Technical Documentation

Chinook Salmon In-Season Bayesian Risk Assessment Tool

Contents

This document is intended to be a companion to the Chinook Salmon In-Season Bayesian Risk Assessment Tool (hereafter, "Bayes' Tool" or simply "the Tool") that describes the motivations for its development and the technicalities of what it does. While this document highlights some of the features of the Tool, it does not provide insights about how to use it. It is the hope of the developers that the Tool is intuitive to use on its own after being introduced to it, but to address any questions or confusion, detailed instructions are documented in the user manual.

The Tool was developed for Kuskokwim River Chinook salmon only. Hereafter, all references in this document to salmon, fish, runs, escapement, and harvest are about the drainage-wide stock of Chinook salmon in the Kuskokwim River.

1   Motivations and Philosophy

The Tool was designed to address several difficulties with in-season run size assessment and decision-making in the Kuskokwim River Chinook salmon fishery. The primary in-season assessment information in the Kuskokwim River comes from the Bethel Test Fishery (BTF, Bue and Lipka 2016). The BTF provides daily catch-per-unit-effort (CPUE) data that provide an indication of the daily in-river run abundance and species composition. When historical years are considered in aggregate, the data provide a rich characterization of the variability in run timing as well. A strong temptation exists to look at the cumulative CPUE (hereafter CCPUE) on a given day and interpret it as an indicator of run size by comparing it to previous years' values on the same date, or to compare the accumulation time series of CPUE to previous years. The notion would be that by doing this one could ascertain if the run is large or small and set harvest targets based on that inference. However, this interpretation (without a model to appropriately express the uncertainty) is subject to failure because of two primary problems:

1. Because of **variability in run timing**, two runs can be of largely different sizes yet show similar CCPUE on the same date. That is to say, early/small and large/late runs provide the indication of an average-sized run with average timing early in the season. Without knowing with a high degree of certainty one of either run size or timing, it is difficult to ascertain the other with any confidence.
2. Because of **sampling variability** at the daily and annual time scale, the same end-of-season CCPUE counts have been observed in years with quite different run sizes.

The acknowledgement of these two issues is not to say the BTF data are not useful, but rather an accounting for them may help ensure BTF data are interpreted with the appropriate level of confidence.

The other assessment information used in the Kuskokwim River Chinook salmon fishery (and many other salmon fisheries) is the pre-season run size forecast. It is the view of

the Tool Developers that the BTF data should be used to *update* rather than *replace* the pre-season forecast in-season. Early in the season, when the BTF data provide neither accurate or confident run size estimates, the pre-season information should still have the most influence on the in-season perception of run size and any resulting management decisions. Later in the season, when BTF data provide a more accurate and confident run size estimate, they can start to overwhelm the pre-season information by either reinforcing what was previously thought or changing it. This philosophy has been applied to other salmon stocks, namely for Bristol Bay sockeye salmon (Fried and Hilborn 1988; Hyun et al. 2005). Two more considerations come to mind when the BTF data are thought of in this way:

3. How should the BTF data be used to update the pre-season forecast (i.e., by what statistical method)?
4. How should the the updated information be used to inform responsive management strategies?

The primary motivations of the Tool are to address items 1 – 3 in a statistically and scientifically defensible way that is accessible to a wide audience, and to present the information in such a way as to be amenable to meeting the goal laid out in item 4.

2 Bayes' Theorem

As an answer to item 3 above, a rational way to perform the updating of pre-season forecasts is to use *Bayesian Inference*. At its core, Bayesian methods say that previous information should not be totally discarded when new information is available, but that new information should update the old with the new. Bayesian Inference relies on a theorem derived from the laws of probability called *Bayes' Theorem*. Bayes' Theorem defines how our beliefs of the outcomes of uncertain events should rationally change as new information becomes available. In this framework, different outcomes are termed *hypotheses* and beliefs that any one hypothesis (or any subset of hypotheses) is true are expressed as probabilities. In the context of salmon run size updating, different run sizes for the current year represent different competing hypotheses and new information is composed of the current BTF CCPUE data. Bayes' Theorem takes the following form:

$$\Pr(H_i \mid data) = \frac{\Pr(data \mid H_i)\Pr(H_i)}{\sum_{j=1}^{J}\Pr(data \mid H_j)\Pr(H_j)} \tag{1}$$

Here, $\Pr(H_i \mid data)$ is known as the *posterior* probability of each hypothesis i, which is the updated understanding of how likely it is that hypothesis H_i is true after considering the new information contained in *data*. It is expressed as a conditional probability, and is read as "the probability of H_i **given** we know the data were observed." The posterior probability

is obtained from Bayes' Theorem, and is proportional to the *likelihood* of observing the *data* if we assumed hypothesis H_i was true times the *prior* probability that hypothesis H_i was true before considering the *data*. The denominator in Bayes' Theorem is needed so that the posterior probabilities sum to one – it serves the purpose of being a normalizing constant. The denominator is often written as $\Pr(data)$, that is, the probability of observing the data at all, integrated across all possible hypotheses.

On each day of the run, the prior probability of each run size $\Pr(H_i)$ is always the pre-season forecast, which has uncertainty about which run size will be the true run this year that can be expressed probabilistically. That is to say, using the forecast, we can assign probabilities to the hypothesis that the run will be less than, say, 100,000 fish. This is somewhat contrary to the way pre-season forecasts for Kuskokwim River Chinook salmon have been presented in the past, but the developers believe that there *is* information in the forecast to suggest that some outcomes within the range are more likely than others. This claim is based on the performance of the *current forecasting method* (which uses last year's run size as the anticipated run for this year): the midpoint of the forecast range makes smaller mistakes more often than it makes larger mistakes, thereby providing justification for placing more faith in run sizes near the center of the range than at the ends of the range.

After some BTF data have been collected (*data*), we can calculate the likelihood of observing those data if the run was truly less than 100,000 fish. Using Bayes' Theorem, we can calculate the posterior probability that the run is less than 100,000 fish after observing the *data*, thus providing an updated run estimate off of which to base management decisions about how much harvestable surplus there is. This process and the accompanying calculations are worked out in an example below.

2.1 Bayes' Theorem Example

Consider the simple (and highly hypothetical) case in which managers are interested in only three hypotheses regarding the run size of the current year:

- H_{small}: less than 100,000 fish
- H_{medium}: greater than 100,000 fish but less than 150,000 fish
- H_{large}: greater than 150,000 fish but less than 220,000 fish.

We realize that these categories are not appropriate when taken in context of all historical run sizes that have occurred in the Kuskokwim River [1976 – 1976 average: 213,000; range: 75,000 – 399,000; Larson (2022)], but this example is for illustration purposes only, so please forgive us for this simplification for the time being. Using the pre-season forecast, managers assign the following prior probabilities:

TABLE 1. Hypothetical prior probabilities for three run size hypotheses.

Run Hypothesis	Prior (Forecast) Probability
H_{small}	0.3
H_{medium}	0.6
H_{large}	0.1

Now suppose the managers have access to the historical run sizes and the corresponding BTF CCPUE on each day of each year since 2008. They divide each of the historical run sizes into the three size categories as defined by the three hypotheses. The specific categories do not matter for this example, but managers divide the BTF CCPUE observed on June 12th into small, medium, and large categories as well. They then create a table that shows how many years the BTF suggested the run was in the correct or incorrect category:

TABLE 2. One way of calculating the frequency with which the BTF correctly and incorrectly indexed the run size on June 12th in previous years since 2008.

| True Run Size | BTF Suggested Run Size | | | |
	Small	Medium	Large	Total
Small	2	0	1	3
Medium	2	3	3	8
Large	1	1	1	3
Total	5	4	5	14

The cases where the BTF correctly indexed the run as of June 12th are those where the row and column run size is the same: e.g., the BTF correctly indexed 2 of the 3 truly small runs as being small, 3 of 8 runs as being medium, but 1 of the 3 runs as being large. All other cells represent an incorrect assignment. Overall, the BTF correctly assigned 43% of the years, and incorrectly assigned 57% of the years. This suggests that the BTF is an unreliable index of run size on June 12th, so managers decide to wait a week to interpret the BTF data as an index of run size.

On June 19th, managers create the same table but using the BTF CCPUE data through June 19th instead:

TABLE 3. One way of calculating the frequency with which the BTF correctly and incorrectly indexed the run size on June 19$^{\text{th}}$ in previous years since 2008.

	BTF Suggested Run Size			
True Run Size	**Small**	**Medium**	**Large**	**Total**
Small	2	0	1	3
Medium	2	4	2	8
Large	1	0	2	3
Total	5	4	5	14

In this case, the BTF has had a 57% correctness rate. For medium runs in particular, the BTF assigned 4 runs out of 8 runs correctly, but assigned 4 incorrectly, and 2 of these runs were assigned to the large category.

This information can be used to formulate the likelihood part of Bayes' Theorem: $\Pr(data \mid H_i)$. Suppose that as of June 19$^{\text{th}}$, the BTF has suggested a small run. We can calculate the likelihood of observing a small BTF CCPUE given each run size is the true one using the information in Table 3:

- When the run was truly small, 2 of 3 runs were correctly assigned as small. Thus, $\Pr(BTF_{\text{small}} \mid H_{\text{small}}) = 2/3 = 0.67$.
- When the run was truly medium, 2 of 8 runs were incorrectly assigned as small. Thus, $\Pr(BTF_{\text{small}} \mid H_{\text{medium}}) = 2/8 = 0.25$.
- When the run was truly large, 1 of 3 runs were incorrectly assigned as small. Thus, $\Pr(BTF_{\text{small}} \mid H_{\text{large}}) = 1/3 = 0.33$.

These are the conditional probabilities of observing small BTF CCPUE data in years that were truly of each size. To obtain the posterior probabilities: $\Pr(H_i \mid data)$, we multiply each of these probabilities by the corresponding prior probability from Table 1, and divide each by their sum. This is simply equation (1) in words. First, calculate the three numerators:

- $\Pr(BTF_{\text{small}} \mid H_{\text{small}}) \Pr(H_{\text{small}}) = 0.67 \times 0.3 = 0.2$.
- $\Pr(BTF_{\text{small}} \mid H_{\text{medium}}) \Pr(H_{\text{medium}}) = 0.25 \times 0.6 = 0.15$.
- $\Pr(BTF_{\text{small}} \mid H_{\text{large}}) \Pr(H_{\text{large}}) = 0.33 \times 0.1 = 0.03$.

To standardize these probabilities, we divide each by the sum of all three: 0.2 + 0.15 + 0.03 = 0.38. Thus, the updated probability of a small run now that we have observed that the BTF suggests it will be small is 0.2/0.38 = 0.53, which represents nearly double the chance of a small run of the prior (0.3). The updated probability for a medium run is 0.15/0.38 = 0.39 and for a large run it is 0.03/0.38 = 0.08.

Table 4 below displays the results of these calculations for each of the three possible BTF

observation types. For example, the first three rows show the results of the calculations when a small BTF value is obtained, which were illustrated in the above text.

TABLE 4. Hypothetical output of Bayesian calculations under each of the three possible BTF outcomes on June 19th.

Run Hypothesis	Prior	Likelihood	Posterior
BTF = Small			
Small	0.3	0.67	0.52
Medium	0.6	0.25	0.39
Large	0.1	0.33	0.09
BTF = Medium			
Small	0.3	0.00	0.00
Medium	0.6	0.50	1.00
Large	0.1	0.00	0.00
BTF = Large			
Small	0.3	0.33	0.31
Medium	0.6	0.25	0.47
Large	0.1	0.67	0.21

2.2   Problems with this Example

This is illustrative example is just that: an illustration. Although the BTF and run size data used are real (and exactly those used by the Tool) and the Bayesian calculations are correct, this example suffers from two primary downfalls that prevent it from being the way the Tool actually works. First, the boundaries of the categories (e.g., small runs are less than 100,000) are totally arbitrary. In reality, run size is not categorized this way, but instead estimates of numbers of fish are desired for management purposes. That is to say, we need to consider *many* different hypotheses, each one representing a unique run size (e.g., that the total run will be 100,000 fish, 101,000 fish, etc. and everywhere in between). The second issue is that these calculations only address the historical reliability of the BTF as an index of run size, which is not necessarily a good sample of how it will perform in any given year.

3   Full Description of Run Estimation

3.1   Overview

This section describes how historical run size and BTF data are used to create a continuous version of the example presented above. Staton and Catalano (2019) performed an evaluation of this method for estimating run size in-season based on incomplete index information, and found that Bayesian updating provides more reliable run size estimates than the pre-season forecast starting on approximately June 17th, and that both the Bayesian update and the pre-season forecast were more reliable than estimates from the BTF data alone starting on June 10th.

The calculations presented here rely heavily on the tool of *Monte Carlo simulation* (hereafter abbreviated MC) to propagate uncertainty from one step of the analysis to the next. MC methods sample random values of a quantity based on the known frequency of these quantities given by some statistical distribution (Bolker 2008).

NOTE: *much of the text in this section has been duplicated from Staton and Catalano (2019), though only the portions relevant to calculations conducted by the Tool have been included.*

3.2   Prior Formulation

The pre-season forecast is used to formulate the prior probability distribution to be updated with new information each day. The current forecasting method for the Kuskokwim River assumes the run in the current year will be similar in size to the run that occurred last year. Thus, the forecast states that the expectation (in the absence of any in-season information) is that the run will be the same size as last year, but has some uncertainty. We express this uncertainty using a bias-corrected log-normal distribution. The log-normal distribution is used because it (*a*) has a domain over only positive real numbers and (*b*) has the property of multiplicative variability (i.e., variation in terms of fish is smaller at anticipated small run sizes than at large run sizes). The bias-correction is used so that the expected value of the random variable N is the specified mean parameter N_{t-1}. MC samples from the prior distribution are made as:

$$\log(N_b) \sim \text{Normal}\left(\log(N_{t-1}) - \frac{\hat{\sigma}_F^2}{2}, \hat{\sigma}_F\right),\tag{2}$$

where N_b is an MC sample of the anticipated run size in the current year (of which 1×10^6 were drawn), N_{t-1} is the estimated run size from the previous year, and $\hat{\sigma}_F$ is the standard deviation of the forecast, representing the uncertainty in it. This $\hat{\sigma}_F$ is calculated by applying the forecasting rule to all previous years and calculating the log-scale errors

that the forecast would have made in the past had the rule been applied. That is, for all historical years we calculate the errors as:

$$\varepsilon_{F,t} = \log\left(\frac{N_{t-1}}{N_t}\right), \tag{3}$$

and summarize their variability:

$$\hat{\sigma}_F = \sqrt{\frac{(\varepsilon_{F,t} - \bar{\varepsilon}_F)^2}{n_t - 1}}, \tag{4}$$

where $\bar{\varepsilon}_F$ is the mean of all $\varepsilon_{F,t}$ and n_t is the number of historical years for which $\varepsilon_{F,t}$ can be calculated. Note that when uncertainty is provided as a coefficient of variation (CV), σ can be obtained as:

$$\sigma = \sqrt{\log(\text{CV}^2 + 1)} \tag{5}$$

Staton and Catalano (2019) performed this analysis with run size data current from 1976 – 2017 and found that $\hat{\sigma}_F \approx 0.28$; Figure 1 shows the historical run time series (panel *a*), the time series of historical forecast errors had the method been used starting in 1977 (panel *b*), and the distribution of these errors (panel *c*).
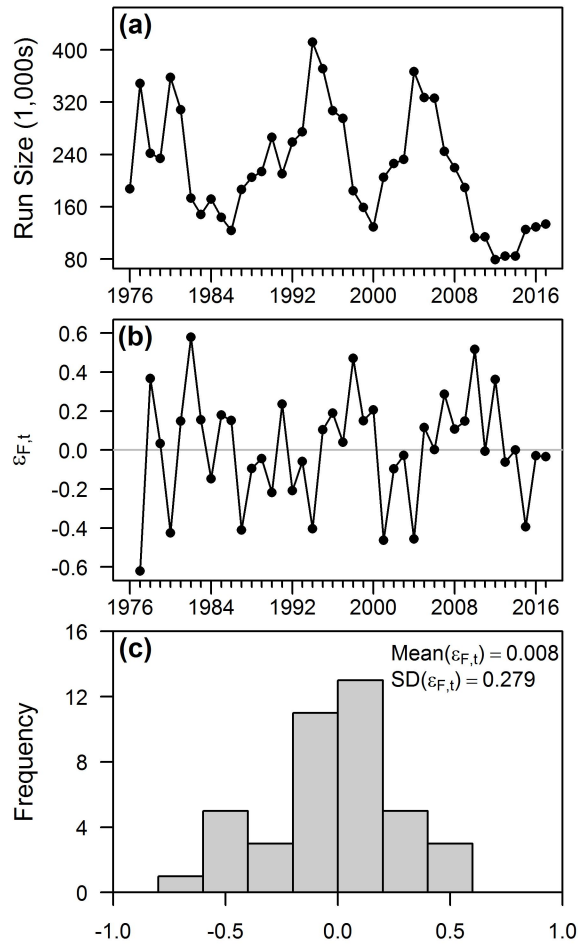
FIGURE 1. (*a*) Estimated run size time series from 1976-2017, (*b*) time series of log scale multiplicative pre-season forecast errors $\varepsilon_{F,t}$, and (*c*) distribution of the $\varepsilon_{F,t}$ values. Duplicated from Staton and Catalano (2019)

.

In some applications of Bayesian Inference, it is possible to use yesterday's posterior as today's prior. In our case this would be statistical folly, given the likelihood from each day contains much of the same information from previous days (the CCPUE). This sort of "double-dipping" is prohibited in Bayesian analyses, and would result in spuriously confident posterior distributions.

3.3 Likelihood Formulation

3.3.1 Historical Relationship

Information about run size is contained in the cumulative BTF CPUE values observed each day of the run $\left(\text{CCPUE}_d = \sum_{j=1}^{d} \text{CPUE}_j\right)$, and thus these data form the foundation for linking in-season abundance index data to different run size hypotheses in a likelihood

framework. The total end-of-season abundance each year (N_t) is related to $\text{CCPUE}_{d,t}$ at various points in the season using multiple linear regression:

$$\log(N_t) = \beta_0 + \beta_{1,d}q_t + \beta_{2,d}\text{CCPUE}_{d,t} + \beta_{3,d}\text{CCPUE}_{d,t}q_t + \varepsilon_{d,t} \tag{6}$$

where $\beta_{j,d}$ are coefficients explaining the relationship on each day, q_t is a binary indicator based on the BTF catchability period, and $\varepsilon_{d,t}$ are normally distributed errors with mean zero and variance equal to σ_d^2. The catchability period term is included because the efficiency of the BTF gear increased substantially beginning in 2008 as a result of a change in net-makers (Bue and Lipka 2016). Regression relationships like those in eq. (6) are commonly used to estimate run size based on in-season index data (Fried and Hilborn 1988; Flynn and Hilborn 2004; Hyun et al. 2005; Michielsens and Cave 2018). This relationship can be fitted to historical data on any day d of the run to obtain estimates of the coefficients ($\hat{\beta}_{j,d}$), their uncertainty in the form of a covariance matrix ($\hat{\Sigma}_d$), and the residual variability ($\hat{\sigma}_d$).

Figure 2 shows these relationships at several points in the run. Note that since all future years are in the post-2008 catchability period, the fitted lines corresponding to the pre-2008 will never be used to inform abundance predictions based on CCPUE, but these data are included in the fitted model to allow them to inform the amount of variability in run size one might expect at a given value of CCPUE_d in some future year.
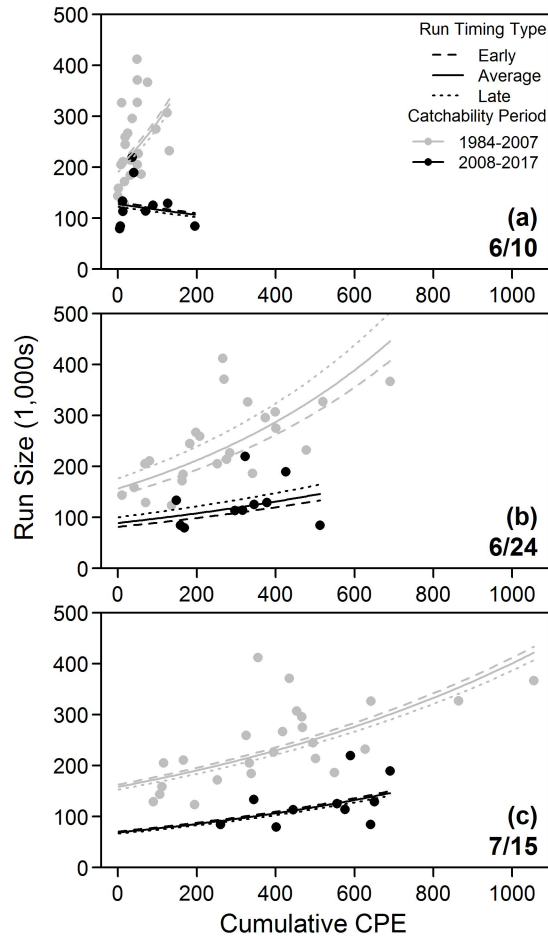
FIGURE 2. The fitted regression relationships from eq. 6 at three points in the season (solid black/grey lines). Duplicated from Staton and Catalano (2019), who investigated the influence of run timing variability on these relationships as well. Different run timing scenarios are shown: dashed lines represent the prediction at a given CCPUE$_d$ for the earliest (in the season) run on record, dotted lines correspond to the latest run on record. Given the fact that these other run timing scenarios explain little variability (relative closeness of the dotted/dashed lines to the solid lines), Staton and Catalano (2019) concluded that run timing could be ignored when using these relationships.

3.3.2   Uncertainty Estimation

The relationships shown in Figure 2 contain much uncertainty about the linkage between N and CCPUE$_d$. There are two primary estimated sources of uncertainty:

1. Uncertainty in the estimated coefficients ($\hat{\beta}_{j,d}$), stored in a covariance matrix ($\hat{\Sigma}_d$). This source of uncertainty is in regard to what the best fitted lines for the data are, and is used to recognize that there are other lines that could possibly explain the data.
2. Variability in N at a given level of CCPUE$_d$, stored as residual variance ($\hat{\sigma}_d^2$).

MC methods are then used to propagate this uncertainty into predictions of run size ($\dot{N}$)

from $CCPUE_d$. First, 1×10^6 MC samples (indexed by b) are drawn from a multivariate normal distribution:

$$(\hat{\beta}_{0,d,b}, \hat{\beta}_{1,d,b}, \hat{\beta}_{2,d,b}, \hat{\beta}_{3,d,b}) \sim \text{MVN}\left((\hat{\beta}_{0,d}, \hat{\beta}_{1,d}, \hat{\beta}_{2,d}, \hat{\beta}_{3,d}), \hat{\Sigma}_d\right), \tag{7}$$

Then, 1×10^6 MC samples of random residuals are obtained (bias-correction included to ensure the mean of $e^{\hat{\varepsilon}_{d,b}} = 1$):

$$\hat{\varepsilon}_{d,b} \sim \text{Normal}\left(-\frac{\hat{\sigma}_\varepsilon^2}{2}, \hat{\sigma}_\varepsilon\right) \tag{8}$$

Finally, these MC samples are combined in a predictive version of the fitted regression model in eq. (6):

$$\log(\dot{N}_{d,b}) = \hat{\beta}_{0,d,b} + \hat{\beta}_{1,d,b}q + \hat{\beta}_{2,d,b}CCPUE_d + \hat{\beta}_{3,d,b}CCPUE_dq + \hat{\varepsilon}_{d,b} \tag{9}$$

The set of MC samples $\dot{N}_{d,b}$ constitutes different plausible values of the true run size this year suggested by the currently observed BTF value $CCPUE_d$. Finally, this distribution is summarized using a one-dimensional kernel density estimator fitted to the 1×10^6 MC samples of $\dot{N}_{d,b}$ to obtain the probability density function (PDF) of the likelihood distribution. The resulting function is hereafter denoted by $\Pr(CCPUE_d \mid N_i)$, where N_i represents a specific continuous run size hypothesis.

3.4 Posterior Formulation

To obtain Bayesian updates of the perceived run size, the log-normal distribution representing uncertainty in the pre-season forecast is used as the prior information each day [denoted $\Pr(N_i)$]. Although this is a simple one parameter Bayesian estimation problem, the likelihood PDF [$\Pr(CCPUE_d \mid N_i)$] does not have a well-defined parametric form which could allow direct analytical calculation of the posterior PDF [$\Pr(N_i \mid CCPUE_d)$] (this is not an uncommon situation). Instead, a custom random walk Metropolis-Hastings Markov Chain Monte Carlo (MCMC) algorithm (Chib and Greenberg 1995) is used. A log-normal proposal distribution is used by the algorithm as opposed to a symmetrical distribution (like the normal distribution) to prevent negative proposals. The standard deviation of the proposal distribution can be tuned such that the acceptance rate of proposals falls between 0.2 and 0.4 (Bédard 2007).

In the Metropolis-Hastings algorithm, random samples are drawn from the posterior. This is conducted by generating a time series of random values by either "jumping" to a proposal value or "staying" at the current value. The probability of staying or jumping is defined by the ratio of the posterior probability of each value of N: $\frac{\Pr(proposed)}{\Pr(current)}$, where

Pr(*proposed*) and Pr(*current*) are calculated using the numerator of Bayes' Theorem in eq. (1). If the proposal has higher posterior probability than the current value, the algorithm will always jump to it and it will be the current value for the next decision in the time series. However, if the proposed value has lower posterior probability, the action to jump will be made with probability equal to the ratio. By carrying out a long time series of such decisions, the values retained as the current value are samples drawn in proportion to their posterior probability.

3.4.1 Posterior Convergence Diagnostics

MCMC is commonly used to estimate a posterior distribution, but its behavior must be assessed to ensure the algorithm has converged to a stable target distribution. Posterior convergence can be assessed in the Tool using two methods found in the "Diagnostic" subtab of the "Estimation" tab after performing the calculations. First, visuals of the MCMC chains and the resulting densities are displayed. Figure 3 displays an example of what these figures should look like if the algorithm has converged. Second, the Potential Scale Reduction Factor (Brooks and Gelman 1998) is used as a quantitative measure of convergence (displayed as "BGR" in the table of Figure 3; values less than 1.1 are indicative of convergence) and the MCMC effective sample size is calculated to ensure enough independent samples have been drawn for adequate inference [values greater than 3,000 – 5,000 are generally considered adequate; Raftery and Lewis (1992)].

Density Plot **Trace Plot**

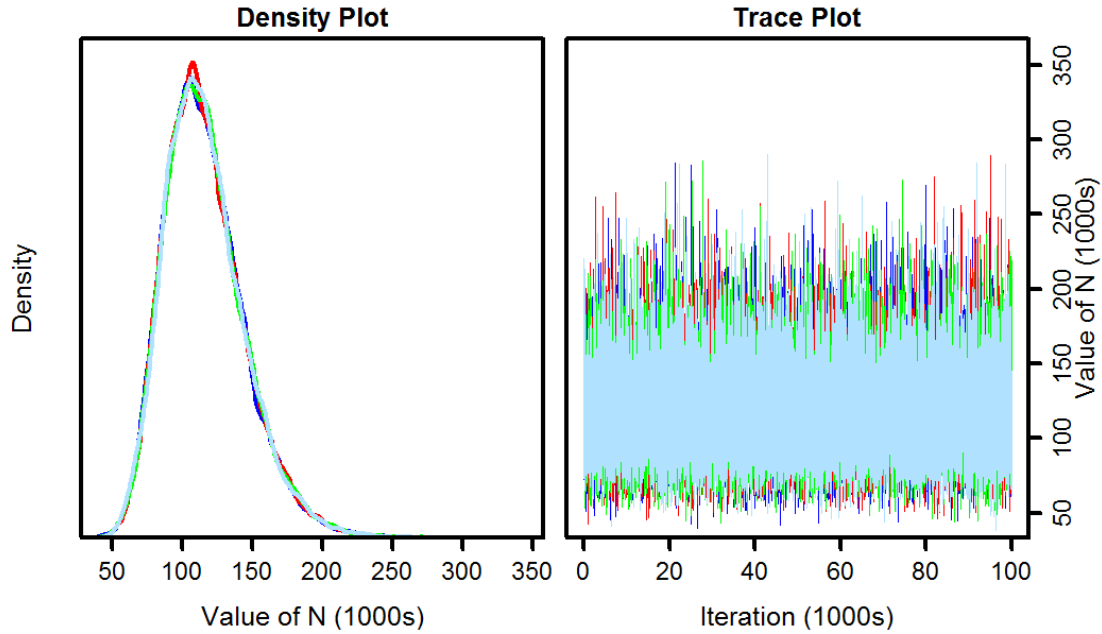| Saved Samples	Effective Samples	BGR	Acceptance Rate
400,000	83,313	1.00	0.36

FIGURE 3. MCMC diagnostics as displayed in the Tool. In the plot on the left, users should look for tight overlap of the four distributions. In the plot on the right, users should look for random jumping patterns that overlap largely for the four colors. These are all indicative of very good convergence of the algorithm.

4   Risk Analysis Calculations

Samples from the three PDFs (prior, likelihood, and posterior) on their own are of little utility to management decisions. Instead, managers might wish to know how different management actions (i.e., different season-wide harvest targets) will affect the likelihood of the occurrence of desirable and undesirable escapement outcomes. This constitutes a risk analysis and the steps we have outlined so far have all been to allow us to perform the calculations that follow.

4.1   Cumulative Probability

Managers may be interested in a quantity Y which represents the probability that drainage-wide escapement (hereafter denoted S) will fall below a limit escapement point L if a target harvest T is taken this season. For this calculation, we will need to introduce the *empirical cumulative density function*, hereafter abbreviated as "CDF." The CDF is the fraction of a random sample that have values less than or equal to some number of interest. If we ignore harvest for the moment, this fraction can be thought of as a probability that the run

size this year will be smaller than some number. For ease of formal presentation, let us define a binary indicator function ψ:

$$\psi(z; x_i) = \begin{cases} 1 & \text{if } x_i \leq z \\ 0 & \text{if } x_i > z \end{cases}$$

x_i is an element in a vector $\mathbf{x}$ of random samples from one of the three PDFs and z is a threshold of interest. We now define the CDF function Φ:

$$\Phi(z; \mathbf{x}) = \frac{\sum_{i=1}^{n} \psi(z; x_i)}{n}, \tag{10}$$

where n is the number of elements in the vector $\mathbf{x}$. If we perform $\Phi(100,000; \mathbf{N})$ and obtain the result 0.3, that indicates there is a 30% chance that the run will be less than or equal to 100,000 fish, *conditional on the information used and assumptions made* to obtain $\mathbf{N}$. Conversely, we can perform $1 - \Phi(100,000; \mathbf{N})$ to obtain the probability that the run will be greater than 100,000 fish.

Now to use the CDF function Φ for risk analysis. Suppose managers are considering harvesting T fish for the season, and wish to know the probability that total escapement (S) will be less than or equal to the limit (L) if *exactly* T fish are harvested this season. We can perform this calculation using $\Phi(L + T; \mathbf{N})$. As an explanation, consider the following *example* (emphasis on example). Managers wish to maintain $S > 65,000$ and are considering allowing a total target harvest of $T = 40,000$ fish for the season. The probability of $S \leq 65,000$ if $T = 40,000$ is the same as the probability that the total run is less than or equal to 65,000 + 40,000.

If U fish have already been harvested (without uncertainty) the calculation for a proposition of harvesting V *additional* fish would be $\Phi(L + V; \mathbf{N} - U)$. If there is uncertainty in U, then the calculation can proceed as $\Phi(L + V; \mathbf{N} - \mathbf{U})$ where $\mathbf{U}$ is a vector of MC samples reflecting the expected harvest already taken as well as the understood uncertainty in it.

If we wish to know the probability that S will be *between* two values A and B if T fish are harvested, where $B > A$, then we can calculate $\Phi(B + T; \mathbf{N}) - \Phi(A + T; \mathbf{N})$ to obtain this probability. If U fish have already been harvested, the calculation would be $\Phi(B + T; \mathbf{N} - \mathbf{U}) - \Phi(A + T; \mathbf{N} - \mathbf{U})$. This calculation may be desired if managers wish to know the probability that S will fall within the escapement goal (or some other) range.

4.2 The Notion of P^*

One approach to using the Tool is to set a critical escapement limit L such that managers would deem the event $S \leq L$ undesirable. Managers can then select a maximum probability of the occurrence of the undesirable event $S \leq L$ they are willing to live with. In the harvest

control systems literature (e.g., Shertzer et al. 2010; Prager et al. 2003), this maximum probability is known as P^* and is an expression of maximum acceptable risk (i.e., risk tolerance). The option in the Tool entitled "Risk Analysis: Choose Harvest Target" uses the P^* notion to pick the maximum allowable harvest that would result in $\Pr(S \leq L) < P^*$. Note that it is entirely possible to obtain the same suggested target harvest using two different levels of P^* and L.

4.3   An Alternative

Some users of the Tool may feel that the rather automated nature of the P^* notion is too mechanical and following it rigorously doesn't leave room for flexibility. These users would be perfectly justified in making such a claim, and thus an additional option is included in the Tool under the "Risk Analysis: Compare Harvest Targets". In this tab, users can propose harvest targets and see what the probabilities of different escapement outcomes would be, and pick the harvest target that is most consistent with their own levels of risk tolerance.

4.4   Concluding Remarks

Some users of the Tool may state that their P^* is zero, that is, they are not willing to accept any level of risk whatsoever that $S \leq L$. If this is true, the management advice would be to rarely catch even a single a single fish, because there is generally some non-zero probability that $N \leq L$, even if it is essentially zero. For this reason, we encourage users to think of these probabilities as long-run frequencies of outcomes. If the user selects $P^* = 0.05$, they are essentially saying that in 5 out of 100 (or 1 in 20) years they are willing to accept the outcome of $S \leq L$ in order to allow some level of harvest.

Furthermore, it is worth noting that the risk assessment framework is not designed to target an exact escapement level. Rather, it is intended to provide some assurance that an undesriable outcome will not happen. The result of this is that P^* and L should be selected to reflect very low acceptable risk of a very bad outcome, and not about a high chance of obtaining the exact desired outcome.

5 Computation

We used Program R (R Core Team 2018) to perform all calculations used in the Tool. R is a coding platform for statistical computation, and is widely used among scientists across many disciplines. The Tool was developed in R using a package called `{shiny}` (Chang et al. 2017), which allows R code to be integrated into a web interface. This document was produced using the packages `{rmarkdown}` (Allaire et al. 2018), `{bookdown}` (Xie 2018), and `{knitr}` (Xie 2014) with typesetting done via LaTeX.

References

Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., Wickham, H., Cheng, J., and Chang, W. 2018. *rmarkdown: Dynamic Documents for R*. R package version 1.9.

Bédard, M. 2007. Weak convergence of Metropolis algorithms for non-i.i.d. target distributions. *The Annals of Applied Probability*, 17(4):1222–1244.

Bolker, B. M. 2008. *Ecological Models and Data in R*. Princeton University Press. ISBN 978-0691125220.

Brooks, S. P. and Gelman, A. 1998. General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7(4):434.

Bue, D. G. and Lipka, C. G. 2016. Characterization of the 2011 salmon run in the Kuskokwim River based on the test fishery at Bethel. Fishery Data Series 16-05, Alaska Department of Fish and Game, Anchorage, AK. Available at: http://www.adfg.alaska.gov/FedAidPDFs/FDS16-05.pdf [last accessed 2/20/2019].

Chang, W., Cheng, J., Allaire, J., Xie, Y., and McPherson, J. 2017. *shiny: Web Application Framework for R*. R package version 1.0.5.

Chib, S. and Greenberg, E. 1995. Understanding the Metropolis-Hastings algorithm. *The American Statistician*, 49(4):327.

Flynn, L. and Hilborn, R. 2004. Test fishery indices for sockeye salmon (*Oncorhynchus nerka*) as affected by age composition and environmental variables. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(1):80–92.

Fried, S. M. and Hilborn, R. 1988. Inseason forecasting of Bristol Bay, Alaska, sockeye salmon (*Oncorhynchus nerka*) abundance using Bayesian probability theory. *Canadian Journal of Fisheries and Aquatic Sciences*, 45(5):850–855.

Hyun, S.-Y., Hilborn, R., Anderson, J. J., and Ernst, B. 2005. A statistical model for in-season forecasts of sockeye salmon (oncorhynchus nerka) returns to the bristol bay districts of alaska. *Canadian Journal of Fisheries and Aquatic Sciences*, 62(7):1665–1680.

Larson, S. 2022. 2021 Kuskokwim River Chinook salmon run reconstruction and 2022 forecast. Regional Information Report 3A.22-0X, Alaska Department of Fish and Game, Anchorage, AK. Available at: https://www.adfg.alaska.gov/FedAidPDFs/RIR.3A.2022.0X.pdf [last accessed 5/XX/2022].

Michielsens, C. G. and Cave, J. D. 2018. In-season assessment and management of salmon stocks using a Bayesian time-density model. *Canadian Journal of Fisheries and Aquatic Sciences*.

Prager, M. H., Porch, C. E., Shertzer, K. W., and Caddy, J. F. 2003. Targets and limits for mangement of fisheries: A simple probability-based approach. *North American Journal of Fisheries Management*, 23:349–361.

R Core Team 2018. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria.

Raftery, A. E. and Lewis, S. 1992. How many iterations in the Gibbs sampler? In *Bayesian Statistics 4*, pages 763–773. Oxford University Press.

Shertzer, K. W., Prager, M. H., and Williams, E. H. 2010. Probabilistic approaches to setting acceptable biological catch and annual catch targets for multiple years: Reconciling methodology with national standards guidelines. *Marine and Coastal Fisheries*, 2(1):451–458.

Staton, B. A. and Catalano, M. J. 2019. Bayesian information updating procedures for Pacific salmon run size indicators: Evaluation in the presence and absence of auxiliary migration timing information. *Canadian Journal of Fisheries and Aquatic Sciences*, 76(10):1719–1727.

Xie, Y. 2014. knitr: A comprehensive tool for reproducible research in R. In Stodden, V., Leisch, F., and Peng, R. D., editors, *Implementing Reproducible Computational Research*. Chapman and Hall/CRC. ISBN 978-1466561595.

Xie, Y. 2018. *bookdown: Authoring Books and Technical Documents with R Markdown*. R package version 0.7.