Introduction to Bayesian Modeling and Inference for Fisheries Scientists

Jason C. Doll[1] and Stephen J. Jacquemin[2]

[1] Quantitative Fisheries Center, Department of Fisheries and Wildlife, Michigan State University, 375 Wilson Road, Room 100, East Lansing, MI 48824.

[2] Department of Biological Sciences, Wright State University - Lake Campus, 7600 Lake Campus Drive, Celina, OH 45822

Formatted for Fisheries

Running Title: Bayesian Inference in Fisheries

Correspondence: Jason C. Doll, dolljas1@msu.edu, Quantitative Fisheries Center, Department of Fisheries and Wildlife, Michigan State University, 375 Wilson Road, Room 100, East Lansing, MI 48824.

Abstract

Bayesian inference is everywhere, from one of the most recent journal articles in Transactions of the American Fisheries Society to the decision making process you go through when you select a new fishing spot. Bayesian inference is the only statistical paradigm that synthesizes prior knowledge with newly collected data to facilitate a more informed decision – and it is being used at an increasing rate in almost every area of our profession. Thus, the goal of this article is to provide fisheries managers, educators, and students with a conceptual introduction to Bayesian inference. We do not assume the reader is familiar with Bayesian inference, however, we do assume the reader has completed an introductory biostatistics course. To this end, we review the conceptual foundation of Bayesian inference without the use of complex equations; present one example of using Bayesian inference to compare relative weight between two time periods; present one example of using prior information about von Bertalanffy growth parameters to improve parameter estimation; and finally, suggest readings that can help to develop the skills needed to use Bayesian inference in your own management or research program.

Introduction

Bayesian inference is rooted in the notion that past experiences or information can be combined with new information to help explain certain events or inform the probability of outcomes associated with specific events. Although you might not be actively using Bayesian inference in your research, you are most likely using it in your everyday life. Bayesian inference is in the minds of card counters at the blackjack table, in the algorithm that picks the pop up advertisements on your favorite social networking site, and in the unconscious decision making process you go through when you select a new fishing spot. It is a way of thinking, learning, and has been proposed as the way our minds process information to make decisions (De Ridder et al. 2014). This natural way of thinking is in contrast to how many analyze their data. For example, data are typically analyzed by calculating the probability of observing the data. This method of analyzing data is referred to as frequentist inference with Null Hypothesis Statistical Testing (NHST). When Bayesian inference is applied to data analysis, probabilities are assigned to certain outcomes given new information and making a decision based on the assigned probability.

If we used NHST in our everyday lives, we might find ourselves in a very disappointing situation. For example, suppose you are an avid Walleye *Sander vitreus* angler that is interested in fishing for nothing else. One day, while vacationing in Florida, you feel the itch to go fishing. However, after a lifetime of experience honing your Walleye fishing technique in the waters of your home state, Minnesota, you realize you don’t know how good Walleye fishing is in Florida. Because you were trained as a frequentist and to use NHST, you decide to perform an experiment to determine the quality of Walleye fishing in Florida. Thus, you form a null hypothesis that fishing for Walleye in Florida is no different than fishing for Walleye in

Minnesota. You know you catch approximately one fish per hour of fishing effort, and based on your null hypothesis of no difference, you predict you will catch one fish per hour of fishing in Florida. To test this hypothesis you fish – for days – and catch no Walleye. Minutes turn into hours, hours to days, and you catch no Walleye thus your trip was a failure. After you have collected enough data from an array of Florida lakes, rivers, and swamps, you calculate the probability of catching zero Walleye per hour in Florida (your data) is so small, that you conclude Walleye fishing in Florida is NOT the same as Walleye fishing in Minnesota! Thus rendering the assumption that Walleye fishing in Florida is no different from Minnesota false (i.e., rejecting the null hypothesis). In this scenario, the time spent collecting new data to make a conclusion about Walleye fishing in Florida could have been reduced if our angler from Minnesota would have used prior information on Walleye fishing in Florida. The reality is, however, that most of us looking to find a new fishing spot would just intuitively know to gather information before setting off to a new area for fishing. This information could come from blogs, overhearing conversations at the local fishing tackle shop (i.e., expert opinion), or distribution maps (i.e., published data). The point is, given a limited travel agenda searching for a new spot to go Walleye fishing; one would find themselves gravitating towards almost anywhere other than Florida….maybe the incredible spring Walleye fishery in western Lake Erie in Sandusky Ohio or even the St. Clair River run in southeastern Michigan.

Bayesian inference has been used to find lost ships, crack the unbreakable Enigma code of World War II, predict the outcomes of elections, forecast nuclear meltdowns, predict Major League Baseball player performances (McGrayne 2012), and most likely, been used at some point in your own lives to find a new fishing spot. Within our own field, and more recently, Bayesian inference has been used in a variety of analyses including generalized linear models, species distribution modeling, incorporating phylogeny into standard models describing trends in abundance, and stock assessments (Punt and Hilborn 1997; Jacquemin and Doll 2014; Rahikainen et al. 2014). Bayesian inference is all around us and commonly used in fisheries science, yet many may not be familiar enough with it to appreciate its flexibility to address both simple and complex problems, and how it can take advantage of all available information to help produce clear and direct inferences. Therefore, the obvious questions and focal points of this article become; "What is Bayesian inference?", "Why should I care about Bayesian inference?", and "What can Bayesian inference do for me?". We attempt to answer these questions here.

The goal of this article is to provide fisheries managers, educators, and students with an introduction to Bayesian inference with minimal equations so one can take the next step towards incorporating Bayesian inference in their quantitative toolbox, be better prepared to critique research that uses Bayesian inference, and teach the next generation of fisheries scientists. Herein, we provide a brief overview of what Bayesian inference is and demonstrate how Bayesian inference can be applied to fisheries data using two examples.

Bayesian Inference

What is Bayesian inference?

Bayesian inference uses a basic law of probability knows as Bayes' theorem. Bayes' theorem was discovered by the Presbyterian minister Thomas Bayes more than 250 years ago and later rediscovered in 1774 by Pierre Simon Laplace who described it in scientific applications. This simple probability rule combines what we already know about an event with new information to provide an updated belief about that event. Conceptually, what makes Bayesian methods unique is the incorporation of that prior information and reallocation of belief.

To better understand Bayesian inference, we find it helpful to draw contrasts to what we already know (frequentist inference and NHST) from introductory biostatistics. Bayesian inference defines probability as a measure of belief about an event or model parameter (e.g., what is the probability of mean catch rates *increasing* under the new management program?). Bayesian inference uses of Bayes' theorem (see below) to combine new data and any prior information. New data and prior information are incorporated by describing each with a probability distribution. The results are a posterior probability distribution that jointly describes the model parameters (e.g., all slope coefficients in a linear regression model). The posterior probability distribution of each parameter is often summarized as credible intervals (CI), which are a direct probability statement about the parameter of interest. Bayesian inference answers the basic question; "What is the probability of a hypothesis given our observed data and any prior information we might have?". This is in contrast to frequentist inference where probability is defined as how often something occurs in the long run (e.g., If I were to hypothetically replicate a study many times, what is the probability of the *observed* or more extreme mean catch rates, if the new management program is not effective?). Frequentist inference treats model parameters as fixed unknown values and the data as random. Frequentist inference makes decisions based on how unlikely the observed values are if there is no effect and draws conclusions about the size of the effect from 95% confidence intervals that are based on hypothetical replicates. The 95% confidence intervals tell us, given a hypothetically large number of surveys, how often (95%) our calculated confidence interval would overlap the true parameter's true value (noting that we have no way of knowing if our calculated 95% confidence interval overlaps the true value or not). This definition effectively renders the use of frequentist probability statements, which only apply to the *sampling*, useless as a direct measure of probability regarding a *specific* parameter. For example, we can say with 95% confidence that this sample of fish is not different from the main stock, but we can't say there is a 95% probability that this sample of fish is the same as the main stock. Frequentist inference with NHST uses *p*-values to answer the question; "What is the probability of observing our data or *more extreme* data given some hypothesis (i.e., specified statistical model) is true?".

Why should I care about Bayesian inference?

There are at least two reasons why you might care about Bayesian inference, either you want to be able to better understand and critique articles that use Bayesian methods or you want to incorporate Bayesian methods into your own quantitative toolbox for analyzing data. Bayesian inference is being used at an increasing rate in fisheries management. Since 2000, fisheries related journals have seen a rise in the number of papers that use Bayesian analyses (based on a topic search conducted June 2017 in Web of Science; Figure 1). Transactions of the American Fisheries Society (TAFS), Canadian Journal of Fisheries and Aquatic Sciences (CJFAS), and Fisheries Research (FR) have exhibited the most consistent increasing trend. In 2016, 7 (6.4%; TAFS), 7 (4.5%; CJFAS), and 11 (4.0%; FR) of their published articles had "Bayesian" in the topical keywords and employed the methodology in their analyses.

Bayesian inference has been gradually gaining momentum over the past few decades because of its many advantages over NHST and *p*-values. Interestingly, there is even a journal, albeit outside of our field (Basic and Applied Social Psychology), that put a blanket ban on NHST and *p*-values in favor of parameter estimation methods including Bayesian inference (Trafimow and Marks 2015). Additionally, the American Statistical Association (ASA) has clarified the use and interpretation of *p*-values by releasing the only formal policy statement released by the association (Wasserstein and Lazar 2016). This statement clarifies that *p*-values are not a measure of probability, do not measure the size of an effect, and cautions that policy decisions should not be made based solely on whether a *p*-value is below some threshold. The ASA policy statement also provides alternatives to *p*-values, such as Bayesian methods, that emphasize estimation over testing. Thus, understanding the general methodology of Bayesian inference and how it is interpreted can help you critique and understand this growing segment of the scientific literature.

There are many advantages to Bayesian inference. Some of the most tangible advantages include improving your ability to draw conclusions conditional on the data (and prior information), easily propagate uncertainty through hierarchical relationships, easily obtaining uncertainty for derived quantities, incorporating latent variables and functions thereof (e.g., hierarchical occupancy models; Royle and Kéry 2007), incorporating prior knowledge, describing more ecologically realistic models, and being able to express your findings in terms of probability that are easier for non-scientists to understand (see Kruschke (2010) for more details on the advantages of Bayesian inference).

What can Bayesian inference do for me?

Have you ever thought to yourself, "I wish I could tell this group of anglers there is some specific probability that the new management program will increase catch rates."? If you have, Bayesian inference can help you do that! Bayes' theorem is the only method of analyzing data to produce probabilities of different hypotheses (Gelman et al. 2014). Concluding probabilities of outcomes based upon different management scenarios has already been widely used in the management of the world's fisheries (methods synthesized in Punt and Hilborn 1997). Two recent examples of applied management studies that have used Bayesian inference include the development of mortality models to assess the outcomes of regulations on Largemouth Bass *Micropterus salmoides* populations, and to predict the results of a new size limit on Snapping Turtle *Chelydra serpentina* harvest (Kerns et al. 2015; Colteaux and Johnson 2017). Bayesian inference has even been used to inform the management of our favorite wandering Florida fisherman's target catch as Tsehaye et al. (2016) estimated probabilities of spawning stock biomass, harvest, and a population crash through the use of a hierarchical age-structured stock assessment model of Walleye. In your own work, Bayesian inference can provide outcome probabilities that can better inform your management decision, regardless of how simple or complex the analysis.

Bayes' Theorem

Bayesian inference uses probability theory as a formal way of incorporating new data with prior information to make a direct probability statement about a hypothesis – this is the foundation of Bayesian inference and is based on Bayes' theorem (Equation 1; Figure 2). According to Bayes' theorem, the posterior probability distribution, $p(\theta|X)$, of model parameters (θ) given observed data (X) is calculated by:

Equation 1: $$p(\theta|X) = \frac{p(X|\theta)p(\theta)}{\int p(X|\theta)p(\theta)d\theta}$$

Where: $p(X|\theta)$, the likelihood, denotes the probability distribution of the data given the parameters, $p(\theta)$ denotes the prior probability distribution of the model parameters, and the denominator is a normalizing parameter calculated by summing across all possible parameter values weighted by the strength of their belief to scale the results to be between 0 and 1. Thus, the posterior probability distribution equals the probability distribution of the data given the parameters, multiplied by the prior probability distribution of the model parameters, all divided by the sum across all possible probability distributions of data multiplied by all possible parameter values weighted by the strength of their belief. Conventionally, $p(X|\theta)$ is denoted as the likelihood. However, $p(X|\theta)$ is calculated from an assumed sampling distribution that is conditional on the data (X) not θ. That is, $p(X|\theta)$ is first defined and after the data (X) are observed, the same function is used and assumed to be proportional to the likelihood, such that $L(\theta|X) \propto p(X|\theta)$. For a thorough review of Bayes' theorem, see Gelman et al. (2014), Carlin and Louis (2008), and McElreath (2016).

The posterior probability distribution is used to make all statistical inference and represents all that is known about the parameter after combining the prior probability distribution with new data. All parameters in a model and all derived quantities (e.g., difference between two parameters; see relative weight example) have a posterior probability distribution. The posterior probability distribution can be summarized by its mean or median with the spread of the distribution summarized with quantiles. The most common summary of the posterior probability distribution to represent full uncertainty is the 95% CI. The 95% CI is the range of values that are bounded by the upper 97.5% and lower 2.5% quantiles of the probability distribution.

Prior information is arguably the most important and greatest advantage of Bayesian inference. Bayesian inference permits researchers to directly incorporate previous knowledge about model parameters in a transparent and defensible manner. Prior probability distributions measure how plausible all potential parameters values are before we see new data. When priors are based on the literature or expert opinions, they are considered "informative priors". In contrast, when the researchers have no basis to construct an informative prior distribution, all possible values are given equal probability and considered "reference priors" or "diffuse priors".

When prior probability distributions are based on reference priors, the mean of the posterior probability distributions, particularly with simple models (e.g., linear regression), are similar to the point estimates of frequentist inference. However, drastically different interpretations remain because of the underlying definitions of probability under the different paradigms (see *What is Bayesian inference?*). Further, some critics argue that the use of informative priors in model building may be considered subjective (Martin et al. 2012). Indeed, prior distributions and the reliance on these "priors" has been the subject of much debate (Dennis 1996; Huelsenbeck et al. 2002; McCarthy and Masters 2005). Nevertheless, the value of prior information cannot be discounted and Bayesian inference provides a transparent mechanism for its inclusion (Kuhnert et al. 2010). We would argue that there are very few situations where the researcher would truly have no prior information, and that their analyses would benefit from the inclusion of available prior information.

Informative prior probability distributions can be categorized in two way; "population" and "state of knowledge". The "population" category includes setting biologically realistic limits on the bounds of a parameter. For example, constraining estimates of detection probability to be between 0 and 1. The "state of knowledge" category includes expert knowledge and published literature. As we demonstrate later in this article, prior information based on the current state of knowledge allows us to make informed decisions where we would otherwise not have biologically relevant parameters (See von Bertalanffy growth model example). Many examples of incorporating informative prior probability distributions in fisheries applications can be found in the stock-assessment literature (McAllister and Ianelli 1997; Romakkaniemi 2015). Other ecological applications outside of fisheries include evaluating impacts of grazing on birds (Martin et al. 2005), estimating Mule Deer *Odocoileus hemionus* survival and abundance

(Lukacs et al. 2009), and estimating poaching mortality of the Wolves *Canis lupus* (Liberg et al. 2011). While incorporating informative prior information can increase the usefulness of your analysis, the prior probability distribution must be carefully selected and be supported by good science.

Combining the likelihood and prior probability distribution into the posterior probability distribution can generally not be accomplished using standard integral approximation and there is often no analytic solution. Thus, sampling techniques that calculate numerical approximations of model parameters are required to overcome these issues. Markov Chain Monte Carlo (MCMC) is the most common method of sampling from the posterior probability distribution, other less common methods of sampling from the posterior probability distribution includes grid search (Kruschke 2015) and sample-importance-resampling (Rubin 1988). A full description of MCMC methods is beyond the scope of this paper, thus we only present a cursory overview here. MCMC methods include several different algorithms that sample from the posterior distribution with a Markov Chain. Constructing a Markov Chain is a process that generates a series of random numbers that are dependent on the previous random number and nothing else. Most MCMC processes begin with one set of random numbers that represent parameters in your model. Then a new set of random numbers are generated and compared with the first. If the new set of random numbers provide a better fit given the data and prior information they are saved and then compared to a new set of random numbers. If they do not provide a better fit given the data and prior information they are not saved and a new set of random numbers are generated and compared to the initial set. This process continues for hundreds and often thousands of iterations until the saved values have converged on the posterior probability distribution. The different algorithms that perform MCMC differ in their proposals and accepting or rejecting criteria and there are a variety of methods to identify convergence to the posterior probability distribution. There is also an initial period in the chain that is removed because they are unlikely to have come from the posterior distribution. The initial part of the chain that is removed is called the "burn-in" period. It is also common to "thin" MCMC chains to remove the correlation between successive iterations and reduce the length of the chain and thus the amount of memory required to save the chain. This is accomplished by only saving every i^{th} step in the chain. The number of iterations to discard between saved steps will be dependent on the amount of correlation and total number of iterations used in the MCMC chain. It is common to set the number of thinning steps to 3 but values greater than 10 are not uncommon for very long chains (e.g., > 10,000 iterations). The end result is a series of "iterations" with values for each parameter being estimated or quantity being derived that represents the joint posterior probability distribution. A more technical description of how MCMC works can be found in Congdon (2007).

Applied fisheries examples

Here we present two increasingly complex fisheries examples to provide an applied framework using actual data. These examples are common and often first introduced in undergraduate fisheries management courses and in popular fisheries text books (Isely and Grabowski 2007; Walters and Martell 2004). Here we emphasize how the results are interpreted as a probability distribution of credible values as opposed to rejecting or failing to reject a null hypothesis. We also demonstrate the use of prior information to improve inference of a common fisheries model.

When conducting any analysis using Bayesian inference there are specific details that need to be included in the narrative that describes the methods. These include; software used, the number of concurrent MCMC chains, the total number of iterations, the number of burn-in steps, the number of thinning steps, the number of saved steps, and convergence diagnostics. The most common software used when fitting Bayesian models are JAGS (Plummer 2003), BUGS (Lunn et al. 2000), and Stan (Carpenter et al. 2017). All three are incorporated into the R programming environment through downloadable R packages. The specific details for the two examples presented here with complete model specification with JAGS and R code is available in the online appendix. Posterior probability distributions of the parameter estimates are summarized with their median and 95% CI. The analyses presented here are not intended to provide a thorough assessment of either fishery. Rather, we use these example as applications of Bayesian inference to common fisheries scenarios.

Comparison of fish condition between years using relative weight (Bayesian t-test)

Relative weight is a common fisheries management metric that is used to monitor the response of a fish population due to regulation changes (Blackwell et al. 2000) and provides a familiar and practical example. Relative weight is the ratio of the weight of an individual fish to a standard weight for a given length scaled to be between 0 and 100.

Equation 2: $$\mathrm{W_{ri}} = \left(\frac{\mathrm{W_i}}{\mathrm{W_{si}}}\right) \times 100$$

Where W_{ri} is the relative weight of individual fish i, W_i is the weight of individual fish i, and W_{si} is a length-specific standard weight predicted by the weight-length regression for individual fish i. The specific equation used to calculate W_s comes from regional species specific weight-length formula (Neumann et al. 2012).

Equation 3: $$\log_{10}(\mathrm{W_s}) = a + b * \log_{10}(\mathrm{TL})$$

Where *a* is the intercept, *b* is the slope, and TL is total length of the individual fish. Here, we will compare W_r's between two groups using the Bayesian two-sample *t*-test (Kruschke 2012).

The Bayesian two-sample t-test is simply the comparison of group means and is analogous to the frequentist two-sample *t*-test. However, the key difference between the two is that the frequentist *t*-test is only comparing group means to determine if they are "significantly" different whereas the Bayesian t-test is comparing the groups mean and uncertainty (i.e., standard deviation) to determine a difference and *how* much different. The Bayesian two-sample t-test describes the data from both groups with a normal distribution (a t-distribution can be used as an alternative to account for outliers).

Equation 4: $y_{ki} \sim normal(\mu_k, \sigma_k)$

Where y_{ki} is the observed W_r for individual *i* in group *k*, μ_k is the mean of group *k*, σ_k is the standard deviation of group *k*. Note the normal distribution in JAGS is parameterized with the mean and precision ($1/\sigma_k^2$). Reference priors are used for μ_k and σ_k.

Equation 5: $\mu_k \sim normal(0{,}1000)$

Equation 6: $\sigma_k \sim uniform(0{,}20)$

The assumptions for a Bayesian two-sample t-test (comparison between group means) do not change because we are using Bayesian inference. (i.e., we assume independent observations and do not assume equal variances).

Application: Yellow Perch (Perca flavescens) relative weight long term (1992 vs 2002) comparisons

Notable changes in management regulations and the Lake Michigan ecosystem have occurred over the past 30 years (Madenjian et al. 2002). Within Indiana waters, commercial fishing was closed in 1997 and a daily recreational creel limit of 15 fish was imposed. In addition to this, invasive species introductions, such as the Zebra Mussel *Dreissena polymorpha*, Quagga Mussels *Dreissena bugensis*, and Round Goby *Neogobius melanostomus* have altered the food web (Griffiths et al. 1991; Lauer et al. 2004; Nalepa et al. 2009). The numerous factors affecting Yellow Perch provide an ideal situation to evaluate changes in mean and standard deviation of relative weight as an applied example of Bayesian inference. We will use a Bayesian two-sample t-test (Kruschke 2012) to determine if average and standard deviation of W_r has changed after a 10-year period (1992 and 2002) and if they have changed, how much have they changed?

The data used for this analysis come from a long-term monitoring program in Southern Lake Michigan. Yellow Perch were sampled at three fixed sites using nighttime bottom trawling at the 5-m depth contour in 1992 and 2002 (other years are available, however we are only using two years of data for demonstration purposes only). For more details about the sampling program see Forsythe et al. (2012). Sites were sampled twice each month (July and August) for a total effort of 12 h each year. After each night, a random subsample of 300 Yellow Perch age ≥ 1 were measured for total length and total weight. Standard weight was calculated by using the parameters reported in Neumann et al. (2012).

A total of 1,701 fish were included in this analysis. Relative weights ranged from 42.2 to 142.4 (Figure 3). Mean W_r in 1992 and 2002 was 86.3 (95% CI = 85.4 to 87.3) and 73.2 (Figure 4; 95% CI = 72.7 to 73.7). The distributions of mean W_r in 1992 and 2002 are clearly different, but one important question the manager will often ask is, *how much* different are mean W_r's in 2002 compared to 1992. This question can easily be answered under the Bayesian approach. In this example, we can simply subtract the posterior probability distributions of the mean W_r in 1992 from 2002. By doing this, we obtain a derived parameter with a measure of uncertainty, a result that is more difficult to obtain under the frequentist paradigm with NHST. The change in W_r corresponds to a decrease in mean W_r of 13.1 (95% CI = 12.2 to 14.2) from 1992 to 2002. These results are interpreted as there being a probability of 0.95 that mean W_r has decreased between 12.2 to 14.2. Estimates of standard deviation in 1992 and 2002 were 11.3 (95% CI = 10.7 to 11.9) and 9.1 (95% CI = 8.8 to 9.5), respectively, indicating a decrease in variability of 2.2 (95% CI = 1.5 to 2.9) from 1992 to 2002.

Because of the rich information contained in the results of Bayesian inference, we can begin to ask questions that have direct and meaningful implications for management. As we have discussed, the results represent probability distributions about parameters (e.g., W_r). Thus, the percentage of the posterior probability distribution that is greater than, less than, or between management benchmarks represent the probability of reaching that specific benchmark. For example, suppose a management benchmark for mean W_r is 73 (or this could be any specific W_r that managers are interested in) and if this benchmark was reached, we would conclude that some management action should be taken. To evaluate this scenario we would calculate the probability that mean W_r in 2002 is 73 or less. This is accomplished by determining the percentage of the posterior probability distribution of the mean W_r in 2002 that is less than 73 (total number of iterations in the posterior probability distribution that are 73 or less divided by the total number of iterations in the posterior probability distribution). In this example, we find that there is a probability of 0.51 that the mean W_r is less than or equal to 73. The fisheries manager can use this calculated probability to make a conclusion on if a new management action should be taken. Similarly, suppose we decide that 73 is too low and a W_r of 80 or less would warrant some management action. Here, the entire posterior probability distribution for W_r in 2002 is less than 80 and thus, there is a probability of 1.00 that W_r is less than 80. Features such as generating probabilities of achieving management benchmarks make Bayesian methods desirable for management decisions.

Evaluating growth using the von Bertalanffy model (non-linear regression)

Understanding how individual organisms change in length over time is one of the fundamental pieces of information used in fisheries management. The change in length over time is typically assessed with length-at-age data acquired from observing annular rings on some bony structure (e.g., otolith, spines, opercle, etc.). Information on growth rates is used to predict future yield (Quist et al. 2010) and set harvest limits (Reed and Davies 1991). To estimate growth rates a biologist must select a growth model that plausibly reflects the relationship between length and age data. The von Bertalanffy growth model is one of the most common models to describe organisms' growth (Doll et al. 2017; Hupfeld et al. 2016; Midway et al. 2015; Ogle et al. 2017).

Equation 9: $$\mathrm{y}_i = L_\infty\left(1 - e^{-\kappa(\mathrm{age}_i - t_0)}\right) + \varepsilon_i$$

Equation 10: $$\varepsilon_i \sim normal(0, \sigma)$$

Where y_i is the length of fish *i*, L_∞ is the hypothetical maximum mean total length achieved, κ is the Brody growth coefficient with units t^{-1}, age_i is the age of fish *i*, t_0 is the age when individuals would have been length 0, and ε_i is a random error term with mean 0 and standard deviation σ. Note the normal distribution in JAGS is parameterized with the mean and precision ($1/\sigma^2$).

Application: Monroe Reservoir Walleye (Sander vitreus) age and growth

For this application, we use Bayesian inference with a non-linear regression model to estimate parameters associated with the von Bertalanffy growth model. We additionally incorporate prior information about model parameters. The data used for this analysis come from

Walleye sampling conducted at Monroe Reservoir (Brown and Monroe Counties, Indiana) in October 2011 using 18 overnight experimental mesh gill net sets. Scale samples were taken from all Walleye for age and growth determination. For more information about the sampling protocol at Monroe Reservoir, see Kittaka (2008).

We estimated parameters using reference prior probability distributions and also extended the model to incorporate informative prior probability distributions (Table 1). The parameters L_∞ and κ were estimated on the log scale to restrict these parameters to be positive. Informative prior probability distributions were obtained from existing Walleye records at FishBase.org (Froese and Pauly 2017). We only included records that were from the United States and had estimates for all parameters, L_∞, κ, and t_0. This resulted in 26 observations for each parameter. Prior probability distributions were specified by taking the arithmetic mean and standard deviation of each parameter. Note that the prior probability distribution for L_∞ and κ are the mean and standard deviation are on the log scale.

Thirty-three fish were included in the analysis. Total lengths ranged from 33cm to 64cm and ages ranged from one to nine. Only one age six and one age nine fish were observed. Estimates of L_∞ were higher with reference prior probability distributions, while κ and t_0 estimates were lower with reference prior probability distributions (Table 2, Figure 5). Reference prior probability distributions resulted in greater uncertainty (i.e., wider 95% CI) compared to informative prior probability distributions (Table 2, Figure 5) for all parameters. Incorporating informative prior probability distributions also resulted in increased standard deviation (Figure 6) to accommodate the data and information in the prior probability distribution.

Looking at the study as a whole something is very apparent – this Walleye dataset contained few older fish, a scenario that is common in routine fisheries surveys. Yet, through the use of informative prior probability distributions we can be better prepared to deal with data sets such as these. If prior information was not included here, the lack of older fish resulted in unrealistic estimates of L_∞ because the curve does not reach an asymptote (Figure 7) and thus limits practical use of the results. Although Walleye have been collected over 70 cm, the majority of individuals are typically under 60 cm (Kittaka 2008). Thus, an average L_∞ greater than 65 cm is not a biologically realistic scenario. Further, our estimate of κ using reference prior probability distribution resulted in the center of the posterior probability distribution (0.06) as being lower than any value reported at FishBase.org in the United States (Table 2, Figure 7). This immediately suggests our estimate without incorporating prior information is biased low. Assessing growth information from limited data can often be misleading due to lack of older fish and inferences drawn using reference prior probabilities can result in inaccurate conclusions. In this example, using informative prior probability distributions resulted in more biologically realistic parameter estimates.

This Walleye example demonstrated two key aspects of Bayesian inference. The first is reallocation of belief. Incorporating new data reallocated the probabilistic belief to a new posterior probability distribution with reference and informative priors (Figure 5; A to B and C to D). The second key aspect of Bayesian inference demonstrated in this example is that prior information can be incorporated in the form of a prior probability distribution. The informative prior probability used in this example is the reason why the posterior probability distribution was more biologically realistic. The biological realism was worked into the model from the beginning by including information from 26 other studies.

Conclusion

Bayesian inference is a powerful and flexible tool that can be useful to all fisheries professionals. Although being able to make direct probabilistic statements about a hypothesis is desirable, perhaps the most advantageous aspect of Bayesian inference is being able to formally incorporate prior information in a defensible and logical way. Our field has grown substantially in its literature base over the past century and it seems worthwhile to stand on these past studies as we reach towards new and higher syntheses in the fisheries world. The literature provides a vast library of data that researchers can use to develop informative prior distributions, and is already being used in fisheries stock assessments (Punt and Hilborn 1997). There is no reason we should not incorporate this historical information into our research and management programs. Herein, we provided one example of how to understand and incorporate prior information into common fisheries models. There are many available sources that provide additional examples and details as to how one can incorporate prior information into your research (Millar 2002; McCarthy and Masters 2005; Kuhnert et al. 2010; Martin et al. 2012).

Our goal with this article is to provide fisheries managers, educators, and students with an introduction to Bayesian inference. This is intended to be the first step towards a more complete understanding of what Bayesian inference is, when to use Bayesian inference, and how to apply Bayesian inference in one’s own research. There are many articles and books available to help readers with the most elementary as well as the most advanced steps (Kéry 2010; Parent and Rivot 2013; Gelman et al. 2014; Kruschke 2015).

Bayesian inference is not a panacea and should not be viewed as a one-size-fits-all method of analysis. Although many do tend to prefer Bayesian methods, one needs to also be pragmatic and view Bayesian inference as another tool to use when needed. Ultimately, what is most important, is that when a problem is approached, that the best statistical method to answer the question at hand is used.

Acknowledgements

We are grateful for the helpful comments and suggestions provided by Derek Ogle and Sandra Clark-Kolaks on earlier drafts. We thank Jeff Schaeffer who encouraged the topic for this article and the anonymous reviewers who helped to improve it. We thank Aaron Lerner for providing the data on the number of papers published in AFS Journals. We also thank Brian Breidert, Sandra Clark-Kolaks, Dave Kittaka, the Indiana Department of Natural Resources, and Ball State University for providing the Yellow Perch data from Lake Michigan and Walleye age and growth data from Monroe Reservoir. This is publication 20YY-NN of the Quantitative Fisheries Center at Michigan State University.

Table 1. Prior probability distributions used in the von Bertalanffy growth model. The normal distribution is parameterized with the mean and precision ($1/\sigma^2$) and the uniform distribution is parameterized with minimum and maximum values. A reference prior for standard deviation (σ) was used under both conditions.

Parameter	Reference prior probability	Informative prior probability
$\mathrm{Log}(L_\infty)$	Normal(0, 1/1000)	Normal(4.27, 0.351)
$\mathrm{Log}(\kappa)$	Normal(0, 1/1000)	Normal(-1.16, 0.546)
t_0	Normal(0, 1/1000)	Normal(-0.47, 0.522)
σ	Uniform(0, 100)	

Table 2. Posterior probability distributions from the von Bertalanffy growth model based on reference and informative prior probability distributions, reported as median (lower and upper 95% Credible Interval). Note, L_∞ (cm) and k (y^{-1}) have been back transformed to the original scale.

Parameter	Reference prior probability	Informative prior probability
L_∞	109 (64, 505)	58 (54, 63)
κ	0.06 (0.01, 0.22)	0.40 (0.31, 0.53)
t_0	-5.75 (-9.24, -2.58)	-1.12 (-1.54, -0.72)
σ	2.83 (2.23, 3.75)	3.33 (2.58, 4.48)

Figure 1. Stacked frequency plot showing the time series of number of published articles that use Bayesian analysis in fisheries related journals by year between 2000 and 2016. Journals grouped by different shades of gray.

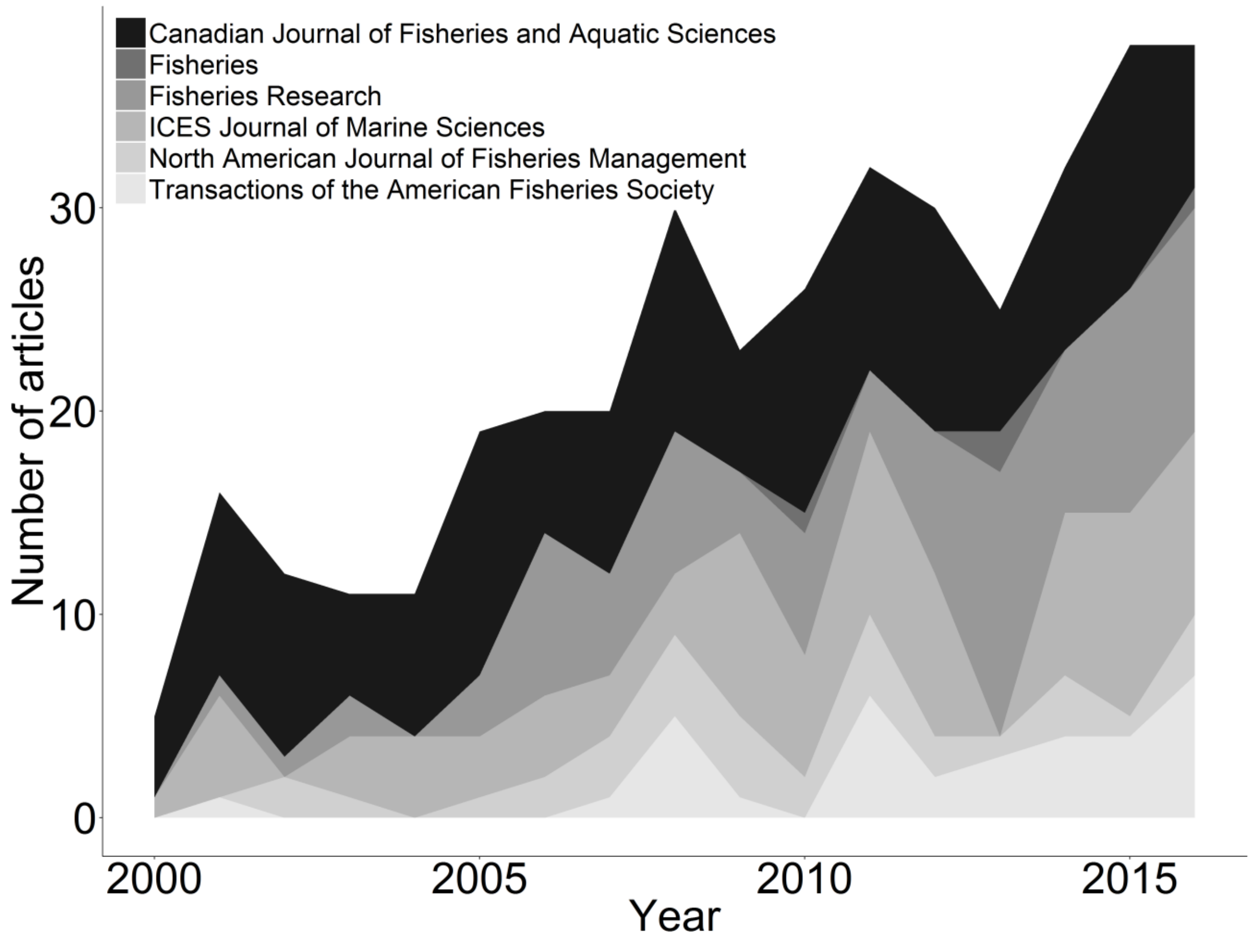


Figure 2. Bayesian inference flow chart using a description (A) and equations (B).

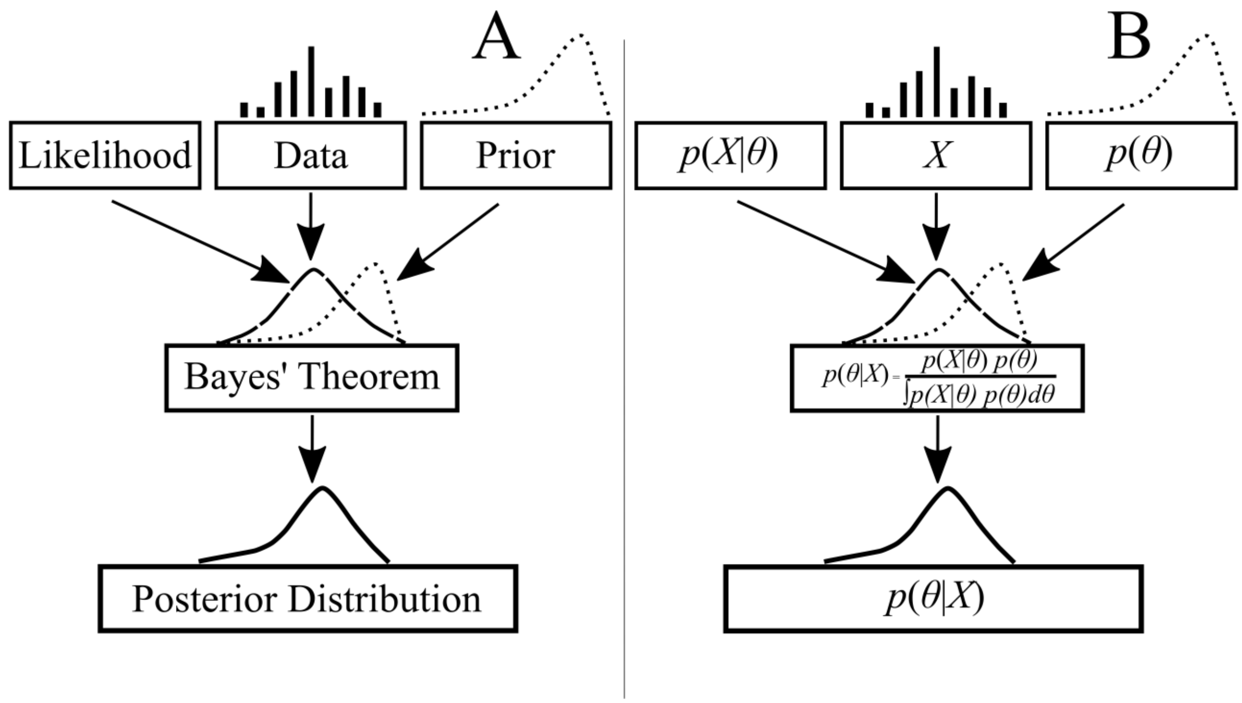


Figure 3. Histogram of data distribution of relative weight from 1992 (light gray) and 2002 (black). Dark gray histogram represents the prior probability distribution of the mean relative weight for each group, see Figure 4 for full prior probability distribution.

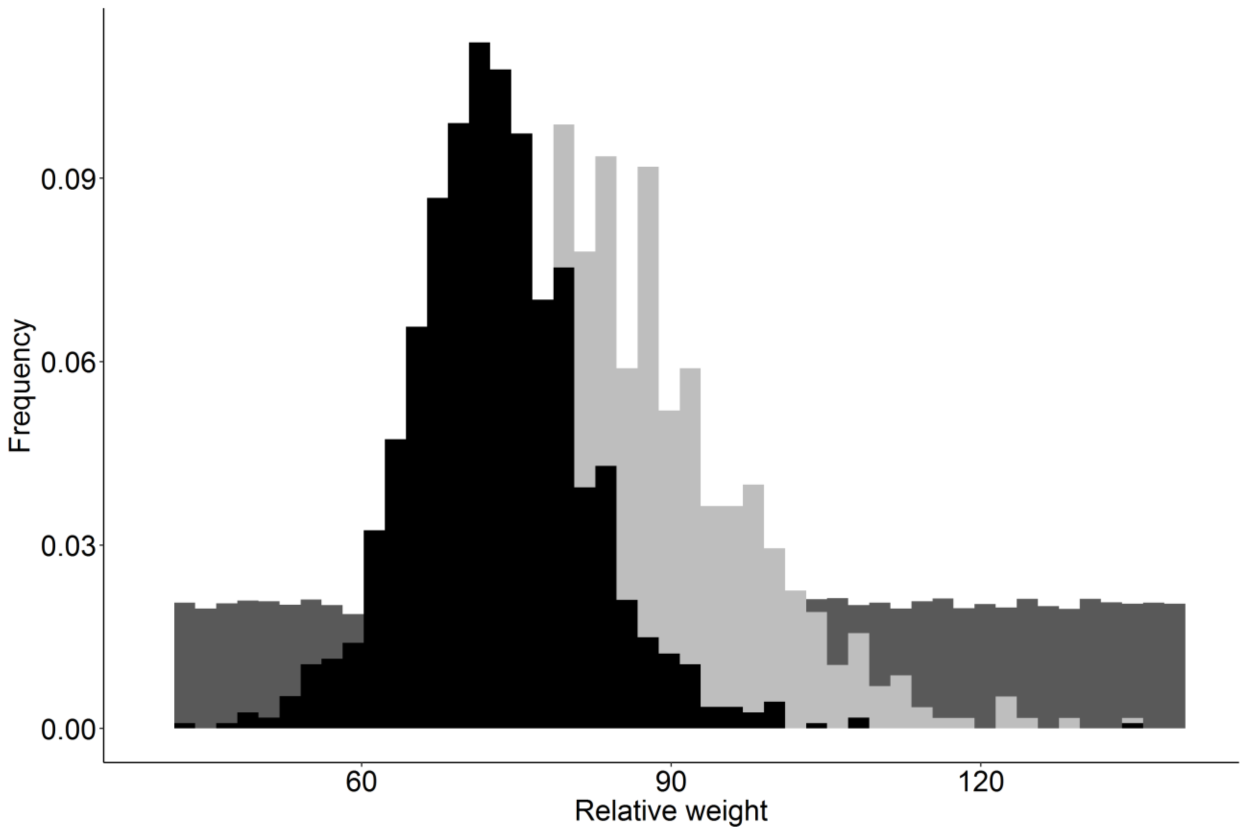


Figure 4. Prior probability distribution (inlayed dark gray histogram and dark grey histogram in main figure) and posterior probability distribution of mean relative weight in 1992 (light gray histogram) and 2002 (black histogram). Prior probability distribution is inlayed to show the full probability distribution because it appears flat when the x-axis is scaled to show details of posterior probability distribution.

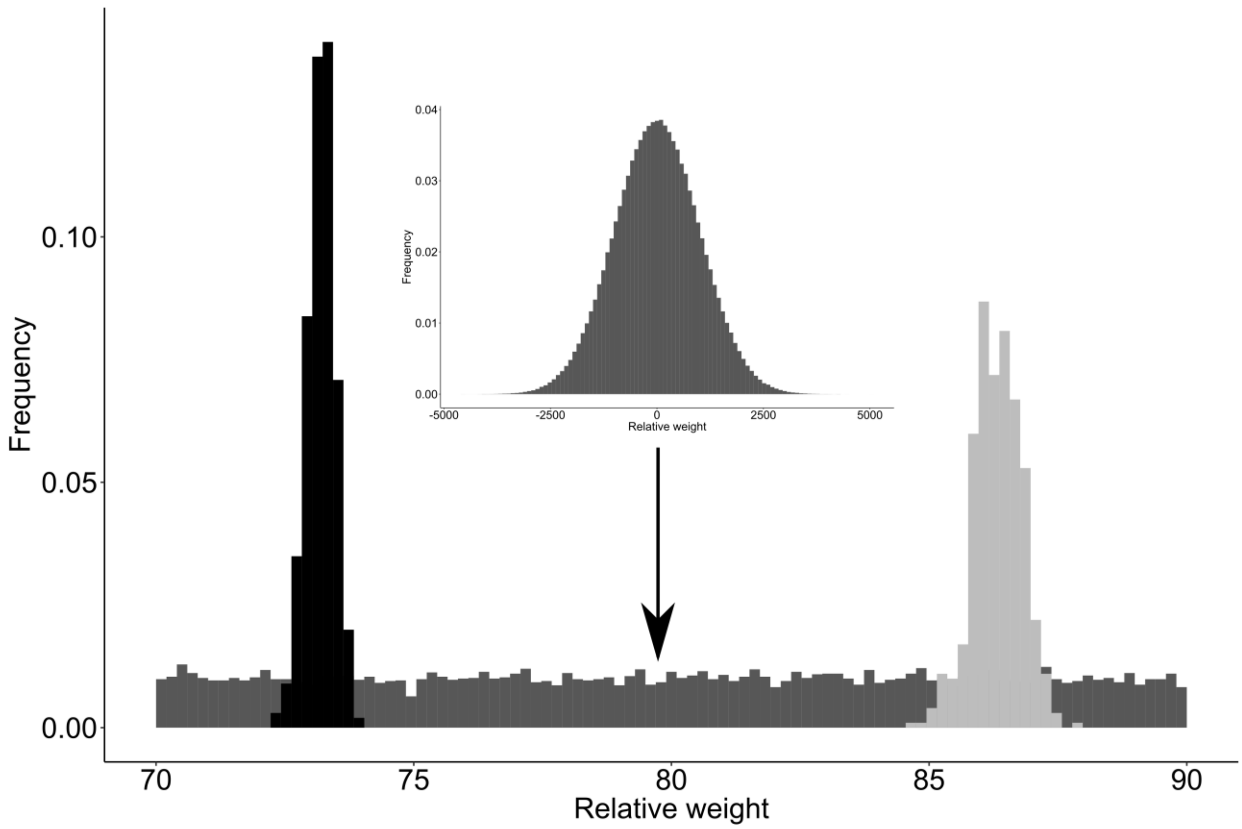
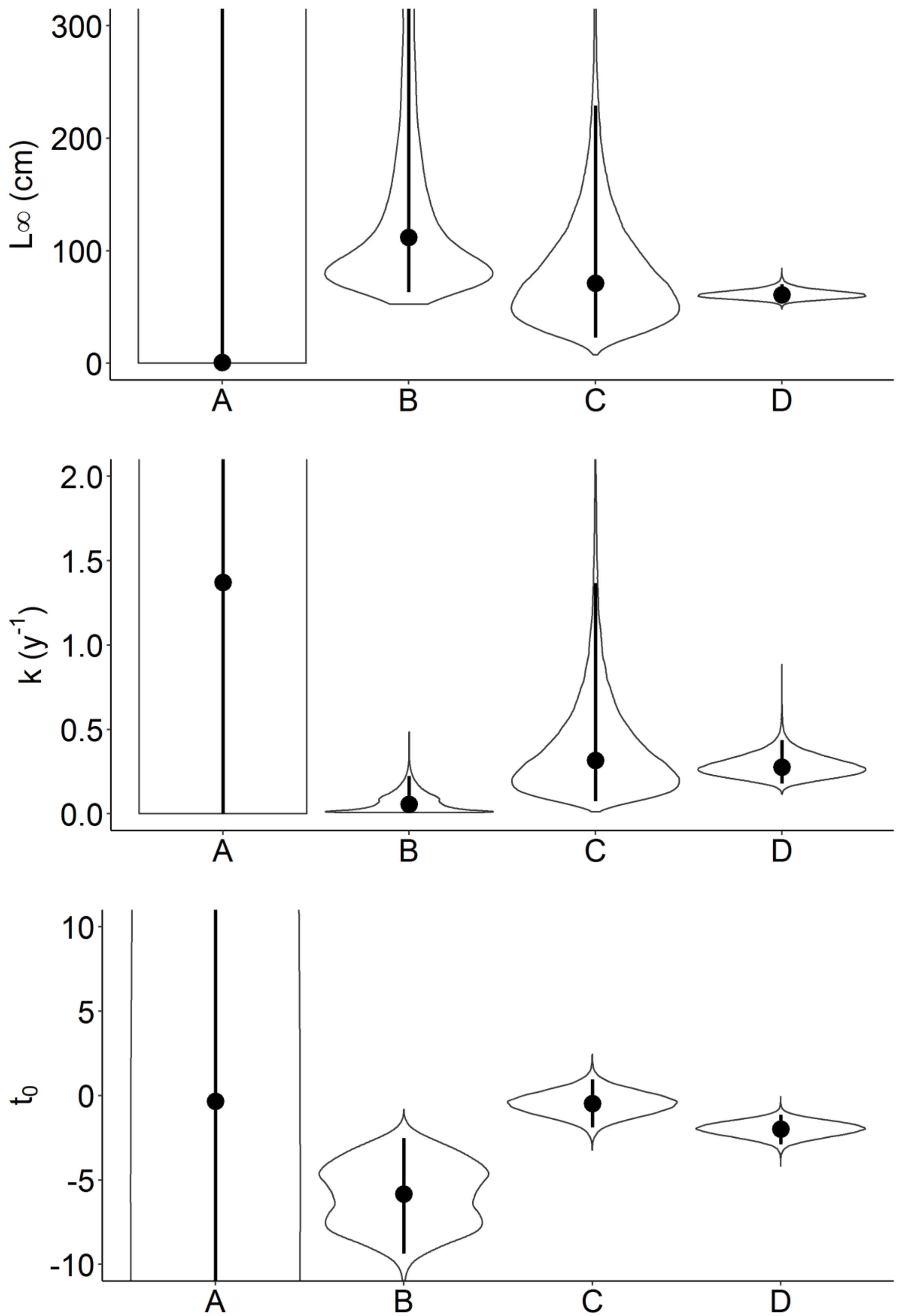


Figure 5. Violin plots of probability distributions for parameters of the von Bertalanffy growth model. Area within the violin plot represent the probability of parameter values, the widest portion of the violin plot indicates the highest probability. Solid points represent median of the probability distribution and solid lines represent 95% Credible Intervals. Group A are reference prior probability distributions for each parameter (because of the extreme uncertainty in the prior, it appears flat), Group B are the posterior probability distributions based on reference prior probability distribution, Group C are informative prior probability distribution for each parameter (see Table 1 for details), and Group D are the posterior probability distributions based on informative prior probability distribution.



525

Figure 6. Posterior probability distribution of the standard deviation from the von Bertalanffy mode using reference prior probability distributions (left) and informative prior probability distributions (right).

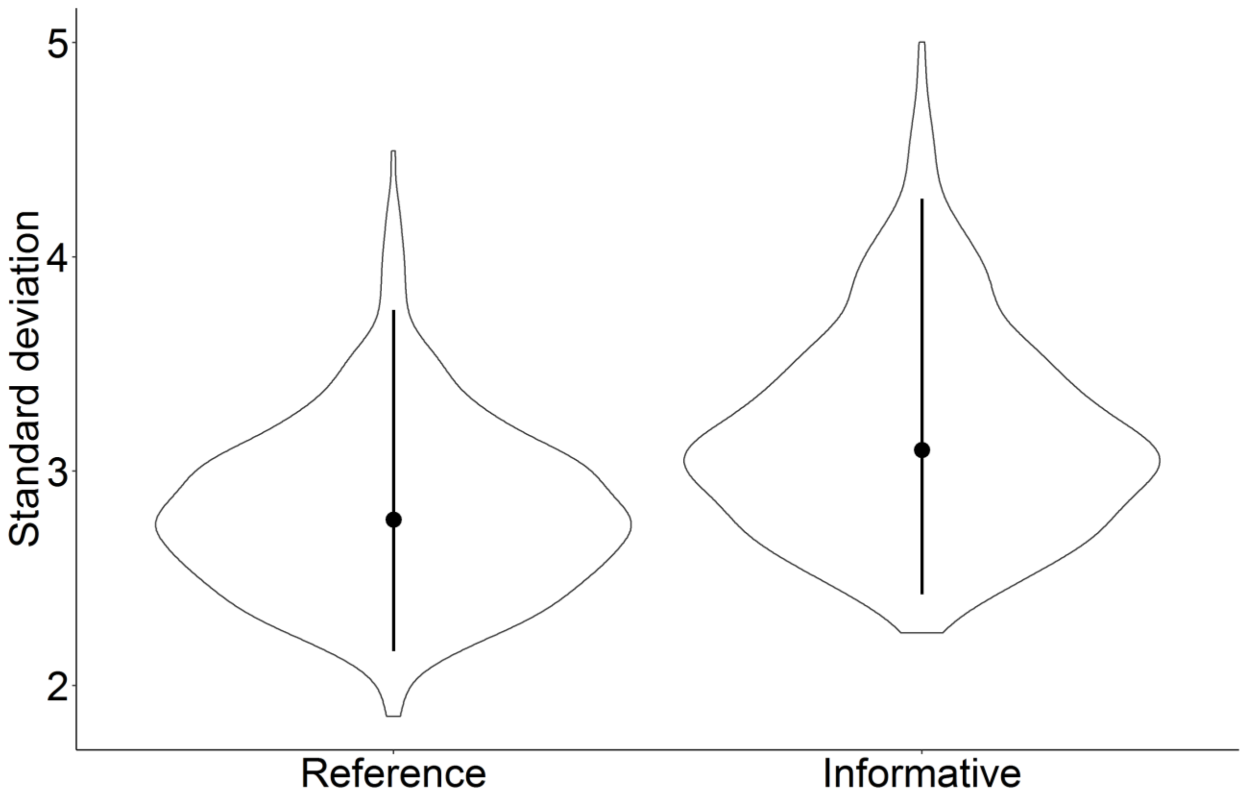
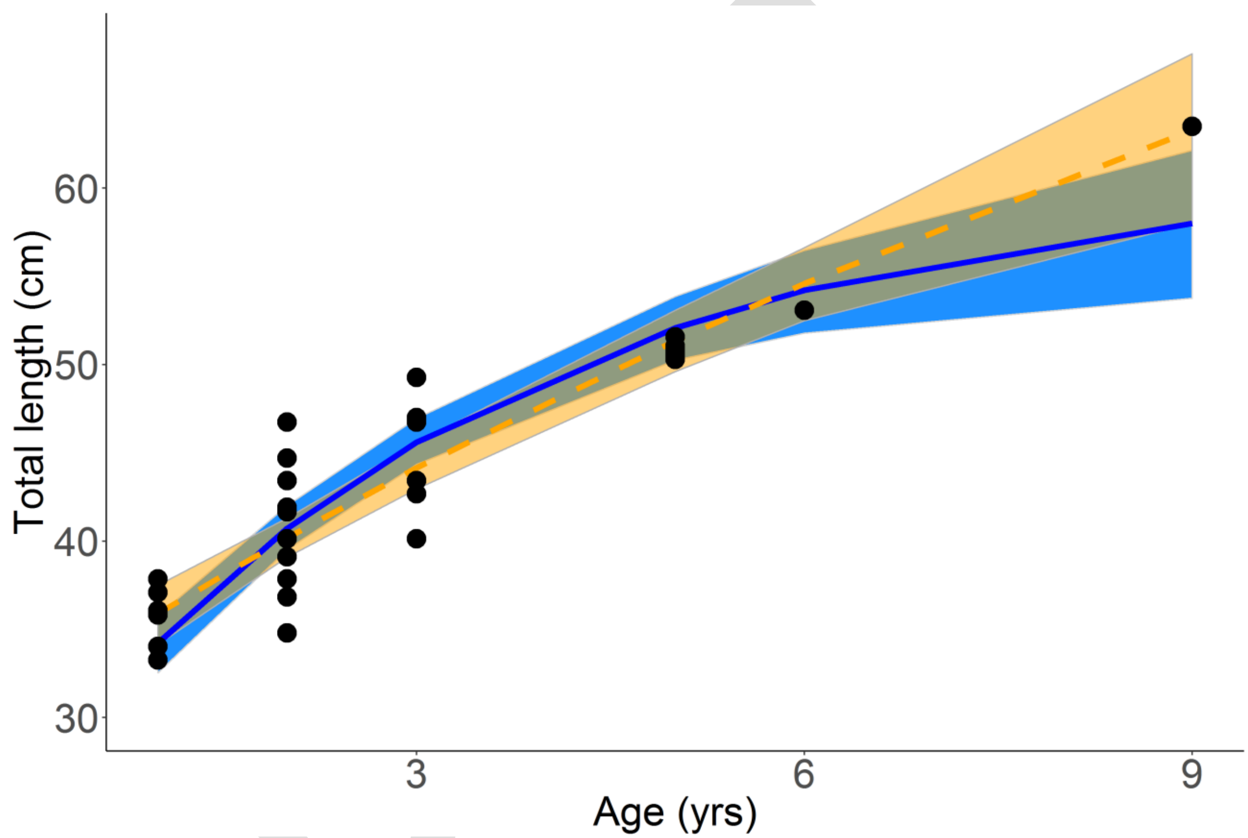


Figure 7. Mean growth curves based on reference (orange) and informative (blue) prior probability distributions. Points represent observed values, dashed line is the median of the posterior probability distribution with reference prior probability distributions, solid line is the median of the posterior probability distribution with informative prior probability distributions, shaded areas represent the 95% credible regions for the reference (orange) and informative (blue) prior probability distributions.



LITERATURE CITED

Blackwell, B.G., M.L. Brown, and D.W. Willis. 2000. Relative weight (W_r) status and current use in fisheries assessment and management. Reviews in Fisheries Science 8(1):1-44.

Carpenter, B., A. Gelman, M.D. Hoffman, D. Lee, B. Goodrich, M. Bentaourt, M. Brubaker, J. Guo, P. Li, and A. Riddell. 2017. Stan: A probabilistic programming language. Journal of Statistical Software 76(1). DOI 10.18637/jss.v076.i01

Carlin, B.P., and T.A. Louis. 2008. Bayesian methods for data analysis. Third Edition. CRC Press, Taylor & Francis Group. Boca Raton, Florida.

Colteaux, B.C., and D.M. Johnson. 2017. Commercial harvest and export of snapping turtles (*Chelydra serpentina*) in the United States: trends and the efficacy of size limits at reducing harvest. Journal of Nature Conservation 35:13-19.

Congdon, P. 2007. Applied Bayesian Modelling. Second Edition. Wiley & Sons, Ltd. United Kingdom.

De Ridder, D., S. Vanneste, and W. Freeman. 2014. The Bayesian brain: Phantom percepts resolve sensory uncertainty. Neuroscience and Biobehavioral Reviews 44:4-15.

Dennis, B. 1996. Should ecologists become Bayesians? Ecological Applications 6(4):1095-1103.

Doll, J.D., T.E. Lauer, and S. Clark-Kolaks. 2017. Yield-per-recruit modeling of two piscivores in a Midwestern reservoir: A Bayesian approach. Fisheries Research 191:200-210.

Forsythe, P.S., J.C. Doll, and T.E. Lauer. 2012. Abiotic and biotic correlates of Yellow Perch recruitment to age-2 in southern Lake Michigan, 1984-2007. Fisheries Management and Ecology 19:389-399.

Froese, R. and D. Pauly. 2017. FishBase. World Wide Web electronic publication. www.fishbase.org. (accessed 07/01/2017).

Gelman, A., J.B. Carlin, H.S. Stern, D.B. Dunson, A. Vehtari, and D.B. Rubin. 2014. Bayesian Data Analysis. Third Edition. CRC Press, Taylor & Francis Group. Boca Raton, Florida.

Griffiths, R.W., D.W. Schloesser, J.H. Leach, and W.P Kovalak. 1991. Distribution and dispersal of the Zebra Mussel (*Dreissena polymorpha*) in the Great Lakes region. Canadian Journal of Fisheries and Aquatic Sciences 48(8):1381-1388.

Huelsenbeck, J.P., B. Larget, R.E. Miller, and F. Ronquist. 2002. Potential applications and pitfalls of Bayesian inference in phylogeny. Systematic Biology 51(5):673-688.

Hupfeld, R.N., Q.E. Phelps, S.J. Tripp, and D.P. Herzog. 2016. Mississippi River Basin Paddlefish population dynamics. Fisheries 41(10):600-610.

Isely, J.J., and T.B. Grabowski. 2007. Age and Growth. Pages 187-228. *in* C.S. Guy and M.L. Brown, editors. Analysis and interpretation of freshwater fisheries data. American Fisheries Society, Bethesda, Maryland.

Jacquemin, S.J., and J.C. Doll. 2014. Body size and geographic range do not explain long term variation in fish populations: A Bayesian phylogenetic approach to testing assembly processes in stream fish assemblages. PLos ONE 9(4): e93522. doi:10.1371/journal.pone.0093522

Kerns, J.A., M.S. Allen, J.R. Dotson, and J.E. Hightower. 2015. Estimating regional fishing mortality for freshwater systems: a Florida Largemouth Bass example. North American Journal of Fisheries Management 35:681-689.

Kéry, M. 2010. Introduction to WinBUGS for ecologists. Elsevier, Burlington, Massachusetts.

Kittaka, D.S. 2008. Monroe Reservoir, Monroe and Brown Counties, 2007 Fish community survey report. Fisheries Section, Indiana Department of Natural Resources

Kruschke, J.K. 2010. What to believe: Bayesian methods for data analysis. Trends Cognitive Sciences 14, 293–300

Kruschke, J.K. 2012. Bayesian estimation supersedes the t test. Journal of Experimental Psychology: General 142(2):573-603. doi: 10.1037/a0029146

Kruschke, J.K. 2015. Doing Bayesian data analysis: a tutorial with R and BUGS. Second edition. Elsevier, Burlington, Massachusetts.

Kuhnert, P.M., T.G. Martin, and S.P. Griffiths. 2010. A guide to eliciting and using expert knowledge in Bayesian ecological models. Ecology Letters 13:900-914.

Lauer, T.E., P.J. Allen, and T.S. McComish. 2004. Changes in Mottled Sculpin and Johnny Darter trawl catches after the appearance of Round Gobies in the Indiana waters of Lake Michigan. Transactions of the American Fisheries Society 133(1):185-189.

Liberg, O. G. Chapron, P. Wabakken, H.C. Pedersen, N.T. Hobbs, and H. Sand. 2011. Shoot, shovel, and shut up: cryptic poaching slows restoration of a large carnivore in Europe. Proceedings of the Royal Society B 279:910-915.

Lukacs, P.M., G.C. White, B.E. Watkins, R.H. Hahn, B.A. Banulis, D.J. Banulis, D.J. Finley, A.A. Holland, J.A. Martens, and J. Vayhinger. 2009. Separating components of variation in survival of Mule Deer in Colorado. The Journal of Wildlife Management 73(6):817-826.

Lunn, D.J., A. Thomas, N. Best, and D. Spiegelhalter. 2000. WinBUGS – a Bayesian modeling framework: concepts, structure, and extensibility. Statistics and Computing 10:325-337.

Madenjian, C. P., G. L. Fahnenstiel, T. H. Johengen, T. F. Nalepa, H. A. Vanderploeg, G. W. Fleischer, P. J. Schneeberger, D. M. Benjamin, E. B. Smith, J. R. Bence, E. S. Rutherford, D. S. Lavis, D. M. Robertson, D. J. Jude, and M. P. Eberner. 2002. Dynamics of the Lake Michigan food web, 1970-2000. Canadian Journal of Fisheries and Aquatic Sciences 59:736-753.

Martin, T.G., M.A. Burgman, F. Fidler, P.M. Kuhnert, S. Low-Choy, M. McBride, and K. Mengersen. 2012. Eliciting expert knowledge in conservation science. Conservation Biology 26(1):29-38.

Martin, T.G., P.M. Kuhnert, K. Mengersen, and H.P. Possingham. 2005. The power of expert opinion in ecological models using Bayesian methods: Impact of grazing on birds. Ecological Applications 15(1):266-280.

McAllister, M.D., and J.N. Ianelli. 1997. Bayesian stock assessment using catch-age data and the sampling-importance resampling algorithm. Canadian Journal of Fisheries and Aquatic Sciences 54:284-300.

McCarthy, M.S. and P. Masters. 2005. Profiting from prior information in Bayesian analysis of ecological data. Journal of Applied Ecology 42:1012-1019.

McElreath, R. 2016. Statistical rethinking: A Bayesian course with examples in R and Stan. CRC Press, Taylor & Francis Group, Boca Raton, FL.

McGrayne, S.B. 2012. The Theory that would not die: How Bayes' rule cracked the Enigma code, hunted down Russian submarines, and emerged triumphant from two centuries of controversy. Yale University Press, New Haven, Connecticut.

Midway, S.R., T. Wagner, S.A. Arnott, P. Biondo, F. Martinez-Andrade, and T.F. Wadsworth. 2015. Spatial and temporal variability in growth of southern flounder (*Paralichthys lethostigma*). 167:323-332.

Millar, R.B. 2002. Reference priors for Bayesian fisheries models. Canadian Journal of Fisheries and Aquatic Sciences 59:1492-1502.

Nalepa, T.F., D.L. Fanslow, and G.A. Lang. 2009. Transformation of the offshore benthic community in Lake Michigan: recent shift from the native amphipod *Diporeia* spp. To the invasive mussel *Dreissena rostriformis bugensis*. Freshwater Biology 54:466-479.

Neumann, R.M., C.S. Guy, and D.W. Willis. 2012. Length, weight, and associated indices. Pages 637-676 *in* A.V. Zale, D.L. Parrish, and D.W. Willis, editors. Fisheries techniques, 3rd edition. American Fisheries Society, Bethesda, Maryland.

Ogle, D.H., T.O. Brenden, and J.L. McCormick. 2017. Growth estimation: Growth Models and Statistical Inference. Pages 265-353 *in* Quist, M.C. and D. Isermann, editors. Age and Growth of Fishes: Principles and Techniques. American Fisheries Society. Bethesda, Maryland.

Parent, É, and É. Rivot. 2013. Introduction to hierarchical Bayesian modeling for ecological data. Chapman & Hall/ CRC Press. Boca Raton, FL.

Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. *In* Proceedings of the 3rd International Workshop on Distributed Statistical Computing (DSC 2003), 20-23 March, Vienna, Austria. ISSN 1609-395X

Punt, A.E., and R. Hilborn. 1997. Fisheries stock assessment and decision analysis: they Bayesian approach. Reviews in Fish Biology and Fisheries 1997(7):35-63.

Quist, M.C., J.L. Stephen, S.T. Lynott, J.M. Goeckler, and R.D. Schultz. 2010. Exploitation of walleye in a Great Plains reservoir: harvest patterns and management scenarios. Fisheries Management and Ecology 17:522-531.

Rahikainen, M., I. Helle, P. Haapasaari, S. Oinonen, S. Kuikka, J. Vanhatalo, S. Mäntyniemi, and K. Hoviniemi. 2014. Towards integrative management advice of water quality, oil spills, and fishery in the Gulf of Finland: A Bayesian approach. AMIBO 43:115-123.

Reed, J.R., and W.D. Davies. 1991. Population dynamics of black crappies and white crappies in Weiss Reservoir, Alabama: Implications for the implementation of harvest restrictions. North American Journal of Fisheries Management. 11(4):598-603.

Romakkaniemi, A. (Ed.). 2015. Best practices for the provision of prior information for Bayesian stock assessment. ICES Cooperative Research Report NO. 328. 93pp.

Royle, J.A., and M. Kéry. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 8(77):1813-1823.

Rubin, D.B. 1988. Using the SIR algorithm to simulate posterior distributions. Pages 395-402 *in* Bernardo, J.M., M.H. DeGroot, D.V. Lindley, and A.F.M Smith, editors. Bayesian Statistics 3rd edition. Oxford University Press, Cambridge, MA.

Tsehaye, I., B.M. Roth, and G.G. Sass. 2016. Exploring optimal Walleye exploitation rates for northern Wisconsin ceded territory lakes using a hierarchical Bayesian age-structured model. Canadian Journal of Fisheries and Aquatic Sciences 73:1413-1433.

Trafimow, D., and M. Marks. 2015. Editorial. Basic and Applied Social Psychology 37:1-2.

Walters, C.J., and S.J.D. Martell. 2004. Fisheries Ecology and Management. Princeton University Press, Princeton, NJ.

Wasserstein, R.L., and N.A. Lazar. 2016. The ASA's statement on p-vales: context, process, and purpose. The American Statistician 70(2):129-133.