegetatio*, **72**, 81–87.

Damgaard, C. (2014). Estimating mean plant cover from different types of cover data: a coherent statistical framework. *Ecosphere*, **5**, art20.

Damgaard, C. (2009). On the distribution of plant abundance data. *Ecological Informatics*, **4**, 76–82.

Dengler, J., Jansen, F., Glöckler, F., Peet, R.K., De Cáceres, M., Chytrý, M., Ewald, J., Oldeland, J., Lopez-Gonzalez, G. & Finckh, M. (2011). The Global Index of Vegetation-Plot Databases (GIVD): a new resource for vegetation science. *Journal of Vegetation Science*, **22**, 582–597.

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.

Dornelas, M., Magurran, A.E., Buckland, S.T., Chao, A., Chazdon, R.L., Colwell, R.K., Curtis, T., Gaston, K.J., Gotelli, N.J., Kosnik, M.A., McGill, B., McCune, J.L., Morlon, H., Mumby, P.J., Øvreås, L., Studeny, A. & Vellend, M. (2013). Quantifying temporal change in biodiversity: challenges and opportunities. *Proceedings of the Royal Society B: Biological Sciences*, **280**, 20121931.

Gelman, A. & Hill, J. (2007). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press, New York, NY.

Gitzen, R.A., Millspaugh, J.J., Cooper, A.B. & Licht, D.S. (Eds.). (2012). *Design and Analysis of Long-term Ecological Monitoring Studies*. Cambridge University Press, Cambridge, UK.

Guisan, A. & Harrell, F.E. (2000). Ordinal response regression models in ecology. *Journal of Vegetation Science*, **11**, 617–626.

Irvine, K.M. & Rodhouse, T.J. (2010). Power analysis for trend in ordinal cover classes: implications for long-term vegetation monitoring. *Journal of Vegetation Science*, **21**, 1152–1161.

Johnson, P.C.D., Barry, S.J.E., 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.

Jones, J.P.G. (2013). Monitoring in the real world. *Biodiversity Monitoring and Conservation: Bridging the Gap between Global Commitment and Local Action* (eds B. Collen, N. Pettorelli, J.E.M. Baillie & S.M. Durant). John Wiley & Sons, Chichester, West Sussex.

Kent, M. (2012). *Vegetation Description and Data Analysis: A Practical Approach*, 2nd edn. Wiley-Blackwell, Chichester, UK.

King, R., Brooks, S.P., Mazzetta, C., Freeman, S.N. & Morgan, B.J.T. (2008). Identifying and diagnosing population declines: a Bayesian assessment of lapwings in the UK. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, **57**, 609–632.

Lebuhn, G., Droege, S., Connor, E.F., Gemmill-Herren, B., Potts, S.G., Minckley, R.L., Griswold, T., Jean, R., Kula, E., Roubik, D.W., Cane, J., Wright, K.W., Frankie, G. & Parker, F. (2013). Detecting Insect Pollinator Declines on Regional and Global Scales. *Conservation Biology*, **27**, 113–120.

Lindenmayer, D. & Likens, G. (2010). *Effective Ecological Monitoring*. Csiro Publishing, Collingwood VIC.

van der Maarel, E. (2007). Transformation of cover-abundance values for appropriate numerical treatment - Alternatives to the proposals by Podani. *Journal of Vegetation Science*, **18**, 767–770.

van der Maarel, E. (1979). Transformation of cover-abundance values in phytosociology and its effects on community similarity. *Vegetatio*, **39**, 97–114.

van der Maarel, E. & Franklin, J. (2012). *Vegetation Ecology*, 2nd edn. John Wiley & Sons, Chichester, UK.

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

Manley, P.N., Zielinski, W.J., Schlesinger, M.D. & Mori, S.R. (2004). Evaluation of a multiple-species approach to monitoring species at the ecoregional scale. *Ecological Applications*, **14**, 296–310.

Maskell, L.C., Norton, L.R., Smart, S.M., Scott, R., Carey, P., Murphy, J., Chamberlain, P.M., Wood, C.M., Barr, C.J. & Bunce, R.G.H. (2008). *Countryside Survey. Vegetation Plots Handbook*. NERC/Centre for Ecology & Hydrology, Lancaster.

Miller, K.M. & Mitchell, B.R. (2014). A new tool for power analysis of fixed plot data: Using simulations and mixed effects models to evaluate forest metrics. *Ecosphere*, **5**, art110.

Morrison, L.W. (2007). Assessing the reliability of ecological monitoring data: power analysis and alternative approaches. *Natural Areas Journal*, **27**, 83–91.

Mueller-Dombois, D. & Ellenberg, H. (1974). *Aims and Methods of Vegetation Ecology*. John Wiley & Sons, New York.

Olsen, T., Hayden, B.P., Ellison, A.M., Oehlert, G.W. & Esterby, S.R. (1997). Ecological resource monitoring: change and trend detection workshop. *Bulletin of the Ecological Society of America*, **78**, 11–13.

Peet, R.K. & Roberts, D.W. (2013). Classification of Natural and Semi-natural Vegetation. *Vegetation Ecology* (eds J. Franklin & E. van der Maarel), pp. 28–70. Wiley-Blackwell, New York.

Pescott, O.L., Walker, K.J., Pocock, M.J.O., Jitlal, M., Outhwaite, C.L., Cheffings, C.M., Harris, F. & Roy, D.B. (2015). Ecological monitoring with citizen science: the design and implementation of schemes for recording plants in Britain and Ireland. *Biological Journal of the Linnean Society*, **115**, 505–521.

Plummer, M. (2013). JAGS Version 3.4.0 User Manual. http://sourceforge.net/projects/mcmc-jags/files/Manuals/3.x/jags_ user_manual.pdf.

Pocock, M.J.O. & Evans, D.M. (2014). The Success of the Horse-Chestnut Leaf-Miner, *Cameraria ohridella*, in the UK Revealed with Hypothesis-Led Citizen Science. *PLoS ONE*, **9**, e86226.

Podani, J. (2006). Braun-Blanquet's legacy and data analysis in vegetation science. *Journal of Vegetation Science*, **17**, 113–117.

Rodwell, J.S. (Ed.). (1991). *British Plant Communities Volume 1. Woodlands and scrub*. Cambridge University Press, Cambridge, UK.

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

Urquhart, N.S. (2012). The role of monitoring design in detecting trend in long-term ecological monitoring studies. *Design and Analysis of Long-term Ecological Monitoring Studies* (eds R.A. Gitzen, J.J. Millspaugh, A.B. Cooper & D.S. Licht), pp. 151–173. Cambridge University Press, Cambridge, UK.

Venables, W.N. & Ripley, B.D. (2002). *Modern Applied Statistics with S*, 4th edn. Springer, USA.

Wade, P.R. (2000). Bayesian methods in conservation biology. *Conservation Biology*, **14**, 1308–1316.

Walker, K., Dines, T., Hutchinson, N. & Freeman, S. (2010). *Designing a new plant surveillance scheme for the UK*. JNCC, Peterborough.

Walker, K.J., Pescott, O.L., Harris, F., Cheffings, C., New, H., Bunch, N. & Roy, D.B. (2015). Making plants count. *British Wildlife*, **26**, 243–250.

Warton, D.I. & Hui, F.K.C. (2011). The arcsine is asinine: The analysis of proportions in ecology. *Ecology*, **92**, 3–10.

Yu-Sung Su & Yajima, M. (2015). R2jags: Using R to Run 'JAGS'. R package version 0.5-7. http://CRAN.R-project.org/package=R2jags.

Table 1. Domin classes and their equivalent cover/frequency values, with the cover values used in analyses.

Domin class	Frequency/cover values (%)	Interpreted cover (%)
(0)	Absent	0.001 - 0.1
1	<1, 1-2 individuals	0.1-1
2	<1, several individuals	1-3
3	1-4	3-5
4	5-10	5-10
5	11-25	10-25
6	26-33	25-33
7	34-50	33-50
8	51-75	50-75
9	76-90	75-90
10	91-100	90-99

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

Table 2. Values of key parameters used in scenario simulations.

Study variable	Values investigated
Initial species' percentage cover (average)	5%, 10%, 20%, 40%
Number of sites monitored per year	15, 30, 50
Proportional declines in odds over 10 years	30%, 50%, 75%, 90%
Slope (trend) variance	0.15
Intercept (starting abundance) variance	0.05
Residual variance	0.80

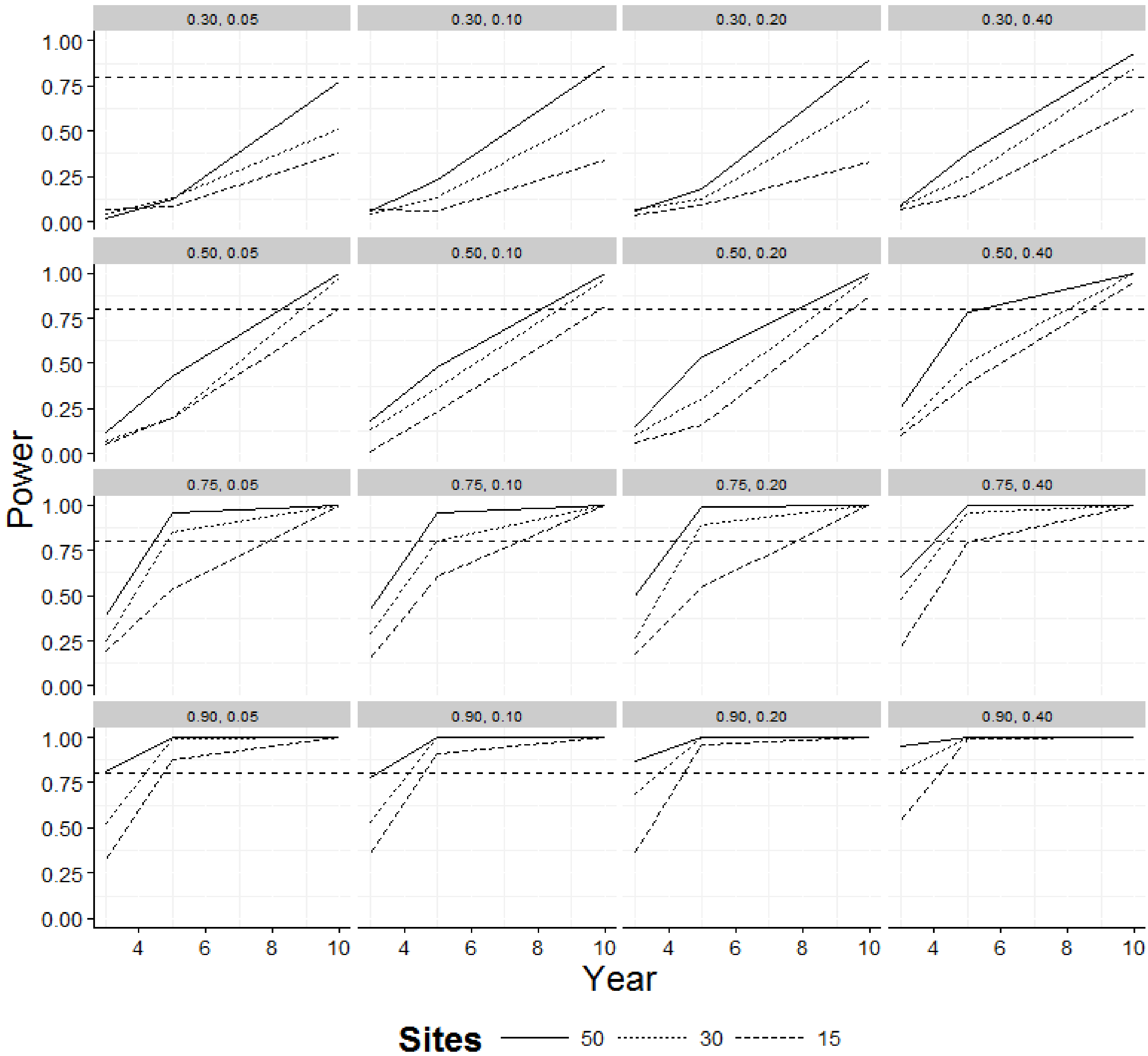

Figure 1. Power for a proportional odds model (Model 1) to detect decreases with α = 0.05 across 100 simulations. Rows represent different proportional declines undergone over a 10 year period (the first number given in the individual graph headers). Columns (the second number given in the individual graph headers) represent different initial starting proportional covers. The dashed horizontal line indicates the conventionally desirable level of 80% power.

PeerJ Preprints | https://doi.org/10.7287/peerj.preprints.2532v1 | CC BY 4.0 Open Access | rec: 16 Oct 2016, publ: 16 Oct 2016

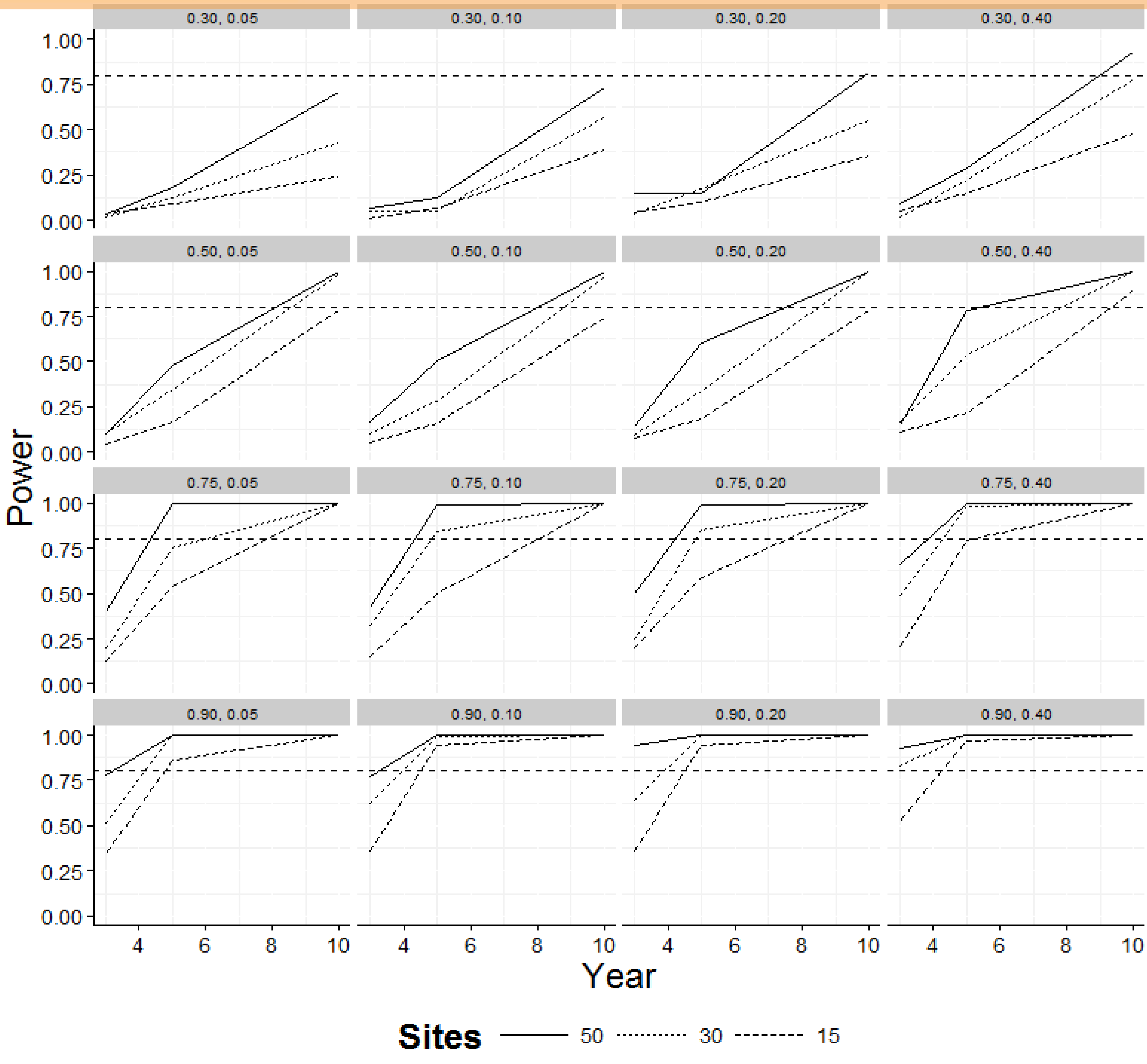

Figure 2. Power for a GLMM using interval class midpoints (Model 2) to detect decreases with α = 0.05 across 100 simulations. Rows represent different proportional declines undergone over a 10 year period (the first number given in the individual graph headers). Columns (the second number given in the individual graph headers) represent different initial starting proportional covers. The dashed horizontal line indicates the conventionally desirable level of 80% power.

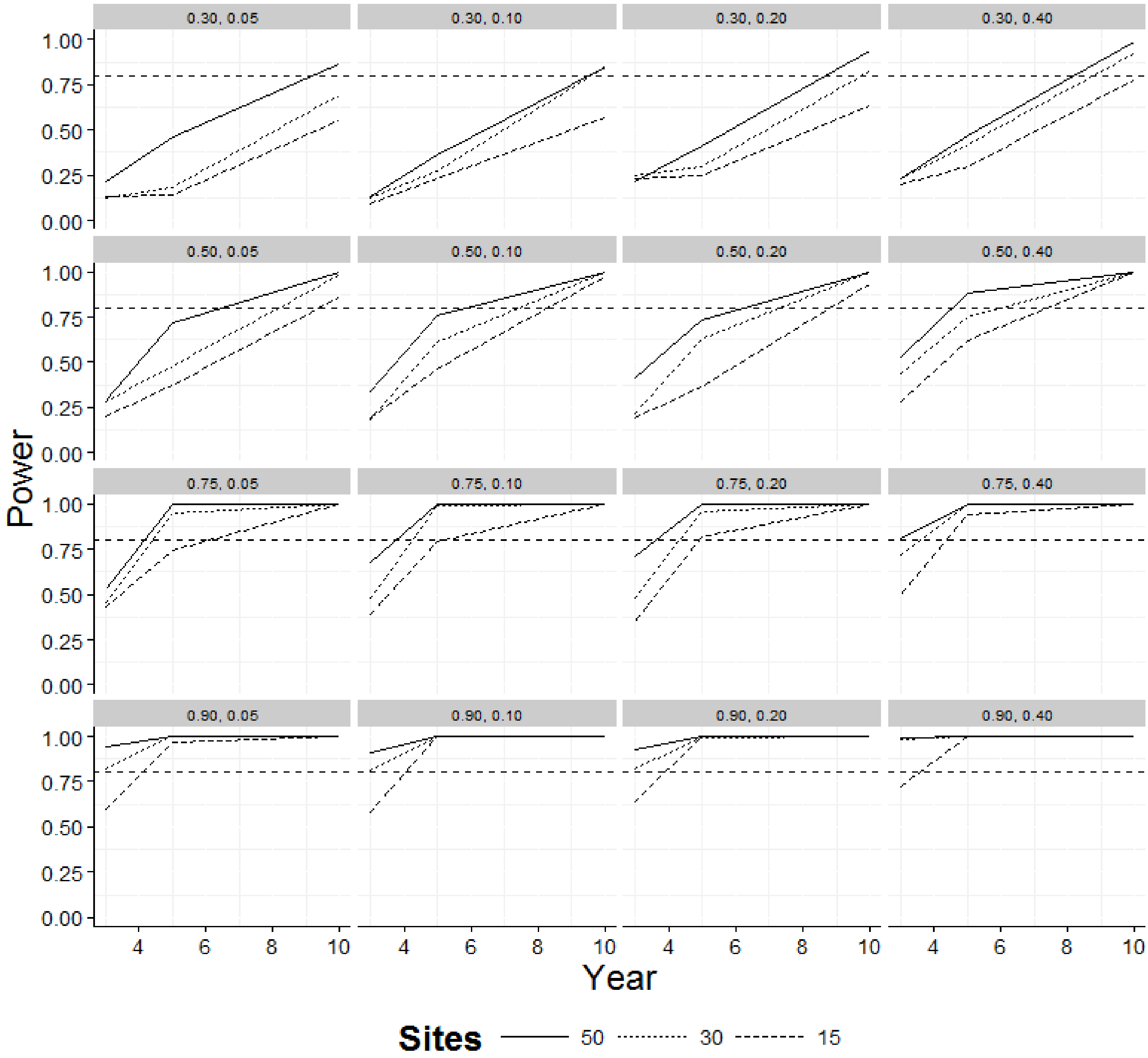

Figure 3. Power for an interval-censored linear model (Model 3) to detect decreases with α = 0.05 across 100 simulations. Rows represent different proportional declines undergone over a 10 year period (the first number given in the individual graph headers). Columns (the second number given in the individual graph headers) represent different initial starting proportional covers. The dashed horizontal line indicates the conventionally desirable level of 80% power.

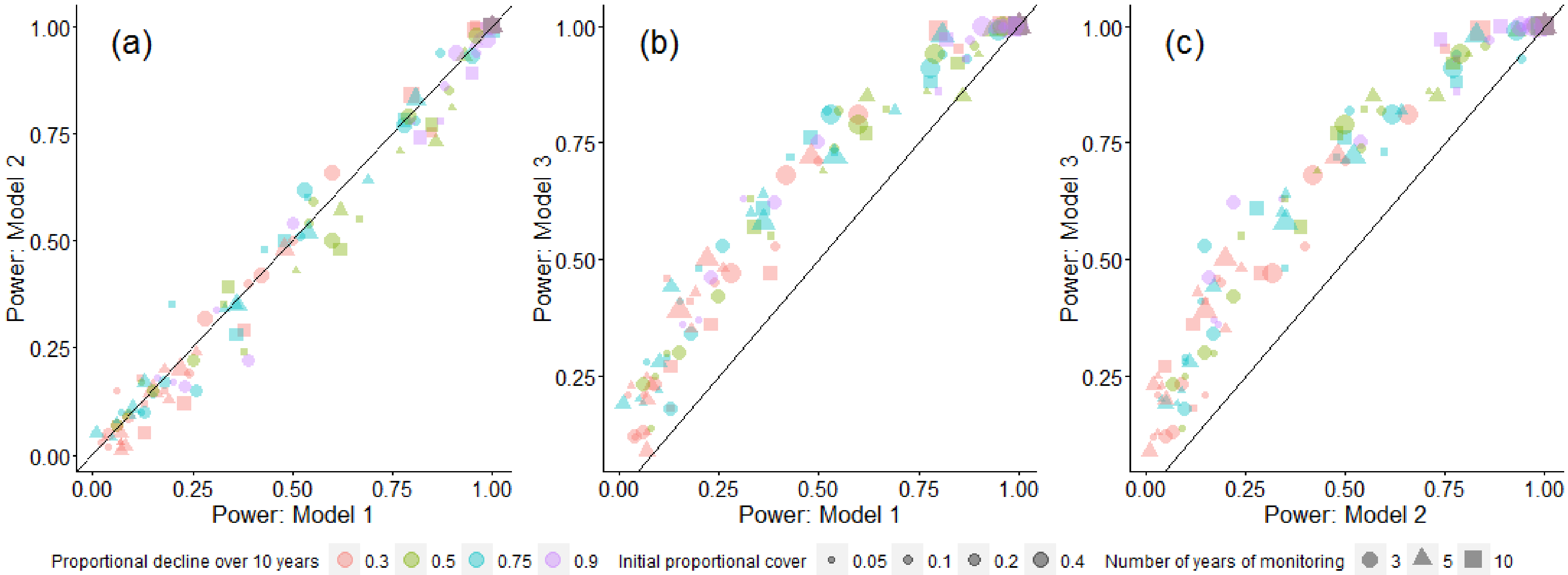

Figure 4. Comparisons of power between frequentist models. (a) Model 1 (proportional odds model) versus Model 2 (GLMM using interval class midpoints); (b) Model 1 (proportional odds model) versus Model 3 (interval-censored linear model); (c) Model 2 (GLMM using interval class midpoints) versus Model 3 (interval-censored linear model).

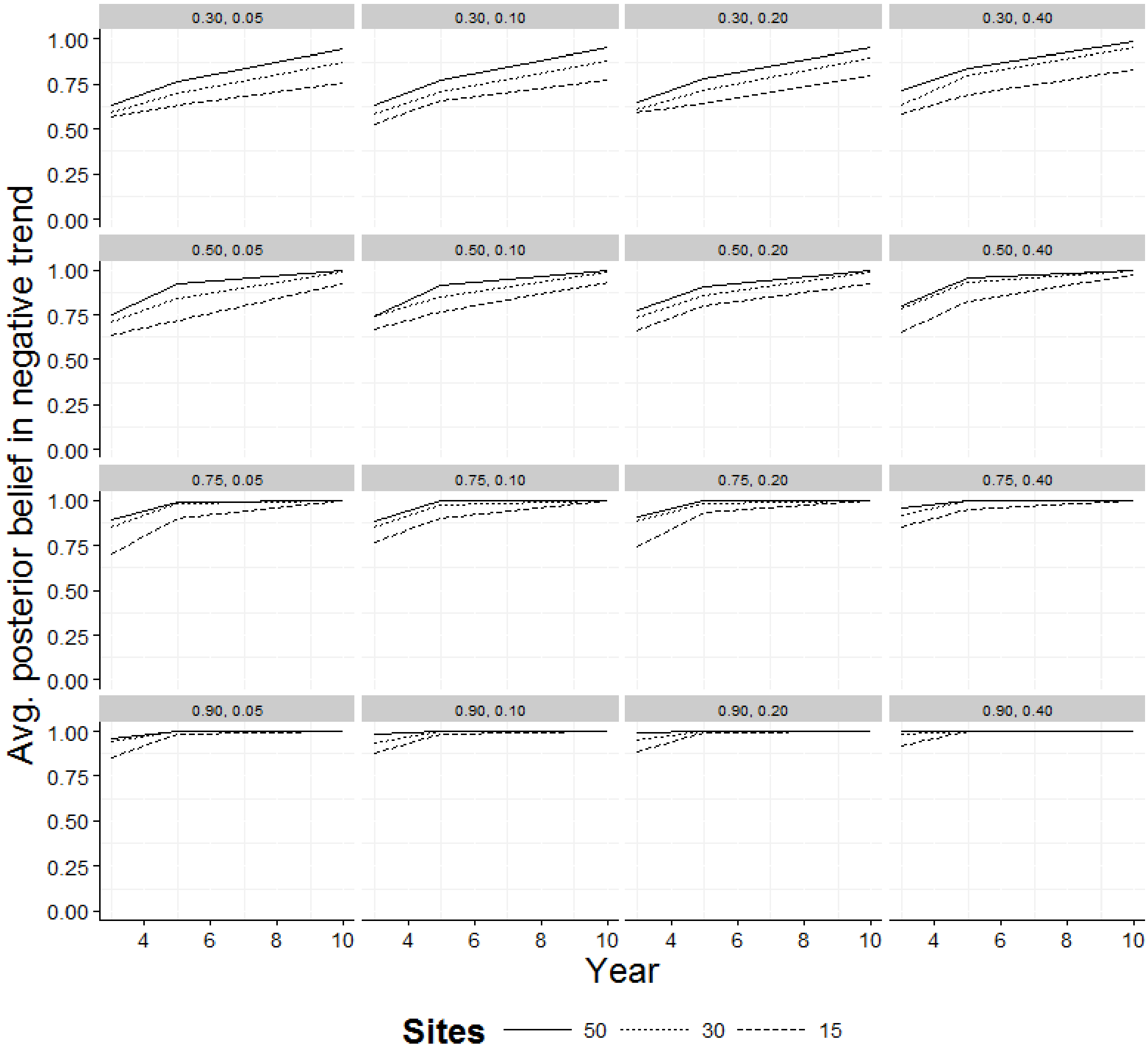

Figure 5. The average belief (proportion of the posterior distribution for the slope coefficient below zero) in a negative trend across 100 simulations from an interval-censored hierarchical Bayesian regression. Rows represent different proportional declines undergone over a 10 year period (the first number given in the individual graph headers). Columns (the second number given in the individual graph headers) represent different initial starting proportional covers.

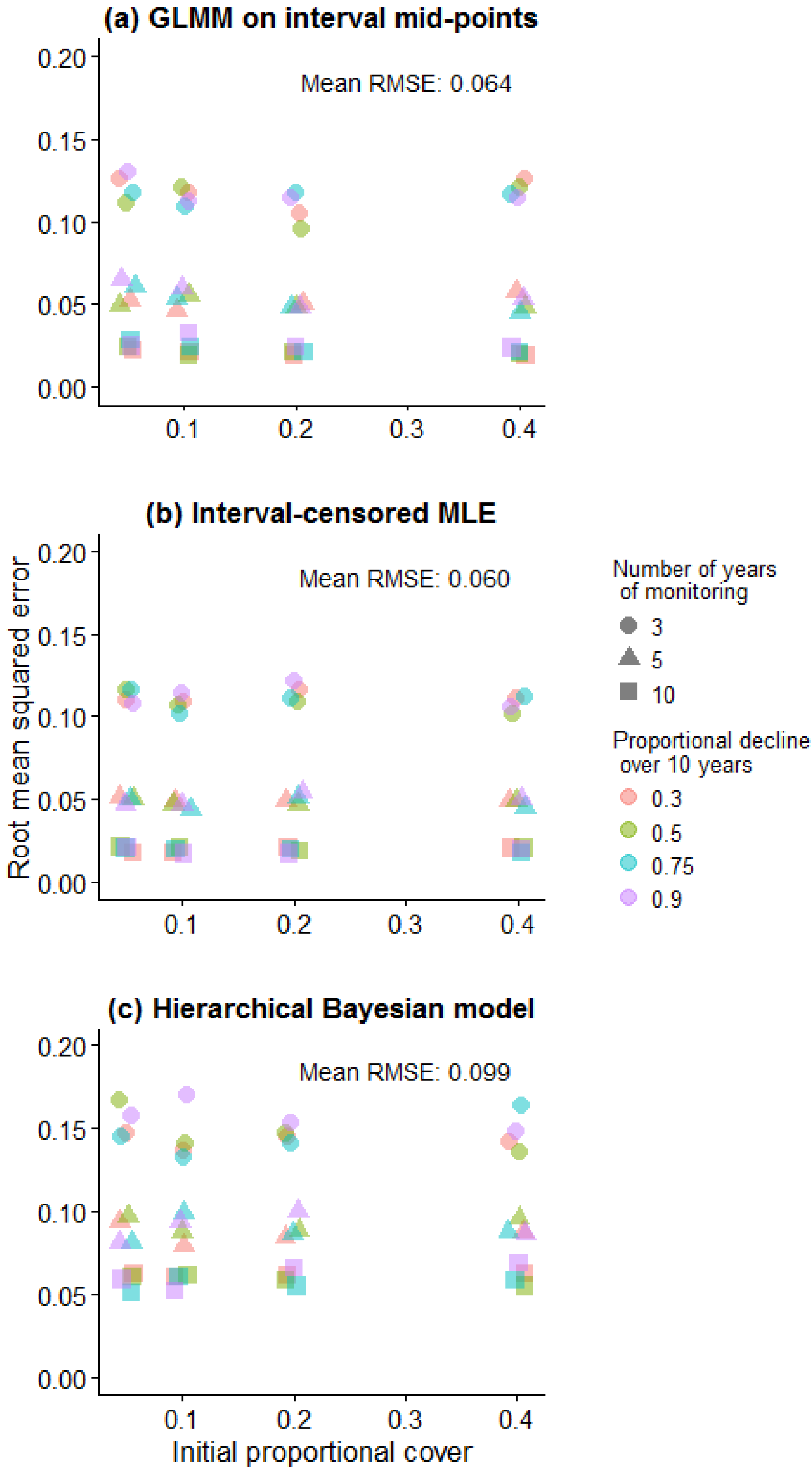

Figure 6. Root mean squared errors for the slope coefficient estimates across all scenarios with 30 monitored sites. (a) GLMM using class interval midpoints (Model 2); (b) interval-censored frequentist model (Model 3); (c) Hierarchical Bayesian model (Model 4). Root mean squared errors in (c) were calculated using the medians of the posterior distributions of the slope coefficient. Points for each level of the proportional initial abundance are jittered to increase their visibility.

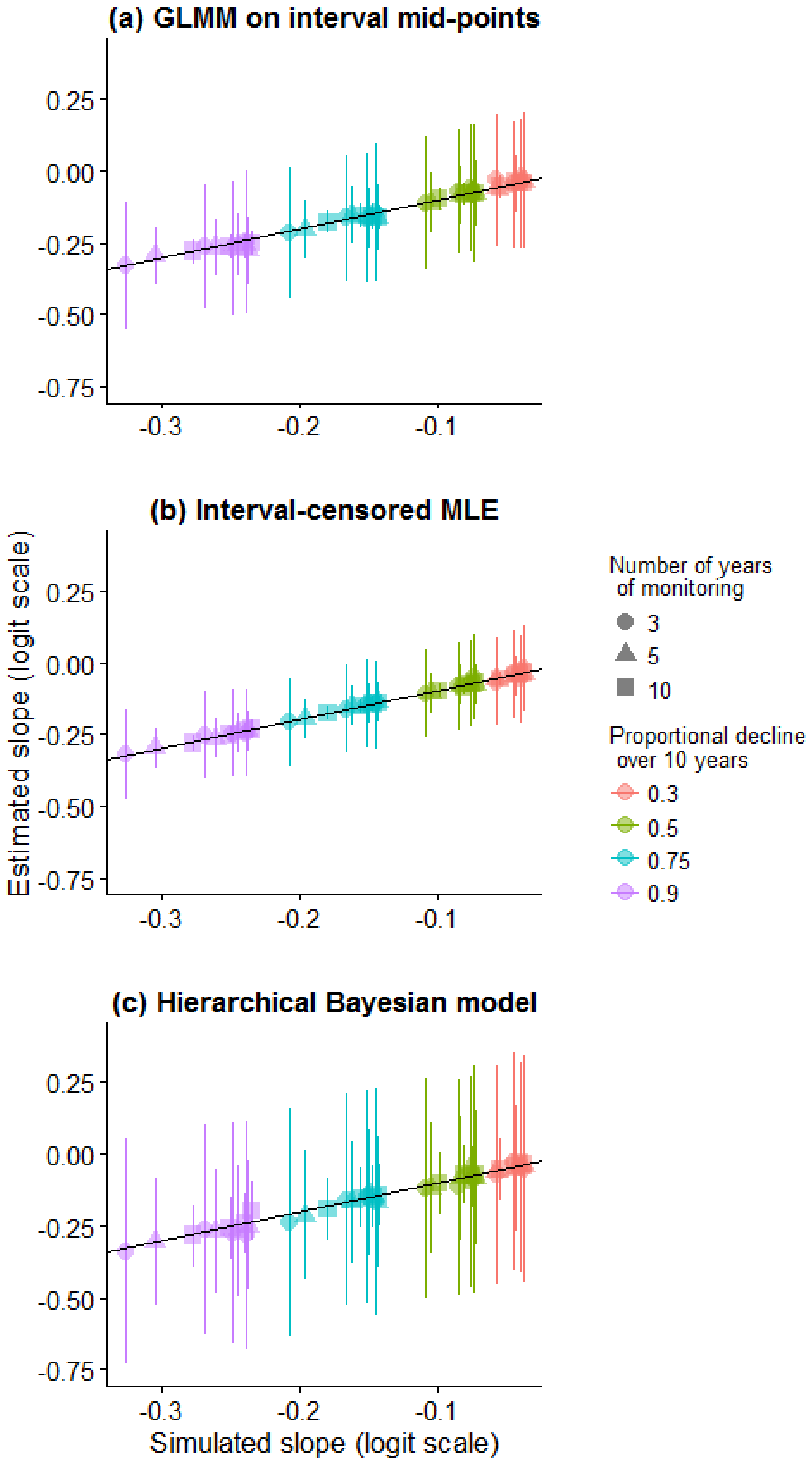

Figure 7. Estimated slope coefficient estimates versus simulated for the 30 site scenario for: (a) GLMM using class interval midpoints (Model 2); (b) interval-censored frequentist model (Model 3); (c) Hierarchical Bayesian model (Model 4). Point estimates in (c) are the medians of the posterior distributions of the slope coefficients. Error bars are 95% confidence intervals (Wald standard error-based naïve estimates in (a)), except in (c) where they are 95% credible intervals. The solid line is the line of equality between simulated and estimated values