Canadian Journal of
Fisheries and
Aquatic Sciences

Article

Redd superimposition mediates the accuracy, precision, and significance of redd counts for cutthroat trout

Jeffrey R. Baldock[a], **Robert Al-Chokhachy**[b], **Timothy E. Walsworth**[c], **and Annika Walters**[d]

[a]Wyoming Cooperative Fish and Wildlife Research Unit, Department of Zoology and Physiology and Program in Ecology, University of Wyoming, Laramie, WY, USA; [b]U.S. Geological Survey, Northern Rocky Mountain Science Center, Bozeman, MT, USA; [c]Department of Watershed Sciences and The Ecology Center, Utah State University, Logan, UT, USA; [d]U.S. Geological Survey, Wyoming Cooperative Fish and Wildlife Research Unit, Department of Zoology and Physiology and Program in Ecology, University of Wyoming, Laramie, WY, USA

Corresponding author: **Jeffrey R. Baldock** (email: jbaldock@uwyo.edu)

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Abstract

Redd counts are commonly applied to estimate spawning population size for salmonids and allow for broad spatial and temporal coverage in monitoring efforts. However, the utility of redd counts may be compromised by observation error, particularly with respect to superimposition, where later arriving spawners construct redds overlapping existing redds. Here, we provide a mechanistic evaluation of the effects of superimposition on the error structure and biological significance of redd count data for Yellowstone cutthroat trout (*Oncorhynchus clarkii bouvieri*) spawning within tributaries to the Snake River, Wyoming. We used a Bayesian framework to parse observation error into distinct components and found low detection of redd clusters (i.e., areas of superimposition) was offset by overestimates of the number of redds per cluster, such that observed counts accurately reflected census redd abundance. However, a saturating relationship between redd counts and spawner abundance indicated that counts is best interpreted as effective reproductive effort rather than spawner abundance. Our results provide a mechanistic understanding of redd count data that can be used to assess their application and interpretation for monitoring.

Key words: trout, superimposition, redd counts, observation error, monitoring, Bayesian model

Introduction

Understanding population status and trends is critical for effectively managing and conserving species of recreational, commercial, or ecological value (e.g., Hilborn et al. 2020). Conservation targets for vulnerable and imperiled species are often based on trends in the reproductive portion of populations. For salmonid fishes, annual estimates of the abundance or density of spawning adults often serve as the basis for management and recovery plans (e.g., Alves et al. 2004; Haak et al. 2010). However, the implementation of a specific methodology is often based on logistical considerations, rather than statistical accuracy, precision, and the ability to meet assumptions (Parsons and Skalski 2010). Reduced accuracy and precision may limit the power of monitoring data to detect and describe important long-term trends (Dauwalter et al. 2009; Pregler et al. 2019). As a result, monitoring programs may lack the ability to inform corrective actions designed to prevent local extirpations (Ham and Pearsons 2000; Al-Chokhachy et al. 2009) and avoid critical thresholds (Stier et al. 2022).

Redd counts (i.e., counts of female spawner nests) are commonly applied to provide indices of population size and assess the drivers of population dynamics for salmonids (Beland 1996; Kovach et al. 2017). Redd counts are easy to perform and non-invasive, allowing for greater spatial and temporal coverage in monitoring than can be attained through traditional approaches such as mark-recapture (Chasco et al. 2014). The validity of redd counts as an index of population size relies on the assumption that observed counts are representative of true redd numbers. There is increasing recognition that this assumption is often not met in practice due to sampling error stemming from inter-observer variability, redd size, age, and density, among other sources (Dunham et al. 2001; Muhlfeld et al. 2006; Howell and Sankovich 2012). A second assumption is that redd counts accurately reflect population status. While redd counts may be correlated with spawner abundance (Hay 1984; Gallagher and Gallagher 2005), inter-annual variation in the population sex ratio may affect this interpretation (Dauble and Watson 1997). Therefore, the use of redd counts as a cost-effective monitoring tool may be limited due to observation error and unknown or variable biological relevance.

An additional challenge in effective redd count monitoring can arise through density dependent effects on spawning behavior. When spawning densities are high, competition for suitable spawning habitat is mediated through redd

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

superimposition, where later spawning females construct redds overlapping those constructed previously (Quinn 2005). Superimposition is assumed to reduce reproductive success via the destruction of previously laid embryos (Hayes 1987; Taniguchi et al. 2000). Redd superimposition is, therefore, one of the primary mechanisms regulating salmonid populations (i.e., density dependent reproductive success; Ricker 1954; Essington et al. 2000). Recent work suggests superimposition may have important effects on the accuracy and precision of redd count data (Murdoch et al. 2019; Auerbach and Fremier 2022). However, the precise effects of superimposition have been ignored as discriminating redds within areas of superimposition is challenging (Dunham et al. 2001; Muhlfeld et al. 2006).

The effect of superimposition on the error structure of redd count data represents a considerable knowledge gap in the use of redd counts as a monitoring tool. In this study, we provide a mechanistic evaluation of the error structure of redd count data and how redd counts can be interpreted with respect to population status for Yellowstone cutthroat trout (YCT, *Oncorhynchus clarkii bouvieri*) spawning in spring-fed tributaries to the Snake River, Wyoming. Our specific objectives were three-fold. First, we quantified the magnitude and drivers of discrete error components in redd count data and the extent of variation among observers, streams, and years. Second, we summarized the net accuracy and precision of redd count data at multiple spatial scales. Third, we explored the biological significance of redd counts in terms of spawning population abundance. We used a fully Bayesian framework to describe the error structure in redd count data, which allowed us to express our results in terms of probability and aids the interpretation of our results for applied fisheries management (Wagner et al. 2013). Our research provides insight into the value and limitations of redd counts as a monitoring tool for salmonid conservation and management.

Methods

Study system and design

Our study took place in the upper Snake River watershed of northwest Wyoming, USA, a priority area for YCT conservation given high river network connectivity, little genetic introgression by non-natives species, and physiologically suitable water temperatures (Haak and Williams 2012; Al-Chokhachy et al. 2018). YCT typically rear and overwinter in mainstem rivers before returning to natal tributaries in the spring and early summer to spawn (Homel et al. 2015). Since the 1960s, the Wyoming Game and Fish Department (WGFD) has used redd counts to monitor YCT spawning in spring-fed streams (those in which streamflow is dominated by groundwater inflow), where superimposition is common (Kiefling 1997).

We assessed the error structure of redd count data on two spring-fed streams (Fig. 1*a*; Table S1). Lower Bar BC (LBBC) drains into the Gros Ventre River ~0.5 km upstream of confluence with the Snake River. Upper Bar BC (UBBC) drains into the Snake River within Grand Teton National Park, ~17 km NNE of LBBC. Both streams are fed by high-volume, cold-water spring seeps, resulting in stable temperature and flow regimes (Table S1). We conducted our study over the course of the spawning period (early May–mid July) in 2019 and 2021. As our objective was to evaluate error structure over a range of redd densities reflective of those observed historically, we divided each stream into reaches. Thus, the unit of analysis in our models was the reach (except for the Redds per Cluster model, *see below*). However, we considered the net error in redd count data at both the reach and stream scales as WGFD redd count monitoring includes both index reaches and whole stream counts. We delineated approximately equidistant reach boundaries according to prominent geomorphic features (e.g., large shallow ponds and deep pools lacking spawning gravel) that separated primary spawning areas (Fig. 1*a*).

Data collection

Census redd surveys

Throughout the spawning period, we conducted detailed spawning ground surveys (i.e., census surveys) twice weekly to monitor the spatial and temporal distribution of redd construction and the condition of redds already constructed. We identified redds by the distinctive pot and tail-spill morphology and lack of periphyton caused by female digging activity (Crisp and Carling 1989; Gallagher and Gallagher 2005). We mapped all redds with a handheld GPS device and assigned a categorical age (Table 1). Categorical redd ages are complementary to redd age in days as local flow hydraulics affect the condition of redds differently; therefore, categorical age is a metric of visual identifiability. For each redd age-1 or age-2 (newly constructed and/or redd features crisp and well-defined), we measured the length and width of the pot and tail-spill to the nearest cm. For all redds, we visually estimated the proportion of the redd surface area (0%–100%) disturbed due to superimposition and assigned a binary cover score (1 if stream or riparian habitat features partially or fully obscured the redd as viewed by an observer standing on the stream bank; 0 otherwise). Finally, we used an electronic tablet (Samsung Galaxy Tab Active 2) to photograph each redd and annotate distinctive features with a drawing application, particularly superimposition of nearby redds (Fig. S1). We used photographs taken during previous surveys to identify redds under active construction and resolve difficulties in redd identification associated with superimposition (Gortázar et al. 2012). Counts from census surveys serve as the best estimate of "true" redd abundance in any given stream reach.

WGFD observer redd surveys

To assess the degree to which observer bias and variability affect the error structure of redd count data, we compared census surveys to redd surveys conducted independently by three WGFD fisheries biologists. Methods generally followed those of Muhlfeld et al. (2006) but were modified to accommodate WGFD redd count protocol, which accounts for

Fig. 1. (*a*) Map of the study area showing Upper Bar BC and Lower Bar BC spring creeks in the upper Snake River watershed, Wyoming. *Arrows* indicate direction of flow. *Alternating colored polygons* represent individual study reaches. *Striped polygons* indicate slack water areas lacking spawning gravel where spawning ground surveys were not performed. *Inset tables* denote lengths (m) of individual stream sections. *Streamlines* and *polygons* were traced from Google Earth Pro 7.3.6 (2022) imagery using the WGS 1984 Web Mercator projection. Proportion of redds not superimposed ("*NA*," redds per cluster = 1) and the distribution of proportional superimposition for redds within redd clusters (redds per cluster > 1) for (*b*) Upper Bar BC and (*c*) Lower Bar BC. *Dark bars* represent data from 22 July 2019; *light bars* represent data from 12 July 2021).

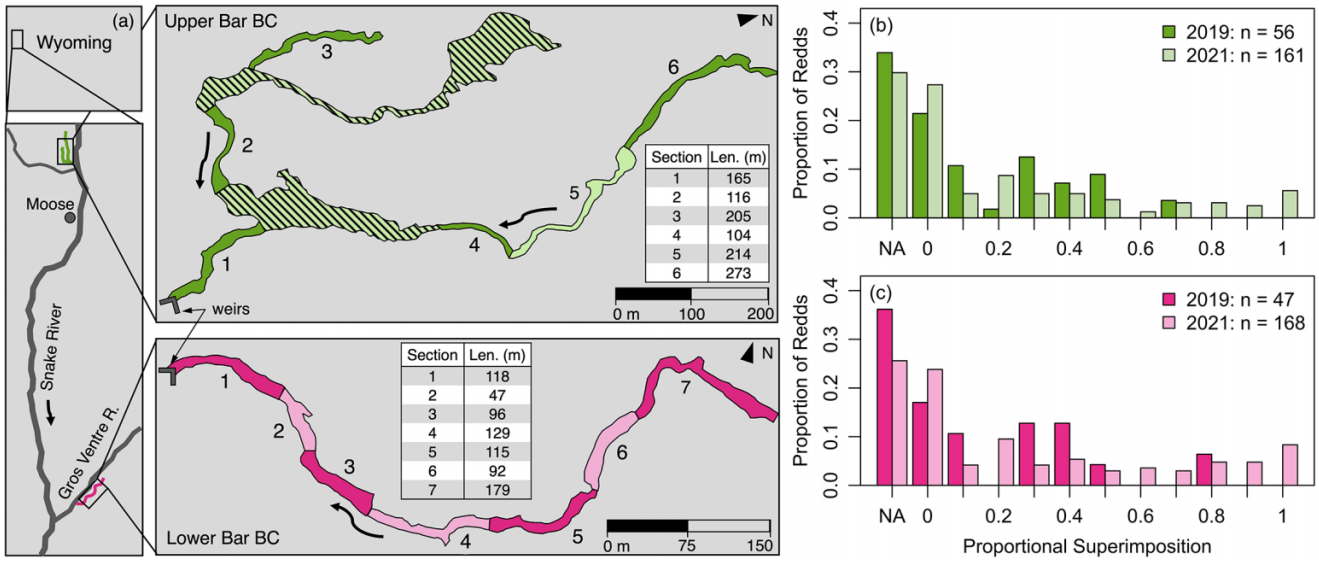


Table 1. Categorical redd age descriptions for spawning ground census surveys.

Age	Description
1	New encountered, not previously surveyed
2	Previously surveyed, features crisp and defined
3	Previously surveyed but features and margins not easily defined
4	Previously surveyed but not identifiable
5	Poor environmental conditions (visibility) preclude redd identification

superimposition. Each observer received departmental training on redd identification and survey protocol; further, prior to our study, each observer had conducted redd surveys on the two focal streams in addition to as many as fifteen other spring-fed streams in the region. Because our objective was to evaluate error structure over a range of redd densities and conditions, WGFD observers conducted redd surveys twice per year: at the peak of spawning and ~2 weeks afterwards. In our analysis we treated these two surveys as statistically independent for two reasons. First, redd densities and the condition of redds (e.g., categorical age) differed between the two surveys such that redds observed during the first survey were virtually unrecognizable during the second survey: the time between surveys (10 days in 2019 and 15 days in 2021) was equivalent to the number of days needed for redds to move between age-1 and age-3 (median = 13 days). Second, because WGFD observers conduct redd counts over the same stream segments annually, observers are already familiar with areas that commonly hold redds and thus the first survey conducted in each year will not inform the second any more than surveys conducted in past years. We validated this assumption of independence by including a hierarchical component for survey (i.e., early vs. late) in all candidate models and examined model outputs for evidence of shifts in the error structure between survey periods.

We provided WGFD observers with detailed maps on which they recorded the number and location of redds in a two-step process. First, redd clusters (single redds or multiple redds superimposed) were identified. Second, observers estimated the minimum and maximum number of redds in each cluster based on the number of distinct pots and total area of disturbance as compared to the average footprint of a single redd. For the purposes of our analysis, we considered the average of the minimum and maximum number of redds per cluster, consistent with WGFD protocol when reporting the final redd count.

Historical redd counts and weir operation

We assessed the biological significance of redd count data by comparing redd counts to direct measures of spawner abundance. Beginning in 1971, the WGFD has conducted redd surveys while simultaneously operating a weir during the spawning migration (late May–early July) to enumerate the number and size structure of YCT returning to spawn in LBBC. The weir captured adult YCT during their migration between the Snake River and the LBBC spawning grounds where redd counts were performed. WGFD biologists counted and determined the sex of all YCT captured each day. Historical redd

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

count data collection followed the procedure described above and was conducted once annually on approximately the date of peak spawning activity (early July).

Statistical analysis

To describe and quantify the error structure of redd count data we compared WGFD surveys to census surveys conducted on the same day. We matched redd clusters identified by observers to clusters identified and georeferenced during census surveys. We specifically noted any clusters not identified by observers (missed detections). Some clusters identified by observers could not be matched to known clusters and were therefore considered to be streambed features mistakenly identified as redds (false identifications). To calculate the size of each redd, we treated the pot and tail-spill as ellipses, calculated the area of each from field measurements, and summed the two values. To calculate the size of redd clusters, we summed the areas of all redds in that cluster, penalized (inverse weight) by the proportional superimposition for each redd.

Objective 1: decomposing error components and their drivers

Muhlfeld et al. (2006) defined observed redd counts as the sum of two independent, random processes: the number of redds detected plus the number of false identifications. We extended this framework to account for superimposition per WGFD protocol, which distinguishes redd clusters and redds within clusters. Specifically, we used a Bayesian hierarchical generalized linear modeling (GLM) framework to parse counting errors into three components: (1) imperfect detection of redd clusters, (2) false identifications, and (3) imperfect assignment of the number of redds per cluster. We used a fully crossed hierarchical structure (i.e., crossed random effects) to investigate variability among the three observers, two streams, and two years in our study. As our primary objective was to estimate the variability in error components among each level of the three grouping variables, we fit all models using a variable intercepts-fixed slopes structure. This allowed us to evaluate differences in error components (intercepts) among observers, streams, and years while simultaneously exploring the effects of redd cluster and stream habitat features on mean error rates.

We estimated cluster detection probabilities at the reach scale using a binomial GLM to test for differences among observers, streams, and years and to quantify the effect of cluster and habitat features on mean detection probability (Table S2):

$$(1)\quad D_{ijklm} \sim Binomial\left(p_{ijklm}, V_{jklm}\right)$$

$$(2)\quad logit\left(p_{ijklm}\right) = \alpha_1 + \alpha'_{1,i} + \alpha''_{1,j} + \alpha'''_{1,k} + \alpha''''_{1,l} + \widehat{\beta}_1 X_m$$

where D_{ijklm} is the number of clusters detected by observer i, in stream j, year k, survey period l, and reach m; p_{ijklm} is the estimated detection probability; V_{jklm} is the census number of clusters; α is the global intercept; α'_i is the observer-specific offset to the global intercept; α''_j is the stream-specific offset to the global intercept; α'''_k is the year-specific offset to the global intercept; α''''_l is the survey-specific offset to the global intercept; $\widehat{\beta}$ is a vector of parameter coefficients; and X_m represents the centered and scale covariate data summarized at the reach scale and the interaction terms. The numeric component of the α and β parameter subscripts reflects the relevant error component model: *1* for cluster detection, *2* for false identifications, and *3* for redds per cluster. All offsets to the global intercept were drawn from normal distributions with a mean of 0.

We modeled rates of false identifications at the reach scale using a negative binomial GLM to account for overdispersion in the data (Gelman and Hill 2007). All candidate models included an offset term to account for variable reach lengths (Table S3):

$$(3)\quad F_{ijklm} \sim Negative\,Binomial\left(p_{ijklm},\ size\right)$$

$$(4)\quad p_{ijklm} = size/\left(size + \lambda_{ijklm}\right)$$

$$(5)\quad \ln\left(\lambda_{ijklm}\right) = \ln\left(dist_l\right) + \alpha_2 + \alpha'_{2,i} + \alpha''_{2,j} + \alpha'''_{2,k} + \alpha''''_{2,l} + \widehat{\beta}_2 X_m$$

where F_{ijklm} is the number of false identifications committed by observer i, in stream j, year k, survey l, and reach m; p_{ijklm} and $size$ are parameters of the negative binomial distribution; λ_{ijklm} is the estimated rate of false identifications; $dist_l$ is the length of reach l; and all other parameters are defined as above.

Finally, we transformed the average observed number of redds per cluster into an error rate:

$$(6)\quad E_{ijklm} = \frac{O_{ijklm} - T_{ijklm}}{T_{ijklm}}$$

where E_{ijklm} is the error rate for observer i, stream j, year k, survey l, and cluster m; O_{ijklm} is the observed number of redds per cluster (average of minimum and maximum estimates); and T_{ijklm} is the census number of redds per cluster. Despite E_{ijklm} having a lower bound of -1, error rate data were approximately normally distributed around 0 (Fig. S2). We therefore modeled error rates in the number of redds per cluster following a normal distribution with covariates summarized at the scale of the redd cluster to avoid averaging over important variation in cluster attributes (Table S4):

$$(7)\quad E_{ijklm} \sim N\left(\mu_{ijklm},\ \sigma\right)$$

$$(8)\quad \mu_{ijklm} = \alpha_3 + \alpha'_{3,i} + \alpha''_{3,j} + \alpha'''_{3,k} + \alpha''''_{3,l} + \widehat{\beta}_3 X_m$$

where σ is the residual variance drawn from an exponential distribution (rate = 1), X_m represents the centered and scale covariate data summarized at the cluster scale and the interaction terms, and all other parameters are defined as above.

Objective 2: summarizing net error in redd count data

To determine the net effect of discrete error components on the overall accuracy and precision of redd count data, we

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

considered the linear relationship between observed (WGFD surveys) and census redd density at the reach scale and how this relationship varied among observers, streams, and years. We used density as opposed to raw count data to account for reaches that varied in length. Data were log-transformed to satisfy linear model assumptions of homoscedasticity. We fit the net error model with variable intercepts and variable slopes, which provide distinct but complementary information about error structure:

$$O_{ijklm} \sim N\left(\mu_{ijklm}, \sigma\right) \tag{9}$$

$$\mu_{ijklm} = \alpha + \alpha_i' + \alpha_j'' + \alpha_k''' + \alpha_l'''' + \beta\, T_{ijklm} + \beta_i'\, T_{ijklm} + \beta_j''\, T_{ijklm} + \beta_k'''\, T_{ijklm} + \beta_l''''\, T_{ijklm} \tag{10}$$

where O_{ijklm} is the observed number of redds per 100 m (logged) for observer i, stream j, year k, survey l, and reach m; T_{ijklm} is the census number of redds per 100 m (logged); σ is the residual variance drawn from an exponential distribution (rate = 2); β is the global slope; β_i' is the observer-specific offset to the global slope; β_j'' is the stream-specific offset to the global slope; β_k''' is the year-specific offset to the global slope; β_l'''' is the survey-specific offset to the global slope; and all other parameters are defined as above. Slope values greater (or less) than 1 (assuming intercept = 0) indicate relative bias increases (or decreases) as redd density increases. Intercept values greater (or less) than 0 (assuming slope = 1) indicate a constant positive (or negative) relative bias.

As monitoring is often based on data collected at the stream scale, we expanded our net error model to stream-scale data. We did not model the relationship between observed and census counts (e.g., eq. 10) due to small sample size that affected model convergence. Instead, we compared the relative bias in redd counts at the stream scale (R) to 0, where positive values indicate overestimates and negative values indicate underestimates of total redd count:

$$R_{ijkl} = \frac{O_{ijkl} - T_{ijkl}}{T_{ijkl}} \tag{11}$$

$$R_{ijkl} \sim N\left(\mu_{ijkl}, \sigma\right) \tag{12}$$

$$\mu_{ijkl} = \alpha + \alpha_i' + \alpha_j'' + \alpha_k''' + \alpha_l'''' \tag{13}$$

Where O_{ijkl} is the observed number of redds for observer i, stream j, year k, and survey l; T_{ijkl} is the census number of redds; and all other parameters are defined as above.

To determine how the precision of redd count data changes with redd density, we used simulations to explore how the standard deviation (SD) and coefficient of variation (CV) of observed redd densities changes with increasing census redd density. On a log-scale, we used eq. 10 to simulate 1000 estimates of observed redd density at each step in a sequence of hypothetical census redd densities (n = 100) using random draws from the posterior probability distributions for group-level intercepts, slopes, and the global SD. We back-transformed predictions and calculated the SD and CV at each step. We then used LOESS smoothing functions to qualitatively describe how the SD and CV of simulated observed redd densities change with census redd density.

Objective 3: exploring the biological significance of redd count data

We compared historical (1971–2021) redd counts and direct measures of female spawner abundance (weir data) to explore the biological significance of redd count data. Both datasets are subject to uncertainty stemming from observation error (redd counts) and an abbreviated monitoring window relative to the complete migration period (weir counts). We corrected redd counts using a regression model structurally identical to eq. 10, but instead of regressing observed against census redd density, response and predictor variables were swapped so that the error term would be applied to the estimate of census redd density. We used this model to generate unbiased estimates of historical redd counts. To generate estimates of female spawner abundance unbiased by variable operational periods of the weir, we used a hierarchical Bayesian migration timing model described by Adkison and Su (2001) and Walsworth and Schindler (2015). Field observations indicate little to no overwinter residency by adult YCT (WGFD and J. Baldock, *unpublished data*); therefore, we considered the model output to be the best estimate of female spawner abundance (see supplementary materials for migration timing model methods and results).

We modeled redd counts as a function of female spawner abundance using four functional relationships representing distinct hypotheses regarding the biological significance of redd counts (Table S5). A linear relationship implies that female spawners construct a proportional number of redds regardless of spawner density. A logarithmic relationship implies that spawners construct proportionally fewer redds as spawner abundance increase, but redd counts continue to increase with spawner abundance. An exponential decay relationship (increasing form) suggests that there is a point at which the spawning grounds become saturated with redds and spawner abundances above a threshold value will not yield greater redd numbers. A broken stick relationship has similar interpretation as the exponential decay relationship, but instead suggests that below the saturation point the relationship between spawners and redd counts is linear (rather than concave). We forced all candidate models through the origin as redds cannot be constructed without spawners.

Model fitting, evaluation, and significance testing

All Bayesian models were analyzed in the Just Another Gibbs Sampler MCMC sampling environment (JAGS; Plummer 2003), implemented through R (R Core Team 2021) using the HDInterval (Meredith and Kruschke 2020), lubridate (Grolemund and Wickham 2011), MCMCvis (Youngflesh 2018), R2jags (Su and Yajima 2021), and tidyverse (Wickham et al. 2019) packages. We assessed model convergence based on large effective sample sizes, low R-hat values (<1.001), and visual inspection of MCMC trace plots and posterior probability distributions (Gelman and Hill 2007). JAGS

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

models were run with parameter and/or hyperparameter prior probability distributions specified in Table S6.

Discrete error component candidate models only included predictor variables (centered and scaled) and interactions we considered biologically relevant (Tables S2–S4). We evaluated candidate models using leave-one-out cross-validation (LOO), in which the model with the greatest support is that with the lowest expected log pointwise predictive density ($elpd_{loo}$; Vehtari et al. 2017). We qualitatively considered the results of candidate models with similar support (i.e., the standard error of the $elpd_{loo}$ difference was greater than the difference itself); however, for simplicity, we only present graphical results from the single top model.

We did not use model selection to test for differences in error structure among groups, as quantifying these differences was our primary objective and model selection can fail in this regard (Kruschke 2014). Instead, we evaluated variability among groups post-hoc by examining the distributions of within-sample credible differences (i.e., difference between parameter estimates for each MCMC sample; Kruschke 2014). For each distribution of differences, we calculated the mean, probability of direction (pd), and probability within the region of practical equivalence (p-ROPE; Fig. S3; Kruschke 2014; Makowski et al. 2019). pd is the proportion of the distribution that is of the median's sign and varies between 50% and 100%; it can be interpreted as the probability that a difference (or, more generally, a parameter) is strictly positive or negative (Makowski et al. 2019). p-ROPE is the proportion of the distribution that lies within a specified range of a null value, where values within that range are considered practically equivalent. We applied these same metrics to the posterior probability distributions for parameters within the top models for each discrete error component. Together, these metrics provide insight into the existence and relevance of differences between hierarchical groups and covariate effects.

Bayesian indices of relevance (p-ROPE) are ultimately dependent on the pre-defined width of the ROPE, which we chose based on prior studies, published recommendations, and expert opinion. For the error component intercepts, we defined the ROPE as $0 \pm 5\%$ for cluster detection (<SD of detection probabilities from Muhlfeld et al. (2006) in which detection among observers differed only slightly), 0 ± 0.2 false identifications per 100 m (approx. the SD from Muhlfeld et al. (2006) in which rates did not differ among observers), and $0 \pm 10\%$ for the redds-per-cluster error rate (which would offset missed detections given that the average number of redds-per-cluster is ~2). For generalized linear model slope parameters, we defined the ROPE as 0 ± 0.1, following published recommendations (Kruschke 2014, Makowski et al. 2019).

Results

In 2019, a total of 48 and 59 redds were constructed in LBBC and UBBC, respectively, while 168 and 162 redds were constructed in the two creeks in 2021. Redd size varied from 0.16 to 3.05 m^2 (mean = 0.92 m^2). As shown in previous studies of superimposition for other trout species (Essington et al. 1998; Gortázar et al. 2012), we found that superimposition occurred frequently despite dramatic differences in redd densities among years: 65% and 63% of redds were superimposed in LBBC and UBBC in 2019 and 74% and 69% were superimposed in 2021. Of the superimposed redds, the proportion of redd surface area disturbed due to superimposition (i.e., proportional superimposition) ranged from 0–0.8 and 0–0.7 in LBBC and UBBC in 2019 and 0–1 in both streams in 2021 (Fig. 1*b* and 1*c*). Overall, the mean number of redds per cluster was 1.89 (median = 1) and 64% of clusters contained a single redd. In 2019, redd clusters contained 1–4 and 1–5 redds in LBBC and UBBC, respectively, and 1–14 and 1–13 redds in 2021. In 2019, mean cluster size was 1.34 (range = 0.22–5.61) and 1.03 m^2 (0.23–3.83) in LBBC and UBBC, respectively, and 1.37 (0.16–5.29) and 1.14 m^2 (0.21–7.16) in the two creeks in 2021. WGFD observers conducted redd counts on 12 July 2019, 22 July 2019, 28 June 2021, and 12 July 2021.

Long-term (1965–2021) LBBC weir sampling by WGFD captured as few as 39 YCT in 1973 (12 females, 28 May–2 July) and as many as 624 YCT in 2017 (337 females, 23 May 23–13 July; Fig. S4*a* and S4*b*). Weir sampling was conducted in every year except 1991 and 2016 due to private property access issues. Historical WGFD redd counts ranged from 30 to 379 total redds (Fig. S4*c*).

While we observed small differences in the magnitude of discrete error components between the early and late WGFD surveys conducted in each year, the magnitude of differences was small and likely equivalent to 0; thereby providing support for our assumption of independence between survey periods (Fig. S5; Table S7).

Estimation of discrete error components

We found detection probabilities of redd clusters were relatively low (global mean = 0.641). While we observed differences in detection among observers (Table 2; Fig. 2*a*–2*c*), post-hoc pairwise comparisons indicate that these differences were small (<0.060) and largely irrelevant (Table 2). Differences in detection among streams and years showed similar patterns in that median differences were small (0.039 among both streams and years), somewhat uncertain, and likely equivalent to 0 (Table 2). In general, detection decreased with the proportion of clusters covered (Fig. 2*d*) and increased with the proportion of clusters ≤ age-2 (redd features crisp and well-defined, i.e., more visible); however, the effect of age diminished when cover was high (Fig. 2*e*). The direction of these effects was highly certain and unlikely to be equivalent to 0 (Table 3). Cover also mediated the effect of cluster size: when cover was low, cluster size had little effect on detection probability, but when cover was high, detection increased dramatically with size (Table 3; Fig. 2*f*). The effect of cluster size alone was small and practically equivalent to 0 (Table 3; Fig. 2*f*). Additionally, detection decreased with mean cluster age in days (Fig. 2*g*). Candidate models with similar, but less, statistical support did not include an effect of cluster age in days, suggesting uncertainty in the effect as indicated by a 95% credible interval overlapping 0 (Table 3; Table S2).

We found false identifications were committed infrequently (global mean = 0.130 per 100 m). While we observed differences in rates of false identifications among observers, streams, and years (Fig. 3*a*–3*c*), post-hoc pairwise

Table 2. Post-hoc pairwise credible differences between each level of the three grouping variables (observers, streams, and years) for each Bayesian hierarchical model.

Model	Comparison	Median (95% CI)	pd (%)	p-ROPE (%)
Cluster detection probability	Obs 1–Obs 2	0.005 (−0.065, 0.076)	55.7	83.3
	Obs 1–Obs 3	−0.054 (−0.124, 0.015)	93.9	44.6
	Obs 2–Obs 3	−0.060 (−0.128, 0.011)	95.2	39.2
	Lower–Upper	0.039 (−0.024, 0.102)	88.8	62.8
	2019–2021	0.039 (−0.073, 0.153)	75.7	51.4
Rate of false identifications (per 100 m)	Obs 1–Obs 2	0.106 (−0.006, 0.262)	98.2	89.2
	Obs 1–Obs 3	−0.081 (−0.298, 0.118)	82.5	87.8
	Obs 2–Obs 3	−0.188 (−0.384, −0.049)	99.9	55.6
	Lower–Upper	−0.088 (−0.227, 0.035)	93.6	94.3
	2019–2021	0.220 (0.054, 0.456)	99.9	41.4
Redds per cluster error rate (%)	Obs 1–Obs 2	24.8 (14.0, 35.6)	100.0	0.4
	Obs 1–Obs 3	−26.9 (−37.5, −16.2)	100.0	0.1
	Obs 2–Obs 3	−51.7 (−62.5, −41.1)	100.0	0.0
	Lower–Upper	9.5 (0.3, 18.9)	97.9	53.7
	2019–2021	12.8 (2.5, 23.3)	99.1	29.6
Net accuracy (intercept)	Obs 1–Obs 2	0.098 (−0.243, 0.441)	71.6	38.0
	Obs 1–Obs 3	−0.168 (−0.508, 0.180)	83.3	28.3
	Obs 2–Obs 3	−0.266 (−0.614, 0.078)	93.6	15.2
	Lower–Upper	−0.037 (−0.326, 0.256)	60.0	49.2
	2019–2021	0.465 (0.162, 0.777)	99.9	0.9
Net accuracy (slope)	Obs 1–Obs 2	0.031 (−0.113, 0.175)	66.8	79.5
	Obs 1–Obs 3	0.006 (−0.140, 0.146)	52.9	82.9
	Obs 2–Obs 3	−0.026 (−0.171, 0.119)	63.6	80.0
	Lower–Upper	0.071 (−0.049, 0.195)	87.6	67.6
	2019–2021	−0.001 (−0.137, 0.136)	50.6	84.9

Note: *Median (95% CI)* indicates the posterior median and 95% credible interval of the difference between groups, *pd* is the probability of direction (the proportion of the posterior mass that is of the median's sign), and p-ROPE is the proportion of posterior mass that lies within the region of practical equivalence.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

comparisons indicate that these differences were small and largely irrelevant (Table 2). In general, the rate of false identifications increased with the census number of clusters (Table 3, Fig. 3*d*). The effect of the proportion of clusters ≤ age-2 alone was small, directionally uncertain, and had a moderate probability of being equivalent to 0 (Table 3); however, the proportion of clusters ≤ age-2 and the census number of clusters interacted additively, such that the rate of false identifications was maximized when both the census number of clusters and the proportion of clusters ≤ age-2 were large (Table 3; Fig. 3*f*). A candidate model with similar, but less, support did not include effects of age or the interaction term, indicating that the rate of false identifications was primarily driven by the census number of clusters (Table S3).

We found that the number of redds per cluster was consistently overestimated (global mean error rate = 38%). Mean error rates were 37%, 12%, and 64% for observers 1, 2, and 3, respectively (Fig. 4*a*). Differences in error rates among the three observers were substantial, highly certain (pd = 100%), and unlikely to be equivalent to 0 (p-ROPE < 1%; Table 2). In contrast, while differences in error rates between streams (10%) and years (13%) were highly certain (pd = 98% and 99%), differences were somewhat likely to be equivalent to 0 (p-ROPE = 54% and 30%; Table 2; Fig. 4*b*–4*c*). We found that error rates in the number of redds per cluster were negatively related to mean proportional superimposition: observers overestimated redds per cluster when superimposition was low but underestimated redds per cluster when superimposition was high, though this relationship was most pronounced when minimum categorical age was high (Fig. 4*d*). The effect of categorical age alone was small and likely equivalent to 0 (Table 3; Fig. 4*e*). In contrast, we found that error rates in the number of redds per cluster increased with the minimum redd age in days and this effect was most pronounced when clusters were very large (Fig. 4*f*). However, the effect of cluster size alone was small and likely equivalent to 0 (Table 3; Fig. 4*g*). While the direction of both interaction terms was highly certain, the magnitude of these effects were small and likely equivalent to 0 (Table 3). A model with similar, but less, support indicated that the effect of cluster size diminished when superimposition was high (Table S4), but the strength of this interaction was modest.

Net accuracy and precision in redd count data

We assessed the net accuracy of observed redd count data by comparing the linear relationship between observed

Fig. 2. Hierarchical Bayesian generalized linear model output for the cluster detection probability model. (*a–c*) Posterior probability distributions for crossed hierarchical intercepts (i.e., mean detection probability) for observers, streams, and years. Colored polygons represent 95% credible intervals of the posterior distributions. (*d–g*) Effects of cover, categorical age, cluster size, and absolute age (days) on detection probability. *Thick black lines* and *grey polygons* represent the main (i.e., marginal) effects and 95% credible intervals for the global distribution of means, respectively. *Colored lines* represent one or more interactive effects, with the interactive variable held at its minimum (*lighter shade*) or maximum (*darker shade*) value.

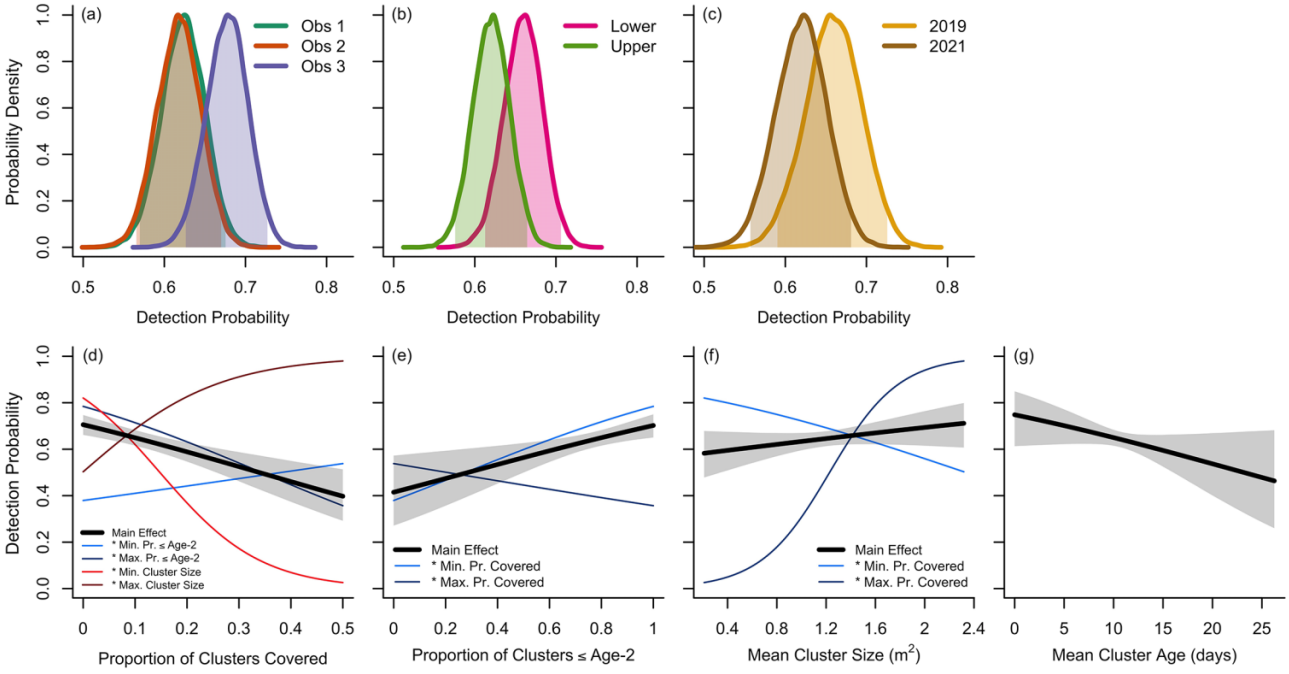


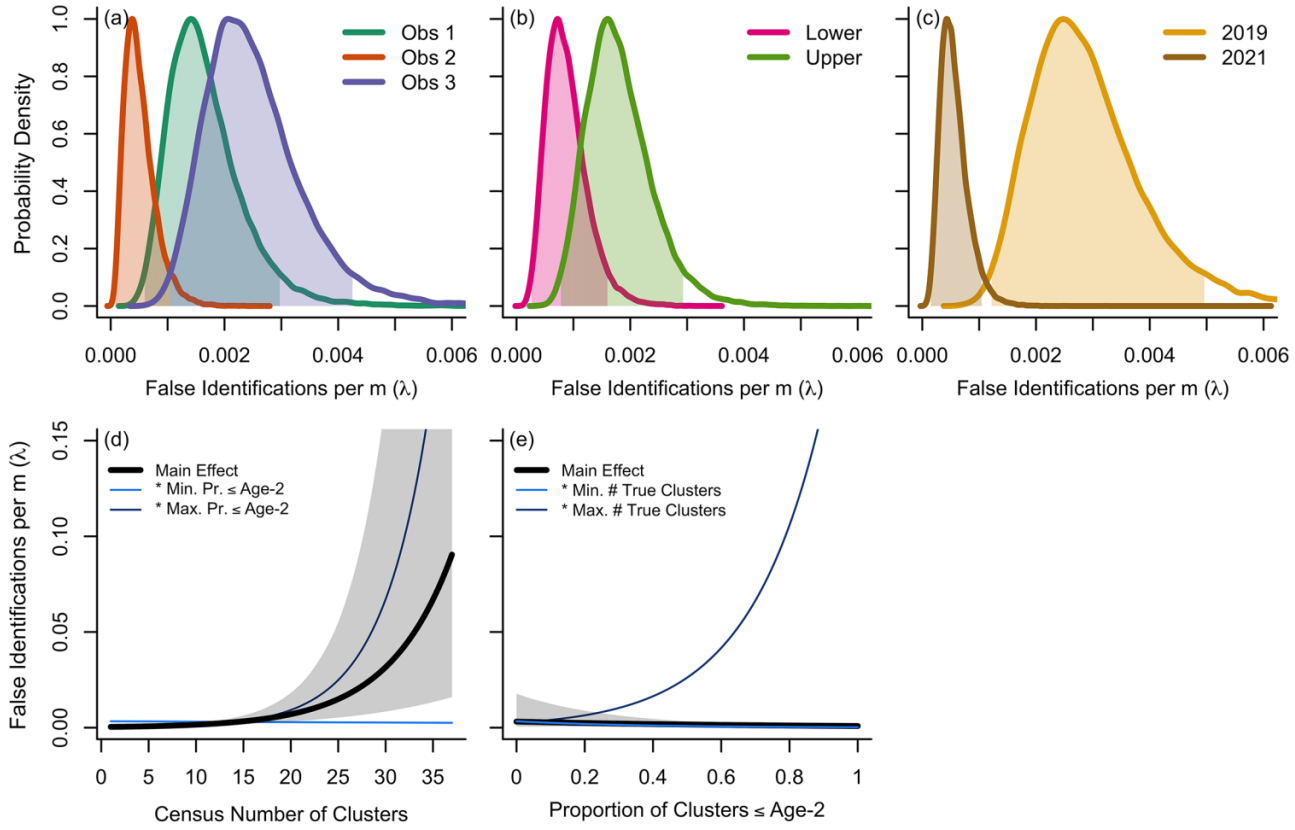
Table 3. Summaryof the drivers of discrete error components.

Model	Parameter	Median (95% CI)	pd (%)	p-ROPE (%)
Cluster detection probability	cov	−0.355 (−0.518, −0.189)	100.0	0.0
	agecat	0.388 (0.122, 0.652)	99.9	0.0
	clsize	0.133 (−0.067, 0.329)	91.1	35.9
	agedays	−0.292 (−0.649, 0.070)	94.6	12.8
	cov * agecat	−0.225 (−0.363, −0.086)	99.9	3.8
	cov * clsize	0.584 (0.365, 0.800)	100.0	0.0
Rate of false identifications	true	1.067 (0.602, 1.634)	100.0	0.0
	agecat	−0.411 (−1.156, 0.319)	87.3	11.6
	true * agecat	0.472 (0.095, 0.900)	99.4	2.0
Redds per cluster error rate	sup	−0.284 (−0.346, −0.225)	100.0	0.0
	agecat	−0.082 (−0.142, −0.022)	99.7	72.3
	agedys	0.222 (0.156, 0.290)	100.0	0.0
	clsize	0.046 (−0.023, 0.116)	90.2	93.9
	sup * agecat	−0.060 (−0.111, −0.011)	99.2	94.2
	clsize * agedys	0.084 (0.018, 0.152)	99.4	67.7

Note: *Median* (*95% CI*) indicates the posterior median and 95% credible interval for the parameter in question, *pd* is the proportion of the posterior mass that is of the median's sign, and *p-ROPE* is the proportion of posterior mass that lies within the region of practical equivalence (limits = 0 ± 0.1). Cluster detection predictor variables are defined as follows: *cov* = proportion of clusters covered by stream habitat features, *agecat* = proportion of clusters with a minimum categorical age of 1 or 2, and *clsize* = mean cluster size in m^2 at the reach scale. False identification predictor variables are defined as follows: *agecat* = proportion of clusters with a minimum categorical age of 1 or 2 and *true* = true number of redds present. Redds per cluster error rate predictor variables are defined as follows: *agecat* = minimum categorical age, *clsize* = cluster size in m^2 at the cluster scale, *agedys* = mean redd age in days at the cluster scale, and *sup* = mean proportional superimposition. All predictor variables were centered and scaled prior to model estimation.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Fig. 3. Hierarchical Bayesian generalized linear model output for the rate of false identification model. (*a–c*) Posterior probability distributions for crossed hierarchical intercepts (i.e., number of false identifications per m) for observers, streams, and years. *Colored polygons* represent 95% credible intervals of the posterior distributions. (*d*, *e*) Effects of the number of true clusters and categorical age on the rate of false identification. *Thick black lines* and *grey polygons* represent the main (i.e., marginal) effects and 95% credible intervals for the global distribution of means, respectively. *Colored lines* represent interactive effects, with the interactive variable held at its minimum (*lighter shade*) or maximum (*darker shade*) value.



and census redd densities to 1:1. In general, we found that observed redd densities accurately characterized census densities (Fig. 5). There was a 53% probability that the global slope (0.902) was practically equivalent to 1 (p-ROPE; Table S8), suggesting minimal bias as a fixed rate. Similarly, there was an 80% probability that the global intercept (0.008) was practically equivalent to 0 (p-ROPE; Table S8), suggesting minimal bias as a fixed offset. We observed differences in slopes among observers, streams, and years, but differences were small (range = 0.001–0.071), relatively uncertain, and practically equivalent to 0 (Table 2; Fig. S6). Similarly, differences in intercepts among observers and streams were minimal (Table 2; Fig. S6). Interestingly, the intercept for 2021 was considerably less than that for 2019 (Table 2; Fig. 5*c*; Fig. S6).

Our analysis of the net error in redd count data at the stream scale supported results at the reach scale (Fig. 5, *inset boxplots*). On average, observed counts underestimated census counts by 14.7% (pd = 99%), but there was a 21% probability that this difference is practically equivalent to 0 (p-ROPE). Relative error at the stream scale tended to differ among observers, but these differences were generally small (Table S9). Additionally, relative error tended not to vary among streams but did vary among years (Table S9).

Our simulation-based analysis of the precision of redd count data shows that the SD of observed redd densities increases with census redd density but the CV is constant (Fig. S7). This finding is consistent with our modeling of net accuracy using log-transformed data to satisfy assumptions of homoscedasticity (mean and variance scale positively). Variation in the effect of redd density on the SD of predictions among observers, streams, and years is consistent with our net accuracy model (Fig. S7*a*), whereas variation in the CV is minimal (Fig. S7*b*).

Biological significance of redd count data

Visual inspection of the data revealed a single outlier in our dataset (year 2017: model-corrected redd count = 434 ± 185, model-corrected female abundance = 371 ± 32; mean ± SD). This data point represents the first year of redd counts conducted by a specific WGFD observer, which happened to coincide with the largest spawning population abundance on record. Given limited experience and difficult conditions for redd enumeration, we did not feel our net accuracy model could provide a robust estimate of the census redd count. Preliminary data analysis also revealed that this data point exerted considerable leverage over parameter estimates

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Fig. 4. Hierarchical Bayesian linear model output for the error rate (%) in the number of redds assigned per cluster. (*a*-*c*) Posterior probability distributions for crossed hierarchical intercepts (i.e., error rate) for observers, streams, and years. *Colored polygons* represent 95% credible intervals of the posterior distributions. (*d*–*f*) Effects of mean proportional superimposition, minimum categorical age, minimum age (days), and cluster size on the error rate in the number of redds assigned per cluster. *Thick black lines* and *grey polygons* represent the main (i.e., marginal) effects and 95% credible intervals for the global distribution of means, respectively. *Colored lines* represent interactive effects, with the interactive variable held at its minimum (*lighter shade*) or maximum (*darker shade*) value.

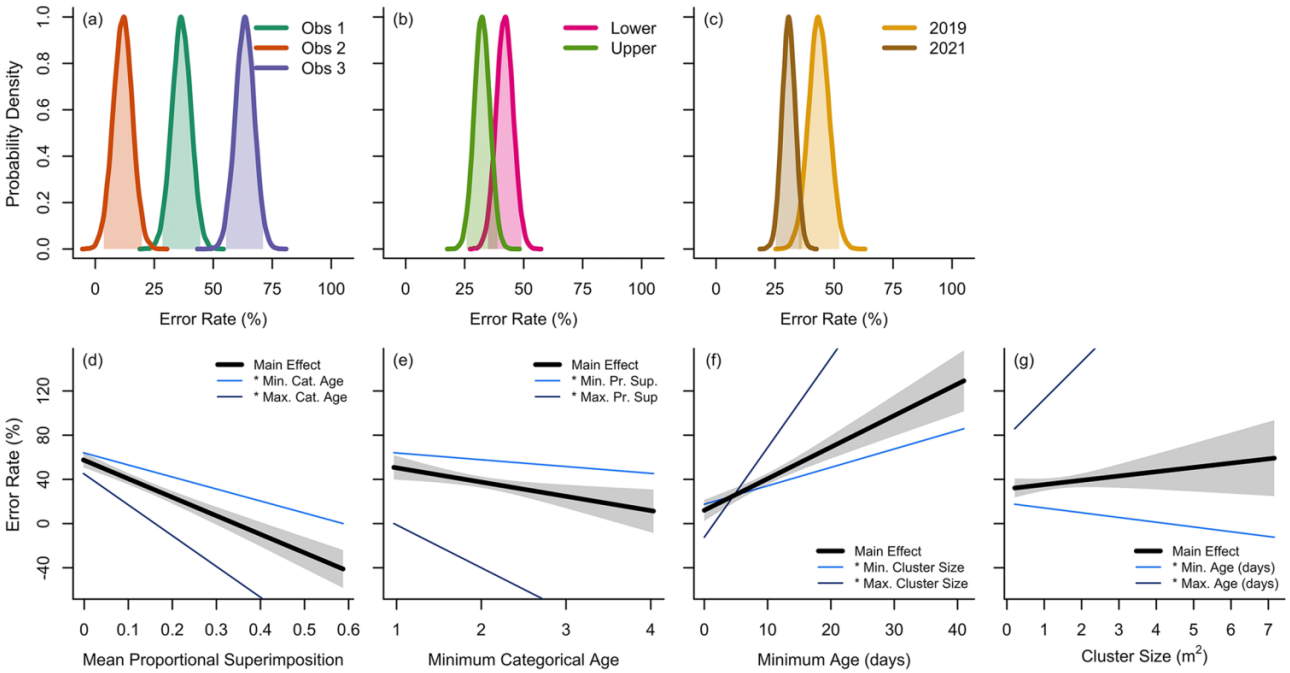
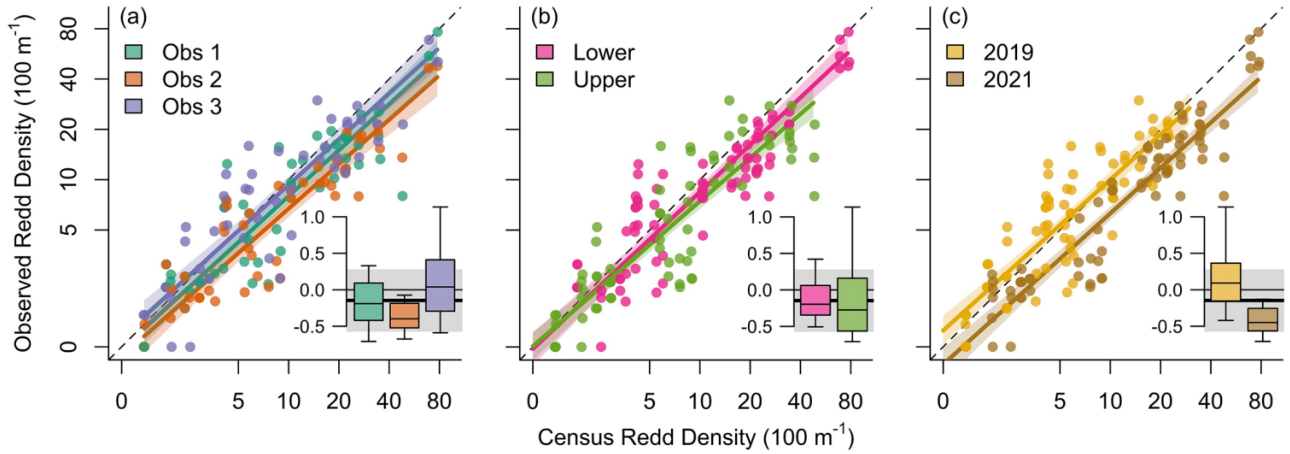


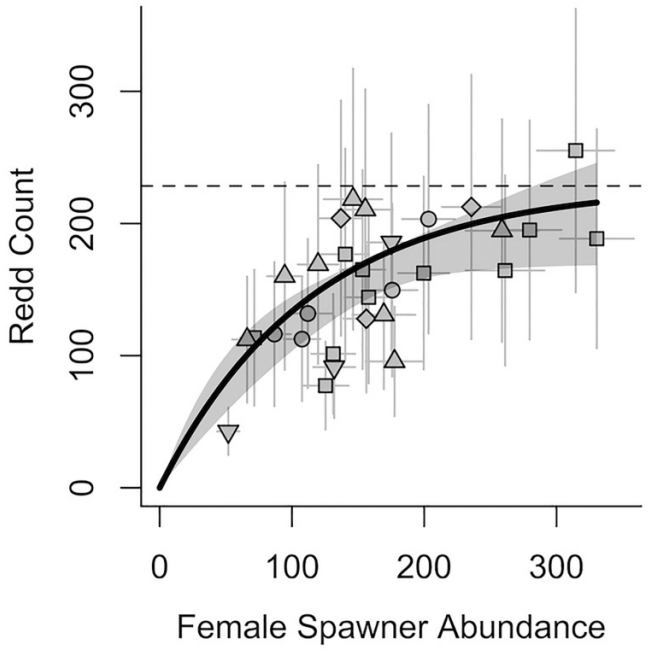
Fig. 5. Hierarchical Bayesian linear model output of the relationship between observed and census redd densities (log scale, redds per 100 m) faceted by (*a*) observers, (*b*) streams, and (*c*) years. *Dashed lines* denote 1:1 for visual reference. *Inset boxplots* summarize the relative error in redd counts at the scale of the entire stream, where 0 indicates no differences between observed and true counts. *Thick horizontal lines* and *grey polygons* represent the global mean ± SD relative error among groups (−0.15 ± 0.43). *Thin horizontal lines* denote 0 (no difference) for visual reference.



and reshuffled candidate model performance and selection using LOO, fundamentally altering our conclusions. We, therefore, removed this data point from all analyses. Model selection and evaluation supported a saturating (exponential decay, increasing form) relationship between female spawning abundance and redd count (Fig. 6). The asymptotic redd count was estimated at 228 redds (posterior mean). Increasing the abundance of spawning females, particularly above approx. 200 females, produces diminishing marginal increases in the abundance of redds in this system.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Fig. 6. Redd count corrected for observation error as a function of the female spawner abundance corrected for the timing of weir operation. *Error bars* represent standard deviations. *Black line* and *grey polygon* represent the top model output and 95% credible interval (exponential decay, increasing form). *Dashed line* represents the asymptotic limit of the exponential function. *Point shape* denotes unique observers.



Discussion

Redd counts are commonly used to monitor the status and trends of salmonid populations at broad spatial and temporal scales (Rieman and Myers 1997; Kovach et al. 2017). While effects of inter-observer variability and bias on the error structure of redd count data are well documented (Dunham et al. 2001; Muhlfeld et al. 2006; Howell and Sankovich 2012), understanding how superimposition mediates the accuracy and precision of redd count data has remained elusive. We found that redd superimposition is the leading cause for variability in error rates among observers in our system (Figs. 2–4). However, discrete error components act antagonistically, such that observed redd counts accurately reflect census redd abundance (Fig. 5). While precision declined with increasing redd density, the CV was constant (Fig. S7), indicating that redd counts are well suited for long-term monitoring. Finally and critically, we found that estimates of spawner abundance derived from redd count data may be misleading, especially under high redd density conditions (Fig. 6). Instead, we argue that redd count data are best interpreted as effective reproductive effort.

Sampling error in redd counts is driven by superimposition

Our investigation into discrete error components revealed that inter-observer variability was minimal for the cluster detection and false identifications error components. In contrast, error rates in the number of redds per cluster varied considerably among observers and variation in the global error rate depended strongly on mean proportional superimposition (Fig. 4*d*). Our results indicate that both the frequency and intensity of superimposition exert considerable influence on the error structure of redd count data; providing explicit evidence for prior assumptions (*sensu*Dunham et al. 2001; Muhlfeld et al. 2006; Murdoch et al. 2019). Further, the effects of superimposition on the error structure of redd count data should be considered when surveying populations spawning at both high and low densities, unlike what has been suggested previously (Lestelle and Weller 2002). That superimposition was common under low density conditions suggests fish have a preference for spawning on top of or near existing redds (*sensu*Essington et al. 1998). Therefore, the accuracy and precision of redd count data are important to consider when designing and implementing monitoring programs, particularly for vulnerable and imperiled populations spawning at low densities.

While our models of the drivers of error components were simple, our results provide insight into the causes and consequences of observation error in redd count data. For example, detection probability decreased with cover and increased with the proportion of clusters $\leq$ age-2 (Fig. 2), lending support to prior work (Heggberget et al. 1986; Dunham et al. 2001). The effect of cover on cluster detection suggests that habitat conditions should be considered when deciding when and where to conduct redd counts, as their effectiveness may vary among reaches and streams with considerably different habitat characteristics. We also found that false identifications were only committed when the census number of clusters and the proportion of clusters $\leq$ age-2 were high (Fig. 3). This result conflicts with prior work showing that false identifications are committed independent of redd density (Muhlfeld et al. 2006), although it is possible that this relationship was masked by a low signal-to-noise ratio as overall redd densities were low in the previous study. However, false identifications were committed infrequently, such that the effect on total redd count can be considered irrelevant with little impact on the efficacy of monitoring.

Redd counts are accurate and well suited for monitoring

Despite variation in the magnitude of discrete error components among observers, we found that error components tended to offset each other such that the relationship between observed and census redd counts approximated 1:1 (on the log scale; Fig. 5; *sensu*Muhlfeld et al. 2006). Our consideration of relative error in redd counts at the stream-scale supports results at the reach-scale: while observed counts tended to underestimate census redd abundance, the mean difference was small and practically equivalent to 0 (Fig. 5). Our results indicate that the error structure we describe at the reach-scale can be applied more broadly to whole stream redd counts. This is promising as redd count monitoring programs often mix index reaches and census counts, or combine reaches of different lengths (Rieman and McIntyre 1996).

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

We found that the uncertainty (i.e., SD) in observed redd counts increased with census redd density (Fig. S6*a*). In a monitoring context, this may limit the ability to detect whether recovery targets are being met, as greater variance at high redd densities may introduce considerable uncertainty into trend estimates (Wagner et al. 2013). However, as conservation efforts are typically most concerned with declining populations (e.g., Reed and Blaustein 1997; Maxwell and Jennings 2005*b*), reduced variance at low redd densities suggests that redd count monitoring will reveal declines that precipitate necessary management intervention. Admittedly, the precision of redd count data documented in this study is relatively low (CV $\cong$ 0.42), which may present difficulties for effective monitoring (Stier et al. 2022). While low precision may limit power to detect impacts of external drivers, such as increased predation, on population abundance (Oken and Essington 2015; Walsworth and Schindler 2016), high accuracy indicates that redd counts can be an efficient monitoring tool given that management plans are robust to uncertainty. For example, trends derived from redd count data should be interpreted as means with a variance and caution should be taken when comparing means to thresholds demarcating when interventions should or should not be employed (Ham and Pearsons 2000; Dauwalter et al. 2009). Further, low precision highlights the importance of long-term monitoring in adequately detecting and describing changes in population status (Al-Chokhachy et al. 2009).

Redd counts are best interpreted as effective reproductive effort

Our analysis of the biological significance of redd count data revealed a saturating relationship between female spawner abundance and redd count (Fig. 6), consistent with past work (Groves et al. 2013). Saturation of redd numbers at high spawner abundances can be attributed to the fact that many redds have been destroyed due to superimposition (McNeil 1964) and are therefore not visible to observers. This is supported by the results of our net accuracy model, for which the intercept declined in 2021 relative to 2019, suggesting increasing negative bias in years with greater redd densities (Fig. 5*c*). While redd counts may reflect changes in spawner abundance when abundance is low, increasing superimposition and density dependence at high abundances weakens the relationship, restricting the use of redd counts to infer abundance (Hay 1984). Therefore, redd count monitoring may not reflect large fluctuations in spawner numbers if changes occur above the redd capacity of the spawning grounds (Groves et al. 2013).

Limited availability of suitable spawning habitat is often considered the primary cause for superimposition (Blanchfield and Ridgway 1997). The occurrence of superimposition has thus been used to justify costly spawning habitat restoration projects in our study system and others (Kiefling 1997; Dudley 2019). However, our result that superimposition was common at low redd densities complements past work in suggesting that salmonids maintain a preference for superimposition (Essington et al. 1998; Youngson et al. 2011). Such preference may result in otherwise suitable spawning habitat going unused, even when population densities are high (Groves et al. 2013). Therefore, the interpretation of redd count is context specific and relates to the redd capacity of the spawning grounds, the life history of the species in question, and the specific goals of the management or monitoring program.

Redd count data are perhaps best interpreted as effective reproductive effort as it is widely appreciated that redd superimposition reduces egg and fry survival (McNeil 1964; Hayes 1987). Therefore, underestimates of total redd numbers due to superimposition are biologically and practically irrelevant: there should be no need or interest in counting redds that will not contribute to recruitment. Instead, because only those redds minimally affected by superimposition produce fry (Baldock et al. *in prep*), redd count data are better interpreted as effective reproductive effort, that is, an index of fry production (Beard and Carline 1991; Beland 1996). Direct measures of spawner abundance may be marginally correlated with recruitment given the effects of superimposition and additional sources of density-dependence on spawning success. Additionally, simple measures of abundance may fail to describe the capacity of populations to recover from periods of decline given that spawning frequency and fecundity depend on body size and condition (Meyer et al. 2003; Haraldstad et al. 2018). Redd counts may, therefore, provide conservation efforts with more relevant information regarding population status, as fry production and recruitment underlie long-term population dynamics in many trout species (Elliot 1994; Lobón-Cerviá 2009; Kanno et al. 2016).

Conclusions

In this study, we provide a mechanistic understanding of the sampling error associated with redd counts for YCT, the resulting net accuracy and precision of redd count data, and how redd counts can be interpreted with respect to spawner abundance. Mechanistic approaches that break down the sources of sampling error not only form the basis for improvements to existing protocols but can also guide decisions regarding when and where monitoring is most appropriately applied (Murdoch et al. 2018). Bayesian approaches allowed us to express results in terms of probability and biological relevance rather than p-values and effect sizes that can be difficult to interpret. Such flexibility is valuable for biologists tasked with designing management plans that are robust to uncertainty (Dauwalter et al. 2009; Pregler et al. 2019). Our results illustrate how reproductive behavior mediates the value and interpretation of population monitoring data. Flexible management strategies that acknowledge the limitations of monitoring techniques are needed to equip practitioners with the tools to manage populations of conservation concern (Schindler and Hilborn 2015).

Acknowledgements

We thank Lukas Brooks, Quincy Harris, Nate Heili, and Sasha Pereira for providing substantial field assistance. Rob Gipson, Clark Johnson, Diana Miller, and Mark Smith of the Wyoming Game and Fish Department provided field assistance, historical data, and expert advice regarding the initial study

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

design and broader implications. Chad Whaley of the National Park Service provided field assistance and expert advice. We thank the Overlock and Morgan families for allowing access to their properties on Lower Bar BC spring creek. We thank Adam Sepulveda and two anonymous reviewers for providing helpful comments on an earlier draft of this manuscript. All YCT were treated humanely, and the methods were approved by the University of Wyoming Institutional Animal Care and Use Committee protocol nos. 20190816AW00389-01 and 20200507AW00423-02. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the U.S. Government.

Article information

History dates

Received: 8 November 2022
Accepted: 4 January 2023
Accepted manuscript online: 20 January 2023
Version of record online: 28 February 2023

Copyright

© 2023 Copyright remains with the author(s) or their institution(s). Permission for reuse (free in most cases) can be obtained from copyright.com.

Data availability statement

Data generated and/or analyzed in this study are available in "Baldock-et-al_ReddCountsSuperimposition_2022" repository on GitHub, https://github.com/j-baldock/Baldock-et-al_ReddCountsSuperimposition_2022.git.

Author information

Author ORCIDs

Jeffrey R. Baldock https://orcid.org/0000-0002-2791-7535
Robert Al-Chokhachy https://orcid.org/0000-0002-2136-5098
Timothy E. Walsworth https://orcid.org/0000-0002-7352-3430
Annika Walters https://orcid.org/0000-0002-8638-6682

Author contributions

Conceptualization: JRB, RA, AW
Formal analysis: JRB
Funding acquisition : JRB, AW
Investigation: JRB
Methodology : JRB, RA, TEW, AW
Project administration: JRB, AW
Supervision: JRB, RA, AW
Validation: JRB, RA, TEW, AW
Visualization: JRB
Writing – original draft: JRB
Writing – review and editing: RA, TEW, AW

Competing interests

The authors declare there are no competing interests.

Funding statement

Funding for this research was provided by the Wyoming Game and Fish Department (grant no. 002 827) and the Jackson Hole One Fly Foundation (grant no. 2019-032). Additional funding for this research was provided by the Department of Zoology and Physiology's graduate student scholarship program at the University of Wyoming (Dennis Anderson Memorial Scholarship from the Jackson Hole One Fly Foundation, the Lyman and Margie McDonald Research Award for Quantitative Analysis in Wildlife and Fisheries Ecology, Vern Bressler Fisheries Fund Scholarship, and the Western Ecosystems Technology Research Award for Quantitative Analysis in Wildlife and Fisheries Ecology).

Supplementary material

Supplementary data are available with the article at https://doi.org/10.1139/cjfas-2022-0267.

References

Adkison, M.D., and Su, Z. 2001. A comparison of salmon escapement estimates using a hierarchical Bayesian approach versus separate maximum likelihood estimation of each year's return. Can. J. Fish. Aquat. Sci. **58**: 1663–1671. doi:10.1139/f01-100.

Al-Chokhachy, R., Budy, P., and Conner, M. 2009. Detecting declines in the abundance of a bull trout (*Salvelinus confluentus*) population: understanding the accuracy, precision, and costs of our efforts. Can. J. Fish. Aquat. Sci. **66**: 649–658. doi:10.1139/F09-026.

Al-Chokhachy, R., Shepard, B.B., Burckhardt, J.C., Garren, D., Opitz, S., Koel, T.M., et al. 2018. A portfolio framework for prioritizing conservation efforts for Yellowstone Cutthroat Trout populations. Fisheries **43**: 485–496. doi:10.1002/fsh.10137.

Alves, J., Krieger, D., and Nesler, T. 2004. *Conservation Plan for Rio Grande Cutthroat Trout in Colorado*. Colorado Division of Wildlife. Denver, CO.

Auerbach, D.S., and Fremier, A.K. 2022. Identification of salmon redds using RPV-based imagery produces comparable estimates to ground counts with high inter-observer variability. River Res. Appl. **39**: 35–45.

Beard, T.D.J., and Carline, R.F. 1991. Influence of spawning and other stream habitat features on spatial variability of wild brown trout. Trans. Am. Fish. Soc. **120**: 711–722. doi:10.1577/1548-8659(1991)120%3c0711:IOSAOS%3e2.3.CO;2.

Beland, K.F. 1996. The relation between redd counts and Atlantic salmon (Salmo salar) parr populations in the Dennys River, Maine. Can. J. Fish. Aquat. Sci. **53**: 513–519. doi:10.1139/f95-216.

Blanchfield, P.J., and Ridgway, M.S. 1997. Reproductive timing and use of redd sites by lake-spawning brook trout (*Salvelinus fontinalis*). Can. J. Fish. Aquat. Sci. **54**: 747–756. doi:10.1139/f96-344.

Chasco, B.E., Ward, E.J., Hesse, J.A., Rabe, C., Kinzer, R., Vogel, J.L., and Orme, R. 2014. Evaluating the accuracy and precision of multiple abundance estimators using state-space models: a case study for a threatened population of chinook salmon in Johnson Creek, Idaho. N. Am. J. Fish. Manag. **34**: 945–954. doi:10.1080/02755947.2014.926302.

Crisp, D.T., and Carling, P.A. 1989. Observations on siting, dimensions and structure of salmonid redds. J. Fish Biol. **34**: 119–134. doi:10.1111/j.1095-8649.1989.tb02962.x.

Dauble, D.D., and Watson, D.G. 1997. Status of fall chinook salmon populations in the mid-Columbia River, 1948–1992. N. Am. J. Fish. Manag. **17**: 283–300. doi:10.1577/1548-8675(1997)017%3c0283:SOFCSP%3e2.3.CO;2.

Dauwalter, D.C., Rahel, F.J., and Gerow, K.G. 2009. Temporal variation in trout populations: implications for monitoring and trend detection. Trans. Am. Fish. Soc. **138**: 38–51. doi:10.1577/T07-154.1.

Dudley, P.N. 2019. S4: a spatially continuous, individual-based model of salmonid redd superimposition. Trans. Am. Fish. Soc. **148**: 352–372. doi:10.1002/tafs.10139.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.

Dunham, J.B., Rieman, B.E., and Davis, K. 2001. Sources and magnitude of sampling error in redd counts for bull trout. N. Am. J. Fish. Manag. **21**: 343–352. doi:10.1577/1548-8675(2001)021%3c0343:SAMOSE%3e2.0.CO;2.

Elliott„ J.M. 1994. Quantitative ecology and the brown trout. Oxford University Press, pp. 304.

Essington, T.E., Quinn, T.P., and Ewert, V.E. 2000. Intra- and interspecific competition and the reproductive success of sympatric Pacific salmon. Can. J. Fish. Aquat. Sci. **57**: 205–213. doi:10.1139/f99-198.

Essington, T.W., Sorensen, P.W., and Paron, D.G. 1998. High rate of redd superimposition by brook trout (*Salvelinus fontinalis*) and brown trout (*Salmo trutta*) in a Minnesota stream cannot be explained by habitat availability alone. Can. J. Fish. Aquat. Sci. **55**: 2310–2316. doi:10.1139/f98-109.

Gallagher, S.P., and Gallagher, C.M. 2005. Discrimination of chinook salmon, coho salmon, and steelhead redds and evaluation of the use of redd data for estimating escapement in several unregulated streams in Northern California. N. Am. J. Fish. Manag. **25**: 284–300. doi:10.1577/M04-016.1.

Gelman, A., and Hill, J. 2007. Data analysis using regression and multilevel/hierarchical modelsVol. **1**. Cambridge University Press, pp. 625.

Google Earth Pro 7.3.6. 2022. Teton County, Wyoming, USA. 43.546556, -110.786726. Satellite imagery [accessed 1 June 2022].

Gortázar, J., Alonso, C., and García de Jalón, D. 2012. Brown trout redd superimposition in relation to spawning habitat availability. Ecol. Freshw. Fish **21**: 283–292. doi:10.1111/j.1600-0633.2011.00546.x.

Grolemund, G., and Wickham, H. 2011. Dates and times made easy with lubridate. Journal of Statistical Software **40**: 1–25. doi: 10.18637/jss.v040.i03

Groves, P.A., Chandler, J.A., Alcorn, B., Richter, T.J., Connor, W.P., Garcia, A.P., and Bradbury, S.M. 2013. Evaluating salmon spawning habitat capacity using redd survey data. N. Am. J. Fish. Manag. **33**: 707–716. doi:10.1080/02755947.2013.793628.

Haak, A.L., and Williams, J.E. 2012. Spreading the risk: native trout management in a warmer and less-certain future. N. Am. J. Fish. Manag. **32**: 387–401. doi:10.1080/02755947.2012.678963.

Haak, A.L., Williams, J.E., Isaak, D.J., Todd, A.S., Muhlfeld, C.C., Kershner, J.L., et al. 2010. The potential influence of changing climate on the persistence of salmonids of the Inland West. No. 2010-1236. US Geological Survey, Reston, VA.

Ham, K.D., and Pearsons, T.N. 2000. Can reduced salmonid population abundance be detected in time to limit management impacts? Can. J. Fish. Aquat. Sci. **57**: 17–24. doi:10.1139/f99-175.

Haraldstad, T., Höglund, E., Kroglund, F., Lamberg, A., Olsen, E.M., and Haugen, T.O. 2018. Condition-dependent skipped spawning in anadromous brown trout (*Salmo trutta*). Can. J. Fish. Aquat. Sci. **75**: 2313–2319. doi:10.1139/cjfas-2017-0076.

Hay, D.W. 1984. The relationship between redd counts and the numbers of spawning salmon in Girnock Burn, Scotland. ICES J. Mar. Sci. **22**: 1–4.

Hayes, J.W. 1987. Competition for spawning space between brown (*Salmo trutta*) and rainbow trout (*S. gairdneri*) in a lake inlet tributary, New Zealand. Can. J. Fish. Aquat. Sci. **44**: 40–47. doi:10.1139/f87-005.

Heggberget, T.G., Haukebø, T., and Veie-Rosvoll, B. 1986. An aerial method of assessing spawning activity of Atlantic salmon, *Salmo salar* L., and brown trout, *Salmo trutta* L., in Norwegian streams. J. Fish Biol. **28**: 335–342. doi:10.1111/j.1095-8649.1986.tb05170.x.

Hilborn, R., Amoroso, R.O., Anderson, C.M., Baum, J.K., Branch, T.A., Costello, C., et al. 2020. Effective fisheries management instrumental in improving fish stock status. Proc. Natl. Acad. Sci. U.S.A. **117**: 2218–2224. doi:10.1073/pnas.1909726116. PMID: 31932439.

Homel, K.M., Gresswell, R.E., and Kershner, J.L. 2015. Life history diversity of Snake River finespotted cutthroat trout: managing for persistence in a rapidly changing environment. N. Am. J. Fish. Manag. **35**: 789–801. doi:10.1080/02755947.2015.1044625.

Howell, P.J., and Sankovich, P.M. 2012. An evaluation of redd counts as a measure of bull trout population size and trend. N. Am. J. Fish. Manag. **32**: 1–13. doi:10.1080/02755947.2011.649192.

Kanno, Y., Pregler, K.C., Hitt, N.P., Letcher, B.H., Hocking, D.J., and Wofford, J.E.B. 2016. Seasonal temperature and precipitation regulate brook trout young-of-the-year abundance and population dynamics. Freshw. Biol. **61**: 88–99. doi:10.1111/fwb.12682.

Kiefling, J. 1997. *A History of the Snake River Spring Creek Spawning Tributaries*. Wyoming Game and Fish Department, Cheyenne, WY.

Kovach, R.P., Al-Chokhachy, R., Whited, D.C., Schmetterling, D.A., Dux, A.M., and Muhlfeld, C.C. 2017. Climate, invasive species and land use drive population dynamics of a cold-water specialist. J. Appl. Ecol. **54**: 638–647. doi:10.1111/1365-2664.12766.

Kruschke, J. 2014. Doing bayesian data analysis: a tutorial with R, JAGS, and STAN. Ed. 2. Elsevier. pp. 759

Lestelle, L.C., and Weller, C. 2002. *Summary Report: Hoko and Skokomish River Coho Salmon Spawning Escapement Evaluation Studies 1986–1990*. Point No Point Treaty Council, Kingston, WA.

Lobón-Cerviá, J. 2009. Recruitment as a driver of production dynamics in stream-resident brown trout (*Salmo trutta*). Freshw. Biol. **54**: 1692–1704. doi:10.1111/j.1365-2427.2009.02218.x.

Makowski, D., Ben-Shachar, M.S., Chen, S.H.A., and Lüdecke, D. 2019. Indices of effect existence and significance in the Bayesian framework. Front. Psychol. **10**: 1–14. doi:10.3389/fpsyg.2019.02767. PMID: 30713512.

Maxwell, D., and Jennings, S. 2005. Power of monitoring programmes to detect decline and recovery of rare and vulnerable fish. J. Appl. Ecol. **42**: 25–37. doi:10.1111/j.1365-2664.2005.01000.x.

McNeil, W.J. 1964. Redd superimposition and egg capacity of pink salmon spawning beds. J. Fish. Res. Board Can. **21**: 1385–1396. doi:10.1139/f64-119.

Meredith, M., and Kruschke, J. 2020. HDInterval: highest (posterior) density intervals. R package version 0.2.2.https://CRAN.R-project.org/package=HDInterval.

Meyer, K.A., Schill, D.J., Elle, F.S., and Lamansky, J.A.J. 2003. Reproductive demographics and factors that influence length at sexual maturity of Yellowstone Cutthroat Trout in Idaho. Trans. Am. Fish. Soc. **132**: 183–195. doi:10.1577/1548-8659(2003)132%3c0183:RDAFTI%3e2.0.CO;2.

Muhlfeld, C.C., Taper, M.L., Staples, D.F., and Shepard, B.B. 2006. Observer error structure in Bull Trout Redd counts in Montana streams: implications for inference on true Redd numbers. Trans. Am. Fish. Soc. **135**: 643–654. doi:10.1577/T05-129.1.

Murdoch, A.R., Frady, C.H., Hughes, M.S., and See, K. 2019. Estimating population size and observation bias for spring Chinook Salmon. Conserv. Sci. Pract. **1**: 1–12.

Murdoch, A.R., Herring, C.J., Frady, C.H., See, K., and Jordan, C.E. 2018. Estimating observer error and steelhead redd abundance using a modified Gaussian area-under-the-curve framework. Can. J. Fish. Aquat. Sci. **75**: 2149–2158. doi:10.1139/cjfas-2017-0335.

Oken, K.L., and Essington, T.E. 2015. How detectable is predation in stage-structured populations? Insights from a simulation-testing analysis. J. Anim. Ecol. **84**: 60–70. doi:10.1111/1365-2656.12274. PMID: 25056097.

Parsons, A.L., and Skalski, J.R. 2010. Quantitative assessment of salmonid escapement techniques. Rev. Fish. Sci. **18**: 301–314. doi:10.1080/10641262.2010.513020.

Pregler, K.C., Hanks, R.D., Childress, E.S., Hitt, N.P., Hocking, D.J., Letcher, B.H., et al. 2019. State-space analysis of power to detect regional brook trout population trends over time. Can. J. Fish. Aquat. Sci. **76**: 2145–2155. doi:10.1139/cjfas-2018-0241.

Quinn, T.P.. 2005. The Behavior and Ecology of Pacific Salmon and TroutVol. **1**. University of Washington Press, Seattle, WA. pp. 320

Reed, J.M., and Blaustein, A.R. 1997. Biologically significant population declines and statistical power. Conserv. Biol. **11**: 281–282. doi:10.1046/j.1523-1739.1997.96339.x.

Ricker, W.E. 1954. Stock and recruitment. J. Fish. Res. Board Can. **11**: 559–623. doi:10.1139/f54-039.

Rieman, B.E., and McIntyre, J.D. 1996. Spatial and temporal variablilty in bull trout redd counts. N. Am. J. Fish. Manag. **16**: 132–141. doi:10.1577/1548-8675(1996)016⟨0132:SATVIB⟩2.3.CO;2.

Rieman, B.E., and Myers, D.L. 1997. Use of redd counts to detect trends in bull trout (*Salvelinus confluentus*) populations. Conserv. Biol. **11**: 1015–1018. doi:10.1046/j.1523-1739.1997.96228.x.

Schindler, D.E., and Hilborn, R. 2015. Prediction, precaution, and policy under global change. Science **347**: 953–954. doi:10.1126/science.1261824. PMID: 25722401.

Stier, A.C., Essington, T.E., Samhouri, J.F., Siple, M.C., Halpern, B.S., White, C., et al. 2022. Avoiding critical thresholds through effective monitoring. Proc. R. Soc. B. **289**: 10.

Su, Y., and Yajima, M. 2021. R2jags: using R to run 'JAGS'. R package version 0.7-1. Available from https://CRAN.R-project.org/package=R2jags.

Taniguchi, Y., Miyake, Y., Saito, T., Urabe, H., and Nakano, S. 2000. Redd superimposition by introduced rainbow trout, *Oncorhynchus mykiss*, on native charrs in a Japanese stream. Ichthyol. Res. 149–156. doi:10.1007/BF02684235.

Vehtari, A., Gelman, A., and Gabry, J. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat. Comput. **27**: 1413–1432. doi:10.1007/s11222-016-9696-4.

Wagner, T., Irwin, B.J., Bence, J.R., and Hayes, D.B. 2013. Detecting temporal trends in freshwater fisheries surveys: statistical power and the important linkages between management questions and monitoring objectives. Fisheries, **38**: 309–319. doi:10.1080/03632415.2013.799466.

Walsworth, T.E., and Schindler, D.E. 2015. Coho salmon escapement and trends in migration timing to a data-poor river: estimates from a Bayesian hierarchical model. Can. J. Fish. Aquat. Sci. **72**: 1807–1816. doi:10.1139/cjfas-2014-0554.

Walsworth, T.E., and Schindler, D.E. 2016. Long time horizon for adaptive management to reveal predation effects in a salmon fishery. Ecol. Appl. **26**: 2693–2705. doi:10.1002/eap.1417. PMID: 27875003.

Wickham, H., Averick, M., Bryan, J., Chang, W., D'Agostino McGowan, L., François, R., et al. 2019. Welcome to the tidyverse. J. Open Source Softw. **4**: 6. doi:10.21105/joss.01686.

Youngflesh, C. 2018. MCMCvis: tools to visualize, manipulate, and summarize MCMC output. J. Open Source Softw. **3**: 3. doi:10.21105/joss.00640.

Youngson, A.F., Piertney, S.B., Thorley, J.L., Malcolm, I.A., and Soulsby, C. 2011. Spatial association of nest construction by brown trout *Salmo trutta*. J. Fish Biol. **78**: 713–725. doi:10.1111/j.1095-8649.2010.02883.x. PMID: 21366568.

Can. J. Fish. Aquat. Sci. Downloaded from cdnsciencepub.com by UNIVERSITY OF WYOMING on 09/28/23
For personal use only.