Received: 19 December 2019 | Accepted: 6 May 2020

DOI: 10.1111/2041-210X.13407

COMMENTARY

Methods in Ecology and Evolution | BRITISH ECOLOGICAL SOCIETY

The use of Bayesian priors in Ecology: The good, the bad and the not great

Katharine M. Banner[1] | Kathryn M. Irvine[2] | Thomas J. Rodhouse[3]

[1]Department of Mathematical Sciences, Montana State University, Bozeman, MT, USA

[2]U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, USA

[3]National Park Service, Oregon State University-Cascades, Bend, OR, USA

Correspondence
Katharine M. Banner
Email: katharine.banner@montana.edu

Funding information
U.S. Geological Survey; Montana State University

Handling Editor: Robert B. O'Hara

Abstract

1. Bayesian data analysis (BDA) is a powerful tool for making inference from ecological data, but its full potential has yet to be realized. Despite a generally positive trajectory in research surrounding model development and assessment, far too little attention has been given to prior specification.
2. Default priors, a sub-class of non-informative prior distributions that are often chosen without critical thought or evaluation, are commonly used in practice. We believe the fear of being too 'subjective' has prevented many researchers from using *any* prior information in their analyses despite the fact that defending prior choice (informative or not) promotes good statistical practice.
3. In this commentary, we provide an overview of how BDA is currently being used in a random sample of articles, discuss implications for inference if current bad practices continue, and highlight sub-fields where knowledge about the system has improved inference and promoted good statistical practices through the careful and justified use of informative priors.
4. We hope to inspire a renewed discussion about the use of Bayesian priors in Ecology with particular attention paid to specification and justification. We also emphasize that *all* priors are the result of a subjective choice, and should be discussed in that way.

KEYWORDS

Bayesian hierarchical models, good statistical practice, sensitivity analysis, subjective priors

1 | INTRODUCTION

Bayesian data analysis (BDA) is now broadly acknowledged as an invaluable tool for modelling ecological data because of its capability to easily account for hierarchical structure, as well as observation and process uncertainties inherent in ecological systems (e.g. Cressie, Calder, Clark, Hoef, & Wikle, 2009; Dorazio, 2016; Ellison, 2004; Hobbs & Hooten, 2015; Kéry & Royle, 2016). Although popular, acceptance of BDA within ecology (and elsewhere) has, in large part, been predicated on the ability to remain 'objective' through the use of 'non-informative' priors. Controversy surrounding the use of 'subjective' (or, more appropriately named, informative) priors (e.g. Lele & Dennis, 2009) has led to a nearly universal adoption of using common 'non-informative' priors without providing justification for this choice or performing a sensitivity analysis (i.e. comparing posteriors created from different priors).

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13407

TABLE 1 Types of prior distributions and our proposed definitions (Gelman et al., 2014). For brevity, we use 'prior' to refer to a 'prior distribution'

Types of priors	Proposed definition
default	Commonly used non-informative priors that are often left unjustified by the user. Examples include, normal priors for regression coefficients with variances as large as 1×10^6, Uniform(0,1) on probabilities or proportions, and other 'non-informative' priors used without justification in software tutorials (e.g. WinBugs manual)
vague, flat, diffuse	A non-informative prior that is used to reflect the prior knowledge that not much is known about the parameter of interest, but is well justified and hyper-parameter values are set to reflect a reasonable range of values for the parameter in the context of the problem.
Jeffreys'	A prior for a single-parameter that, when the parameter is transformed to a different scale (via a 1:1 transformation), the resulting prior for the transformed parameter is exactly the same as the prior for the parameter on the original scale. This approach was introduced by Jeffreys' (Jeffreys, 1946), and is often used to define a non-informative prior for a single-parameter that is invariant to transformations, or *scale-invariant*
weakly informative	Often refers to prior distributions that are used to reflect a diluted (or scaled back) amount of knowledge about the parameters
regularizing	A type of weakly informative prior that is meant to constrain the parameter space to help with estimation of the posterior distribution. Examples include N(0, σ^2 = 2) priors on logistic regression coefficients, and shrinkage priors when the number of predictors is greater than the sample size (i.e. $p > n$ problems)
informative	A prior that is carefully designed to reflect the current knowledge (and uncertainty) of the parameter. By nature, this must be well justified and communicated to the reader
conjugate	A prior distribution that has the same distributional form as the posterior distribution. A common example, is the Beta-Binomial, a beta prior on the parameter for probability of 'success' combined with a Binomial likelihood leads to a beta posterior distribution
proper versus *improper*	Priors for continuous parameters can be *proper* or *improper*. A proper prior does not depend on data and integrates to 1 over the support of the parameter. An improper prior has a non-finite integral over the support of values the parameter can take on

We refer to this blind acceptance of certain priors as choosing *default* priors, and we believe this choice not only promotes poor statistical practice (see Gelman & Hennig, 2017), but has left the full potential of the BDA framework unrealized in ecological applications (e.g. Morris, Vesk, McCarthy, Bunyavejchewin, & Baker, 2015). The practice of choosing default priors is different from choosing non-informative or weakly informative priors by carefully considering the context of the problem and explaining the decision to use such priors. For clarity, we provide definitions for the types of priors we consider in Table 1.

In this commentary, we provide an overview of how BDA is currently being used, highlight sub-fields where knowledge about the system has improved inference through the careful and justified use of informative priors, and discuss implications for inference if current bad practices continue. We hope to inspire a renewed discussion about the use of BDA in Ecology with particular attention paid to the decision of which priors to use in an analysis.

2 | BDA IN ECOLOGY

Despite early criticisms of BDA by ecologists (e.g. Dennis, 1996), most practitioners agree that specification of a reasonable model that approximates the true data-generating process is central to any statistical analysis. Then, the distinguishing characteristic of BDA from frequentist approaches is the treatment of parameters as random variables, which requires careful specification of probability distributions to reflect a priori knowledge about parameter uncertainty (i.e. choosing the priors). In BDA, prior distributions are combined with observed data through the likelihood to reflect updated a posteriori knowledge in the form of posterior probability distributions for all model parameters (i.e. the posteriors). These three components are illustrated for three different priors in a simple Bayesian model in Figure 1, and a full description of the model is provided in Box 1. There are multiple textbooks providing comprehensive treatments of BDA specifically focused on ecological applications (e.g. Hobbs & Hooten, 2015; Kéry & Royle, 2016; Link & Barker, 2010; Royle & Dorazio, 2008), and although the potential advantages of using prior knowledge are discussed (even at length in some of the more recent texts), it seems the use of default priors is far more common in practice.

To begin to understand how priors are currently used in ecological research, we conducted a Web-of-Science search within a subset of Ecology journals. Rather than attempting to develop a sampling frame that encompassed all application papers, we focused our search on journals where we commonly see Bayesian methods applied to ecological data (see Supplement S1 for details about search criteria, our review process and citations). We focused on published articles with 'Bayesian model' or 'Bayesian hierarchical model' in the abstract or keywords. Our search resulted in 238 articles, from which we randomly selected 15 to review. Of the 15 articles selected, there were 12 that were application based. The majority of the applied papers (9 of 12) used default priors with 'standard/pre-set' hyper-parameter values for all parameters in their analyses; the remaining three papers used prior information to adjust

BOX 1 A Beta-Binomial example

Problem description: Consider an ecological random variable that takes on one of two outcomes (i.e. a success/failure outcome variable). For example, let X be the detection of a species on a single visit to a sample unit within a region of interest, where $x = 1$ if the species is detected and $x = 0$ if not; or the survival of a particular individual in a population from year 1 to year 2, where $x = 1$ if the individual survives and $x = 0$ if not. Then, let Y be the total number of successes from a collection of n trials of one of the X variables (e.g. the total number of detections from n sample units). If it is reasonable to model the n trials as independent with a constant probability of success, θ, then Y can be modeled with a Binomial(n, θ) distribution. There are three major components to a Bayesian analysis of these data, each are described in detail below; this type of model is commonly referred to as the *Beta-Binomial* model. Terms associated with definitions in Table 1 are italicized.

- **The prior distribution:** Knowledge about θ before data are collected is expressed through a probability distribution called the *prior*. The *conjugate* family of distributions for θ are the beta(α, β) distributions, which have support from 0 to 1 and are very flexible in shape (dashed lines in Figure 1); although any characterization of knowledge for θ that is positive on (0, 1) and integrates to 1 is a *proper* prior distribution for θ.
- **The model likelihood:** The likelihood function describes how 'likely' each value of θ is, given the data. So, once data are observed, the likelihood for θ can be plotted for the observation that y successes were seen in the sample on n trials (solid line in Figure 1).
- **The posterior distribution:** The result of updating the prior with the likelihood is the posterior distribution. In this case, it is a beta($\alpha^* = y + \alpha$, $\beta^* = n - y + \beta$) distribution, where the hyper-parameter values chosen for α and β in the prior distribution have the practical interpretation of being an expectation of the number of additional successes and failures, respectively.

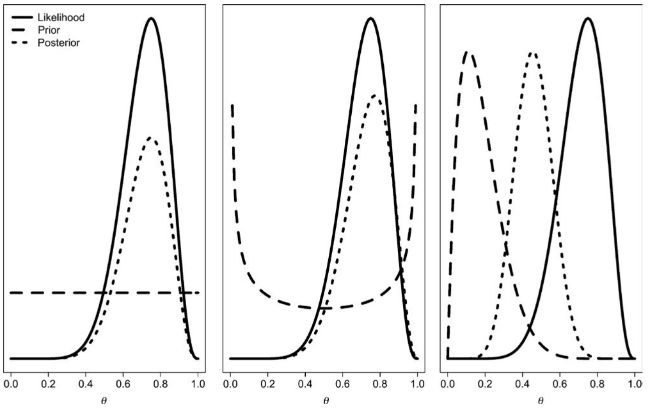

FIGURE 1 A graphical illustration of a Beta-Binomial model (see description in Box 1) with three different choices of priors. Left to right, non-informative uniform prior, scale-invariant Jeffreys' prior, and highly informative prior. *Note that the densities were scaled to aid in the visualization*

An illustration of how three different choices of a prior distribution affects the posterior distribution in a Beta-Binomial model is shown in Figure 1. For all three models, it is assumed that $y = 9$ 'successes' (e.g. detections or survivals) were observed in $n = 12$ trials. That is, the Binomial($n = 12$, θ) liklelihood is displayed for an observed value of $y = 9$. The three different priors considered, shown from left to right, are a *flat* uniform prior ($\alpha = \beta = 1$), a *Jeffreys*' prior that results in the same beta distribution ($\alpha = \beta = 0.5$, see Section 3) for any 1:1 function of θ, and an extremely *informative* prior that is inconsistent with the observed data ($\alpha = 2$, $\beta = 9$). The informative prior distribution effectively adds 11 trials to the 'experiment,' of which only two were 'successes', and thus has a strong influence on the posterior distribution as compared to the non-informative priors that add just one or two hypothetical observations to the data. The extent to which the posterior distribution is influenced by the specific choice of prior distribution is a function of its shape and the amount of information contained in the prior with respect to that contained in the data (e.g. the size of α and β compared to n in the Beta-Binomial model).

hyper-parameter values for at least one of their model parameters, along with vague or reasonably justified non-informative priors for others.

In two of the papers using informative priors, authors provided clear justifications for their choice of hyper-parameters; and in one paper, a discussion of the sensitivity analysis performed was provided in Supplement S1. In the nine articles using default or non-informative priors, authors provided varying levels of justification for their choices—few provided clear and thorough explanations (two papers), some provided no explanation at all (three papers), and the remaining four papers fell somewhere in between. Only two of these nine papers contained evidence of conducting a form of sensitivity analysis. Although our sample of papers contained a wide set of model types, a common theme emerged; discussions surrounding prior specification were all-around insufficient, and noticeably the most deficient when *default* priors were used (Table 1).

2.1 | The good

If the inherent subjectivity surrounding choosing a prior distribution to carry out a Bayesian analysis is more broadly acknowledged, and if practitioners adopt the practice of justifying their use of *any* type of prior (just like what is done for other modelling decisions), then we expect the stigma surrounding the use of informative priors to lessen. The 'classic' informative prior distribution is formulated based on results from related empirical studies (Dupuis & Joachim, 2006; Ellison, 2004) or pilot data that were collected to inform sample sizes and study design (Morris, Vesk, & McCarthy, 2013). In long-term monitoring, if there is a gap in the time series related to limited resources or changes in agency priorities, the posterior distributions based on analysing the first period of data collection can be used as empirically informed prior distributions when analysing the second time period of data. For example, in Rodhouse et al. (2019), a dynamic occupancy model with spatially explicit predictors (e.g. elevation, per cent forest cover) was used to assess evidence of change in bat populations. The posterior mean and standard deviation of the partial regression coefficients from previously published results were used as the mean and standard deviation (hyper-parameter values) in the normal priors for the same predictors in the analysis of the second phase of monitoring. Alternatively, in the absence of empirical data, expert opinions can be elicited, where methods for elicitation are discussed in Martin, Kuhnert, Mengersen, and Possingham (2005) and Kuhnert, Martin, and Griths (2010), among others. Several ecological studies have shown that using prior information can increase estimator precision without compromising accuracy (Morris et al., 2015), ultimately leading to larger effective sample sizes and saved resources.

Several authors propose employing Bayesian decision analysis for natural resource management and conservation (Dorazio & Johnson, 2003; Wade, 2000; Williams & Hooten, 2016). Decision analysis hinges on identifying an appropriate loss function that marries the potential actions based on estimated ecological parameters (e.g. population size or annual trend) with the potential consequences under the various management actions (Wade, 2000; Williams & Hooten, 2016). However, most applications of adaptive management (e.g. Canessa et al., 2015; McDonald-Madden et al., 2010; Runge, Converse, & Lyons, 2011) fall short of a fully Bayesian approach because prior distributions were not used to reflect uncertainty about ecological parameters. Employing a fully Bayesian analysis though, would allow for sequential updating of the posterior distribution for the ecological parameters of interest encouraging 'learning while doing' as proposed for adaptive management (Nichols, 1991).

Many argue Bayesian hierarchical models are essential for ecology because they allow for complexity and explicit consideration of the fact that the field measurement (e.g. a species counts) is one layer removed from the ecological parameter of interest (e.g. population size). The most common examples are occupancy models that explicitly link the detection or observation process conditional on the partially observed, yet of interest, occupancy or occurrence state. Researchers have continued to advance occupancy and count models by allowing for false positives and multiple species (e.g. Chambert et al., 2018; Wright, Irvine, Almberg, & Litt, 2019). However, the increase in model complexity has required the use of regularizing priors or informative priors based on calibration datasets to assist with parameter identifiability issues. An interesting conundrum that, for now, is only solved by using informative priors is that the statisticians advancing the model options tend to be out in front of the practitioners collecting the data. The gap has created a time lag for implementation as the community of practice institutionalizes the need to change field protocols to gather additional information to fuel more complex, yet improved, statistical models.

An underutilized potential of informative priors is as a social science tool to build consensus among stakeholders regarding how a system may respond to a management action. For example, a study on how logging in the rain forest impacts birds and small mammals used priors elicited from differing perspectives on the effect size and then used data to update the prior beliefs (Crome, Thomas, & Moore, 1996). The comparison among posterior distributions under divergent priors was then influential for reaching consensus on a polarizing conservation issue. Advancements in Bayesian model checking (e.g. Broms, Hooten, & Fitzpatrick, 2016; Conn, Johnson, Williams, Melin, & Hooten, 2018) and visualization tools (e.g. Gabry, Simpson, Vehtari, Betancourt, & Gelman, 2019) that aim to help practitioners think both generatively (about consistency between data generated by the system and from the prior), and predictively (about consistency between data generated from the posterior predictive distribution and those from the system) have paved the way for the transparent use of priors in general, and of informative priors when they are appropriate.

2.2 | The bad

In the face of all of this potential, some bad practices still remain. As evidenced in our sample of application papers, it appears there is a misconception among some practitioners that *all* default priors remain non-informative regardless of how the parameter is transformed in the model/likelihood. Invariance with respect to transformations is often true for Jeffreys' priors (Jeffreys, 1946), but there are some cases (albeit somewhat unrealistic) where the Jeffreys' prior does not remain non-informative (i.e. Tuyl, Gerlach, & Mengersen, 2012, Beta-Binomial models when the observed data are all successes or all failures). As a simple example, we show how the Normal(0, σ^2) prior commonly placed on partial regression coefficients in logit-linear regression and occupancy models can be quite informative on the probability scale (Figure 2). The degree to which the 'U'-shaped prior influences the posterior distribution is problem-specific, but one thing is clear from this simple example: normal priors with very large variances on partial regression coefficients in logit-linear models likely do not achieve the shape intended by the practitioner on the probability scale (see Seaman III, Seaman Jr., & Stamey, 2012). This emphasizes the importance of conducting sensitivity analyses to assess implications for final inference. Northrup and Gerber (2018) highlight this unintended consequence for occupancy models and provide an R (R Core Team, 2018) Shiny App for running sensitivity analyses for occupancy models (https://briangerber.shinyapps.io/OccupancyPrior/).

Similarly, Link (2013) demonstrates how the improper uniform prior, which is used with binomial count (*N*) models, can result in biased abundance estimates, and suggests an alternative yet still non-informative prior. For both types of models, when sample sizes are small or the detection probability is low, bias in ecological parameters of interest is more pronounced; this is extremely unfortunate because these are the common scenarios where researchers are often particularly careful about letting the data 'speak for themselves'. When the scale of interest is a status (and/or trend) indicator, this ostensibly good practice can lead to misinformed management and/or conservation decisions about the species of concern.

Another place where default priors can have pathological properties is in the context of model selection or combination (hereafter multimodel inference, or MMI). In a MMI framework, inference is conditional on all models in a model set rather than on a single model,

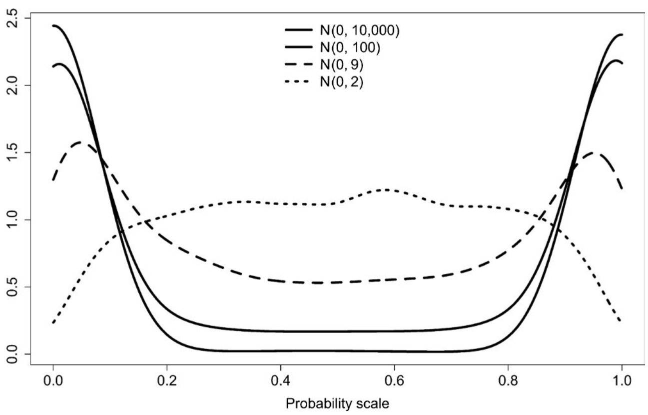

FIGURE 2 Five-thousand realizations of β_0 drawn from its prior distribution N(0, σ^2), where σ = 100, 10, 3 or 1.41, then back-transformed to the probability scale using the inverse-logit function, $\psi = \frac{e^{\beta_0}}{1+e^{\beta_0}}$, and plotted using the density function between 0 and 1. The likelihood assumed is a naïve constant probability model, $logit(\psi) = \beta_0$, but is illustrative of what occurs with more complex logit-linear models (see Northrup & Gerber, 2018; Seaman III et al., 2012)

which holds the appeal of addressing one type of model uncertainty but does not eliminate model uncertainty from a problem all together. To conduct Bayesian MMI, priors must be specified for the model set, and for all parameters in each model. The prior on the model set is categorical; each model is assigned a probability such that, together, the probabilities sum to one. The result is then a jointly estimated posterior for the model set, in the form of posterior model probabilities, and all parameters in each model. When improper or default priors with extremely large variance terms are used for partial regression coefficients in *all subsets* regression, the posterior model probabilities are largely influenced by Bayes factors favouring the simplest hypothesis even when the data say otherwise (see Kass & Raftery, 1995; Kass & Wasserman, 1995). This issue is discussed at length in the statistical literature (e.g. Consonni, Fouskakis, Liseo, & Ntzoufras, 2018; Kass & Raftery, 1995; Kass & Wasserman, 1995), and see Hooten and Hobbs (2015) for a comprehensive review of Bayesian model selection and regularization, and Link and Barker (2006) for a discussion of MMI, both in the context of ecological applications. Other aspects of MMI, specifically certain applications of model averaging, raise complications related to practical implications for inference (i.e. interpretations) and are beyond the scope of this paper, but are treated elsewhere (see Banner & Higgs, 2017; Cade, 2015).

Lastly, common to both the single-model and multimodel frameworks, default priors tend to have unexpected effects on the posterior distribution when variances of normal priors or ranges of uniform priors are set to unrealistically large values (e.g. Uniform(0, 1.0E6) on a standard deviation or Normal(0, σ^2 = 1.0E6) on a partial regression coefficient). We suspect this practice resulted from practitioners following early accessible examples of prior specification and then all subsequent applications followed suit without question, as did we in the beginning. For example, in the WinBUGS manual example for a normal hierarchical model implementation of a random effects growth curve model, the prior specification for the intercept, mean hyper-parameter was Normal(0, σ^{-2} = 1.0E-6). Further, in one of the first textbooks using this software, a multivariate normal prior on the regression coefficients for normal regression models was specified as Normal(0, σ^{-2} = 1.0E-4; Ntzoufras, 2009). Currently, statisticians are aggressively questioning these choices and developing ways to deal with the usual high dimensionality of regression problems through regularizing priors for regression coefficients (see Hooten & Hobbs, 2015, for ecology-specific examples) and tree shrinkage priors (Hefley, Zhang, Gray, & Bouska, 2020).

2.3 | The not great

The ugly truth is that oftentimes default priors with little to no thought about hyper-parameter values are used in place of thinking critically about what is known about the system under study and justifying why priors make sense for the problem at hand—including assessing how prior choice affects inference. It is easy to say, 'we used non-informative priors on *X*, *Y*, and *Z*,' but that is 'not great' statistical practice. Further, it is not much harder to justify reasonable hyper-parameter values that reflect current scientific understanding. For example, to define a prior on a latent state variable for body size, a prior that only places positive probability on body size intervals that are >0 should be used (e.g. a gamma, or a normal centred at a positive value with reasonable variance term), and a default N(0, 10E-6) does not make much sense. Known information about how the body sizes are distributed could help researchers choose a reasonably shaped population distribution for the body sizes (e.g. a gamma if the body sizes are expected to be right skewed and a normal if they are expected to be symmetric). Finally, additional information about the expected value, range, or variance of the body size distribution could help determine sensible hyper-parameter values. If nothing is known, a uniform(0, θ), where θ is chosen to be close to the largest possible value for the species could be sensible. As discussed in Section 2.2, the practice of blindly choosing default priors for common ecological models can ultimately result in poor conservation and management decisions for species of interest or concern. Fortunately, this is something that can change. Resources exist to make sensitivity analyses and model assessment standard components to *any* Bayesian analysis (e.g. Broms et al., 2016; Conn et al., 2018; Hobbs & Hooten, 2015; Link & Barker, 2010; Northrup & Gerber, 2018, among many others). The time is now to change the culture and make it inexcusable to take the easy way out.

3 | DISCUSSION

A strength of the Bayesian framework is that it is inherently subjective, requiring the *choice* and justification of a prior, regardless of the type of prior chosen (i.e. default, regularizing, weakly informative, informative). The choice to reflect 'no knowledge' using a non-informative prior is itself a subjective practice. We believe that 'subjective' Bayesian analysis should no longer be used in a pejorative sense or be a risky proposition. It is clear that the choice of 'no knowledge' using non-informative priors is at least as risky and often less transparent than thoughtful use of subjectively informed priors. There are cases where little prior knowledge exists regarding the parameter space for quantities of interest, but it is imperative that this is clearly stated and that care is taken to assess how the 'lack of information' in the prior affects the posteriors for ecological parameters of interest (which could be on a different scale than the prior). Further, rather than using default priors with pre-set hyper-parameter values (that can be ludicrous), it is important to think hard about the system and constrain prior distributions to give high density to values that make sense and low density to those that do not. Tools and guidelines exist to make this process more accessible to practitioners (e.g. Gabry et al., 2019; Hobbs & Hooten, 2015; Northrup & Gerber, 2018), and it is essential they start being used more consistently.

Potentially even more concerning is that there are many applications where the full potential of BDA is not being realized such as, long-term monitoring and adaptive management applications. The Bayesian approach is not just an 'easier way to fit hierarchical models', but a powerful framework providing the potential to incorporate the full scientific thought process into an analysis. In the context of monitoring, it is arguably never the case that *absolutely nothing* is known about model parameters with ecological importance before designing and conducting a full-scale study and even more-so as a program is refined over time. Thus, when default priors are used without careful consideration the opportunity to (a) hypothesize/predict, (b) study/experiment, (c) analyse data, (d) rinse-and-repeat applying what was learned, is squandered (Hobbs & Hooten, 2015; Morris et al., 2013, 2015). This can have grave consequences for monitoring rare, elusive, sensitive, or endangered species.

Additionally, although not our primary focus, a surprising result from our literature search was insufficient discussion of model assessment in almost all of the application papers in our sample. Because of this, we feel it is worth pointing out the growing body of references that provide guidelines for assessing common ecological models. For example, Broms et al. (2016) provides a prime example of how to conduct a Bayesian analysis start-to-finish for multi-species occupancy models, providing tools for within-sample and out-of-sample model assessment; Wright, Irvine, and Higgs (2019) provides diagnostic plots for residuals from occupancy models; and finally, a more general treatment of the subject can be found in Conn et al. (2018) as well as textbooks specific to BDA in Ecology (e.g. Hobbs & Hooten, 2015).

In conclusion, we echo the sentiments in Gelman and Hennig (2017) that the inherent subjectivity of BDA (whether priors are informative or not) leads to transparency in methods. Specifically, the prior specification step requires a thorough understanding and acknowledgment of the implicit assumptions the prior makes about the system under study, how that information affects the posterior distribution (i.e. sensitivity analysis), and transparent communication to potential consumers of the results. This practice requires practitioners to make tough decisions and think critically; it is something that already exists when informative priors are used (e.g. Morris et al., 2013, 2015; Rodhouse et al., 2019), and should be extended to *all* priors due to their inherent subjectivity.

ACKNOWLEDGEMENTS

Any use of trade, firm or product names is for descriptive purposes only and does not imply endorsement by the US Government. Funding for this research was provided by Cooperative Agreement No. 00CRAG0024 between U.S. Geological Survey and Montana State University through the Institute on Ecosystems. First, we thank the reviewers and editors for their thoughtful feedback, which improved the quality and clarity of this paper immensely. [Correction added on 17 July 2020, after first online publication: the acknowledgement section has been amended in this version.]

AUTHORS' CONTRIBUTIONS

K.M.B., K.M.I. and T.J.R. together, conceived the ideas for this manuscript; K.M.B. and K.M.I. led the writing of the manuscript with significant contributions from T.J.R. All authors contributed to the drafts and gave final approval for submission.

DATA AVAILABILITY STATEMENT

There were no data used in this manuscript. Search criteria for obtaining a sampling frame of papers in the ecological literature that employ Bayesian methods and our sampling strategy are described in Supplement S1.

ORCID

Katharine M. Banner https://orcid.org/0000-0002-7103-0042
Kathryn M. Irvine https://orcid.org/0000-0002-6426-940X
Thomas J. Rodhouse https://orcid.org/0000-0001-5953-9113

REFERENCES

Banner, K. M., & Higgs, M. D. (2017). Considerations for assessing model averaging of regression coefficients. *Ecological Applications*, *27*(1), 78–93. https://doi.org/10.1002/eap.1419

Broms, K. M., Hooten, M. B., & Fitzpatrick, R. M. (2016). Model selection and assessment for multi-species occupancy models. *Ecology*, *97*(7), 1759–1770. https://doi.org/10.1890/15-1471.1

Cade, B. S. (2015). Model averaging and muddled multimodel inference. *Ecology*, *96*(6), 2370–2382. https://doi.org/10.1890/14-1639.1

Canessa, S., Guillera-Arroita, G., Lahoz-Monfort, J. J., Southwell, D. M., Armstrong, D. P., Chadès, I., … Converse, S. J. (2015). When do we need more data? A primer on calculating the value of information for applied ecologists. *Methods in Ecology and Evolution*, *6*, 1219–1228. https://doi.org/10.1111/2041-210X.12423

Chambert, T., Grant, E. H., Miller, D. A., Nichols, J. D., Mulder, K. P., & Brand, A. B. (2018). Two-species occupancy modelling accounting for species misidentification and non-detection. *Methods in Ecology and Evolution*, *9*(6), 1468–1477. https://doi.org/10.1111/2041-210X.12985

Conn, P. B., Johnson, D. S., Williams, P. J., Melin, S. R., & Hooten, M. B. (2018). A guide to Bayesian model checking for ecologists. *Ecological Monographs*, *88*(4), 526–542. https://doi.org/10.1002/ecm.1314

Consonni, G., Fouskakis, D., Liseo, B., & Ntzoufras, I. (2018). Prior distributions for objective Bayesian analysis. *Bayesian Analysis*, *13*(2), 627–679. https://doi.org/10.1214/18-BA1103

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., & Wikle, C. K. (2009). Accounting for uncertainty in ecological analysis: The strengths and limitations of hierarchical statistical modelling. *Ecological Applications*, *19*(3), 553–570. https://doi.org/10.1890/07-0744.1

Crome, F. H. J., Thomas, M. R., & Moore, L. A. (1996). A novel Bayesian approach to assessing impacts of rain forest logging. *Ecological Applications*, *6*(4), 1104–1123. https://doi.org/10.2307/2269595

Dennis, B. (1996). Discussion: Should ecologists become Bayesians? *Ecological Applications*, *6*(4), 1095–1103. https://doi.org/10.2307/2269594

Dorazio, R. M. (2016). Bayesian data analysis in population ecology: Motivations, methods, and benefits. *Population Ecology*, *58*, 31–44. https://doi.org/10.1007/s10144-015-0503-4

Dorazio, R. M., & Johnson, F. A. (2003). Bayesian inference and decision theory – A framework for decision making in natural resource management. *Ecological Applications*, *13*(2), 556–563. https://doi.org/10.1890/1051-0761(2003)013[0556:BIADTA]2.0.CO;2

Dupuis, J. A., & Joachim, J. (2006). Bayesian estimation of species richness from quadrat sampling data in the presence of prior information. *Biometrics*, *62*(3), 706–712. https://doi.org/10.1111/j.1541-0420.2006.00524.x

Ellison, A. M. (2004). Bayesian inference in ecology. *Ecology Letters*, *7*(6), 509–520. https://doi.org/10.1111/j.1461-0248.2004.00603.x

Gabry, J., Simpson, D., Vehtari, A., Betancourt, M., & Gelman, A. (2019). Visualization in Bayesian workflow. *Journal of the Royal Statistical Society Series A*, *182*, 389–402. https://doi.org/10.1111/rssa.12378

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2014). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: Taylor & Francis Group.

Gelman, A., & Hennig, C. (2017). Beyond subjective and objective in statistics. *Journal of the Royal Statistical Society: Series A (Statistics in Society)*, *180*(4), 967–1033. https://doi.org/10.1111/rssa.12276

Hefley, T. J., Zhang, H., Gray, B. R., & Bouska, K. L. (2020). *Model selection for ecological community data using tree shrinkage priors* (Unpublished report). Retrieved from https://arxiv.org/abs/2005.14303

Hobbs, N. T., & Hooten, M. B. (2015). *Bayesian models: A statistical primer for ecologists*. Princeton, NJ: Princeton University Press.

Hooten, M. B., & Hobbs, N. T. (2015). A guide to Bayesian model selection for ecologists. *Ecological Monographs*, *85*(1), 3–28. https://doi.org/10.1890/14-0661.1

Jeffreys, H. (1946). An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London, Series A 186*(1007), 453–461. https://doi.org/10.1098/rspa.1946.0056

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430), 773–795. https://doi.org/10.1080/01621459.1995.10476572

Kass, R. E., & Wasserman, L. (1995). A reference Bayesian test for nested hypotheses and its relationship to the Schwarz criterion. *Journal of the American Statistical Association*, *90*(431), 928–934. https://doi.org/10.2307/2291327

Kéry, M., & Royle, J. A. (2016). *Applied hierarchical modelling in ecology: Analysis of distribution, abundance, and species richness in R and BUGS*. San Diego, CA: Elsevier Press.

Kuhnert, P. M., Martin, T. G., & Griffths, S. P. (2010). A guide to eliciting and using expert knowledge in Bayesian ecological models. *Ecology Letters*, *13*(7), 900–914. https://doi.org/10.1111/j.1461-0248.2010.01477.x

Lele, S. R., & Dennis, B. (2009). Bayesian methods for hierarchical models: Are ecologists making a Faustian Bargain? *Ecological Applications*, *19*(3), 581–584. https://doi.org/10.1890/08-0549.1

Link, W. A. (2013). A cautionary note on the discrete uniform prior for the binomial N. *Ecology*, *94*(10), 2173–2179. https://doi.org/10.1890/13-0176.1

Link, W. A., & Barker, R. J. (2006). Model weights and the foundations of multimodel inference. *Ecology*, *87*(10), 2626–2635. https://doi.org/10.1890/0012-9658(2006)87[2626:MWATFO]2.0.CO;2

Link, W. A., & Barker, R. J. (2010). *Bayesian Inference: With ecological applications*. Boston, MA: Academic Press.

Martin, T. G., Kuhnert, P. M., Mengersen, K., & Possingham, H. P. (2005). The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. *Ecological Applications*, *15*(1), 266–280. https://doi.org/10.1890/03-5400

McDonald-Madden, E., Probert, W. J., Hauser, C. E., Runge, M. C., Possingham, H. P., Jones, M. E., … Wintle, B. A. (2010). Active adaptive conservation of threatened species in the face of uncertainty. *Ecological Applications*, *20*(5), 1476–1489. https://doi.org/10.1890/09-0647.1

Morris, W. K., Vesk, P. A., & McCarthy, M. A. (2013). Profiting from pilot studies: Analysing mortality using Bayesian models with informative priors. *Basic and Applied Ecology*, *14*(1), 81–89. https://doi.org/10.1016/j.baae.2012.11.003

Morris, W. K., Vesk, P. A., McCarthy, M. A., Bunyavejchewin, S., & Baker, P. J. (2015). The neglected tool in the Bayesian ecologist's shed: A case study testing informative priors' effect on model accuracy. *Ecology and Evolution*, *5*(1), 102–108. https://doi.org/10.1002/ece3.1346

Nichols, J. D. (1991). Science, population ecology, and the management of the American Black Duck. *The Journal of Wildlife Management*, *55*(4), 790–799. https://doi.org/10.2307/3809533

Northrup, J. M., & Gerber, B. D. (2018). A comment on priors for Bayesian occupancy models. *PLoS ONE*, *13*(2), https://doi.org/10.1371/journal.pone.0192819

Ntzoufras, I. (2009). *Bayesian modeling using WinBUGS*. Hoboken, NJ: Wiley.

R Core Team. (2018). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.

Rodhouse, T. J., Rodriguez, R. M., Banner, K. M., Ormsbee, P. C., Barnett, J., & Irvine, K. M. (2019). Evidence of region-wide bat population decline from long-term monitoring and Bayesian occupancy models with empirically informed priors. *Ecology and Evolution*, *9*(19), 11078–11088. https://doi.org/10.1002/ece3.5612

Royle, J. A., & Dorazio, R. M. (2008). *Hierarchical modelling and inference in ecology: The analysis of data from populations, metapopulations, and communities*. San Diego, CA: Elsevier Press.

Runge, M. C., Converse, S. J., & Lyons, J. E. (2011). Which uncertainty? Using expert elicitation and expected value of information to design an adaptive program. *Biological Conservation*, *144*, 1214–1223. https://doi.org/10.1016/j.biocon.2010.12.020

Seaman III, J. W., Seaman Jr., J. W., & Stamey, J. D. (2012). Hidden dangers of specifying noninformative priors. *The American Statistician*, *66*(2), 77–84. https://doi.org/10.1080/00031305.2012.695938

Tuyl, F., Gerlach, R., & Mengersen, K. (2012). A comparison of Bayes-Laplace, Jeffreys and other priors. *The American Statistician*, *62*(1), 40–44. https://doi.org/10.1198/000313008X267839

Wade, P. R. (2000). Bayesian methods in conservation biology. *Conservation Biology*, *14*(5), 1308–1316. https://doi.org/10.1046/j.1523-1739.2000.99415.x

Williams, P. J., & Hooten, M. B. (2016). Combining statistical inference and decisions in ecology. *Ecological Applications*, *26*(6), 1930–1942. https://doi.org/10.1890/15-1593.1

Wright, W. J., Irvine, K. M., Almberg, E. S., & Litt, A. R. (2019). Modelling misclassification in multi-species acoustic data when estimating occupancy and relative activity. *Methods in Ecology and Evolution*. https://doi.org/10.1111/2041-210X.13315

Wright, W. J., Irvine, K. M., & Higgs, M. D. (2019). Identifying occupancy model inadequacies: Can residuals separately assess detection and presence? *Ecology*, *100*(6), e02703. https://doi.org/10.1002/ecy.2703

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.

How to cite this article: Banner KM, Irvine KM, Rodhouse TJ. The use of Bayesian priors in Ecology: The good, the bad and the not great. *Methods Ecol Evol*. 2020;11:882–889. https://doi.org/10.1111/2041-210X.13407