Abstract
The primary conservation prioritization tool for spring Chinook Salmon (*Oncorhynchus tshawytscha*), population viability analysis, is often conducted with biased spawner abundance data with no associated statistical uncertainty or error. This study estimated observation error of surveyors counting redds in two spring Chinook Salmon populations where hatchery supplementation is implemented as a conservation tool. Habitat complexity, redd density and the amount of observer experience were important in estimating error rates. Increases in both habitat complexity and experience reduced net error rates. Conversely, net error rates increased as redd density increased. Unbiased estimates of redd abundance were generated and converted to spawner abundance using population specific redd expansion factors. The precision (i.e., coefficient of variation [CV]) of spawner abundance estimates were similar in the Wenatchee (natural CV = 5%; hatchery CV = 6%) and Methow (natural CV = 5%; hatchery CV = 2%) watersheds because average net error rates were similar (Wenatchee = −0.1512; Methow = −0.1748). This study addresses a criticism of population viability analysis (i.e., parameter uncertainty) that should result in more scientifically defensible conservation priorities and recommendations that can be implemented with greater certainty.

**KEYWORDS**
observation error, population viability analysis, redd counts, spawner abundance, spring Chinook Salmon

1 | INTRODUCTION
Redd counts are widely used across the western United States to monitor the status and trend of anadromous salmonid populations (Gallagher, Hahn, & Johnson, 2007). For populations listed under the United States Endangered Species Act, Crawford and Rumsey (2009) recommended that methods used to estimate spawner abundance are unbiased and precise (i.e., coefficient of variation [CV] <15%). Spawner abundance estimates derived from redd counts are often treated to be unbiased (NWFSC, 2015), but this assumption is rarely tested. In Washington State, the majority (79%) of the extant spring Chinook Salmon (*Oncorhynchus tshawytscha*) populations are monitored using redds counts with no accompanying measure of precision or bias (WDFW, unpublished data). In a recent status review of salmon populations listed under the United States Endangered Species Act in the Pacific Northwest, the trends of overall risk for spring Chinook Salmon evolutionary significant units were generally unchanged since the last status review (NWFSC, 2015). The authors acknowledge in the report that population abundance estimates or the number of
adult spawners used in the analysis were derived from data collected by various agencies (i.e., federal, state, and tribal fish managers) and did not include a measurement of variability due to observation error.

Various methodologies have been used to count spring Chinook Salmon redds (Chapman, Peven, Giorgi, Hillman, & Utter, 1995) and few studies have provided estimates of observation error. Chasco et al. (2014) used state-space models to estimate the accuracy and precision of four types of Chinook Salmon spawner abundance estimators. Multi-pass census redd surveys were reported to have similar accuracy (i.e., unbiased) and precision as mark-recapture estimates (Chasco et al., 2014). In that study, census surveys were only conducted 3–4 times per season and the authors concluded that multi-pass census redds surveys may provide adequate estimate of true population size, at a third of the cost, provided additional information about the population is known (i.e., proportion of females and prespawn mortality). More recently, models were recently developed to predict steelhead (Oncorhynchus mykiss) redd observer error rates using river reach specific covariates that included surveyor experience, habitat complexity, observed redd density, river condition, and size (Murdoch, Herring, Frady, See, & Jordan, 2018). In that study, steelhead redd observers found only between 48 and 58% of the known redds. Using a modified Gaussian area-under-the-curve methodology, redd count data and estimated error rates were used to generate an unbiased estimate of steelhead redds with an associated estimate of precision.

The expansion of redd counts to a spawner abundance estimate typically involve multiplying the redd count by a fish per redd (FPR) value. While considerable effort and resources are used to generate an estimate of redd abundance, FPR values often lack a similar level of scientific rigor (Murdoch, Pearsons, & Maitland, 2010). Because spawner abundance is the parameter of interest (e.g., PVA, recovery, and escapement goals) incorporating the uncertainty associated with the FPR into the population estimate is an often overlooked step. The objectives of this study are to (a) measure spring Chinook Salmon redd observer error (b) develop predictive models to estimate observer error, (c) use predictions of observer error to estimate the total number of spring Chinook Salmon redds, and (d) estimate both spawner abundance and precision due to observation error.

2 METHODS

2.1 Study area

The Wenatchee (3,439 km²) and Methow (4,727 km²) watersheds drain the east slope of the Cascade Mountains with 207 and 239 km of streams accessible to anadromous salmonids, respectively (Mullan, Williams, Rhodus, Hillman, & McIntyre, 1992). Hydrologic discharge patterns in both watersheds are predominantly driven by snowmelt and ground water, but glacier melt during the summer low flow periods are of greater influence in the Wenatchee (2,546 acres) than the Methow (519 acre; Mullan et al., 1992). Spring Chinook Salmon typically spawn during minimum base flow periods depending on snowpack levels. Each spring Chinook Salmon population has a branched spatial structure comprised of several major spawning streams.

2.2 Study design

Mark-recaptures studies are expensive and cannot be performed for all populations because a weir or dam is required in order to mark a representative sample of the population. Our study design compared intensive census redd surveys (i.e., twice a week) to the more typical multi-pass census redd surveys (i.e., once every 7–10 days) currently used throughout the Upper Columbia Basin. Study reaches, based on historical reach breaks, were selected to maximize the variability of possible covariates in order to facilitate the development of redd observer error models that can be used to correct any observed bias in redd counts (Figure 1). The advantages of this approach are that a mark-recapture study is not required and the model can be used to estimate error rates in a much larger geographic area.

2.3 Redd surveys

2.3.1 Intensive census redd surveys

Comprehensive multiple pass or census redd surveys of study reaches were conducted in an upstream direction in the Wenatchee watershed and downstream in the Methow watershed. Redd observers with the greatest reach-specific experience surveyed twice a week beginning in early August, before any redds were observed, until no new redds were found (i.e., end of September). Spring Chinook Salmon redds or any features (i.e., hydraulically formed, incomplete or test redd or a redd of another species) that could be misidentified as a spring Chinook Salmon redd were georeferenced using hand held global positioning (GPS) devices, annotated on aerial photographs and on hand drawn maps in cases of high density spawning. Incomplete spring Chinook Salmon redds or test redds were differentiated from complete redds based on the morphology (e.g., incorrect substrate composition, small size, bowl and tail not well defined) and lack of a guarding female. Census surveyors did not attach flagging to nearby vegetation along the streambank or place anything within the redd (i.e., colored stones) that otherwise would identify that location as a redd. The total census count of redds in a study reach was assumed to be without error (i.e., the true number of redds). This assumption was tested by comparing census redd
counts to mark-recapture estimates of spawning female spring Chinook Salmon. In the Wenatchee Basin, female spring Chinook Salmon previously tagged with passive integrated transponder (PIT) tags were detected on redds by census surveyors (Hughes & Murdoch, 2017). Subsequently, all female spring Chinook Salmon carcasses recovered during redd surveys were examined for marks. The estimate of female Spring Chinook Salmon spawner abundance was calculated using a modified Chapman estimator:

\[ \hat{N}_i = \frac{(M_i + 1)(C_i + 1)}{(R_i + 1)} - 1, \]

where \(M_i\) is the number of marked females detected on redds, \(C_i\) is the number of carcasses recovered, and \(R_i\) is the number of carcasses recovered that were marked (Seber, 1973). The variance of the estimate was calculated using:

\[ \text{var}(\hat{N}_i) = \frac{(M_i - 1)(C_i - 1)(M_i - R_i)(C_i - R_i))}{((R_i + 1)(R_i + 1)(R_i + 2))}. \]

2.3.2 | Naïve redd surveys

All naïve surveyors received both classroom and field training, albeit limited, prior to the start of the study. In the training, we emphasized the use of redd characteristics (i.e., size, substrate and morphology) to differentiate complete redds from test redds. Naïve surveyors conducted weekly surveys
of study reaches in a typical downstream direction and recorded both the start and stop time of the survey (i.e., surveyor effort). Similarly, all completed spring Chinook Salmon redds were georeferenced using hand held GPS devices, annotated on aerial photographs and in cases of high density spawning or superimposition delineated using hand drawn maps. Redds identified by naïve surveyors were not flagged or otherwise marked because multiple naïve surveyors conducted surveys in each study reach. Census and naïve surveyors were instructed not to discuss the results of any survey throughout the study period.

### 2.4 | Model covariates and error rates

A series of habitat measurements were taken to assess the environmental variation within and among survey reaches. Ten equidistant transects were measured perpendicular to the flow in each study reach to estimate stream width and water depth. Daily mean river discharge was measured at the nearest stream gauging station. We used the variation in the thalweg depth profile to quantify spawning habitat heterogeneity (Madej, 1999). The mean CV in thalweg depth profile was based on three redd locations per reach randomly selected from the previous (Wenatchee) or the current (Methow) year. Thalweg depth was measured (N = 100) every three meters or a total of approximately 300 m with the redd location as the midpoint. Several other potential habitat complexity metrics were collected by census surveyors (i.e., pieces of large woody debris, number of log jams, and number of gravel bars) but were significantly correlated (Spearman Rank Correlation; \( p < .05 \)) with the mean reach thalweg CV and not included in the model development to prevent collinearity. Thalweg CV was retained as possible covariate because the variability in streambed elevation is the product of both geomorphic and hydrologic processes, including habitat condition, and is easily reproducible. Water clarity can differ within and across study reaches and streams as a result of glacial till. We used a secchi disc to estimate water visibility at the start and end of each survey by placing it in approximately 0.5 m of water with a laminar flow pattern. The census observer would walk upstream to a point when the white and black sections of the disc were no longer discernible. The distance between the observer and secchi disc was measured on the stream bank using a tape measure or range finder. The mean distance of the two measurements was used as an index of horizontal water visibility for that survey, the mean distance across surveys during the season was used as a covariate for that study reach. Redd density (redds/km) was based on the number of redds or redd-like features identified by the naïve observer divided by reach length.

We quantified observer experience based on the number of days per week of redd survey experience for an entire spawning season,

\[
\text{Experience} = \sum_i a_i b_i + \sum_i 0.1 c_i,
\]

where \( i \) indexes the different salmon species, up to a maximum of \( n \), for which the observer has conducted redd surveys, \( a_i \) is the number of years surveys were conducted, \( b_i \) is the number of days per week surveys were conducted and \( c_i \) is the days redd surveys were conducted if not conducted as part of a multi-pass redd survey design. For example, a redd observer who conducted steelhead surveys three times week for two complete seasons and Chinook Salmon surveys twice a week for five complete seasons would receive an experience score of 16. Partial or incomplete seasons were prorated based on a 10-week spawning season (i.e., each day equal 0.1 points). In addition, surveyor effort was calculated for each naïve surveyor by reach and was standardized to be number of minutes spent surveying per kilometer of stream.

We calculated observer error rates at the end of the spawning season by comparing the redd locations of census and naïve redd observers. We defined the commission error rate or the proportion of features identified that were not spring Chinook Salmon redds (e.g., hydraulic scour and depositional areas, test redds, or redds of other species) as:

\[
C_i = \frac{E_i}{F_i},
\]

where \( C_i \) is the error rate of commission, \( E_i \) is the number of features erroneously identified as redds, and \( F_i \) is the total number of redd-like features found by the naïve redd observer \( i \). This corresponds to one minus the accuracy rate described in (Murdoch et al., 2018). We defined the omission error rate or the proportion of actual redds missed by naïve surveyors as:

\[
O_i = \frac{M_i}{V_{cen}},
\]

where \( O_i \) is the omission rate, \( M_i \) is the number of redds found by the census surveyors but missed by the naïve surveyor \( i \), and \( V_{cen} \) is the number of visible redds identified by the census surveyors. This corresponds to one minus the efficiency rate described in (Murdoch et al., 2018). Ideally, commission and omission rates would be zero. The total bias of a naïve surveyor is a combination of commission and omission errors, which may cancel each other out. As a one-step alternative, we explored modeling the overall net error, following Thurow and McGrath (2010) which we defined as:

\[
NE_i = \frac{E_i - M_i}{V_{cen}},
\]
where net error, $NE_i$, greater than 0 implies an overall over counting of redds, while net error less than 0 implies an undercounting. By this definition, an estimate of $V_{cen}$ (i.e., the truth) can be made by using the number of features identified by the naïve surveyor, $F_i$, and the estimate of net error:

$$\hat{V}_{cen} = \frac{F_i}{NE_i + 1}.$$  
(7)

### 2.5 Model selection and performance

Models for omission and commission used a binomial error structure with a logit link, while models for net error were linear regressions with a Gaussian error structure. All models used the same potential covariates, $X$:

$$\text{logit}(C_i) = X\beta_c + e_i,$$

$$\text{logit}(O_i) = X\beta_o + f_i,$$

$$NE_i = X\beta_n + g_i,$$

$$e_i \sim N(0, \sigma_e^2),$$

$$f_i \sim N(0, \sigma_f^2),$$

$$g_i \sim N(0, \sigma_n^2).$$  
(8)

We normalized all covariates to have a mean of zero and standard deviation of one. Stream width and depth were highly correlated with average discharge ($r = 0.77$ and 0.82). We assumed discharge to have more of an effect on observer error than width or depth; therefore, we dropped them from our list of possible covariates. We retained average discharge, mean thalweg CV, mean visibility, observed redd density (features / km), surveyor effort and experience as potential covariates in each model. Because we believed there to be a saturating effect of experience (i.e., at some level of experience, an additional season will not change the error rates) we treated experience levels as a factor. Experience was weakly correlated errors of omission ($r = -0.30$, $p < .05$) and commission ($r = -0.24$, $p < .10$), but not net error ($r = 0.19$, $p = .16$). However, when observers with no experience were excluded, experience was not related to any error rates ($p > 0.20$). Because redd observers with an experience value less than one (i.e., rookie) may have greater error rates then observers with more experience, observers were categorized as rookies ($< 1$) or experienced observers ($> 1$). We developed a list of candidate models by constructing all possible additive combinations of these six covariates, including an intercept-only model. After fitting each of these candidate models, we recorded the Akaike information criterion corrected for small sample size ($\text{AICc}$), and the Akaike weight ($w_i$) for each model. The sum of Akaike weights for all models containing a particular covariate provides a measure of the relative importance of that covariate (Burnham & Anderson, 2002).

We compared several different error rate predictions. Following the advice from Burnham and Anderson (2002), we calculated the predictions from modeling averaging all possible models. Because all models with a $\Delta\text{AICc} < 2$ were considered to be strongly supported by the data, we predicted error rates by model averaging only this subset of models. We also tried predicting error rates using only the full global model, ignoring all the model selection criteria. Finally, we calculated predictions from models that only incorporated the top three covariates in the net error model. The goal was to evaluate the impact of using all the possible covariates versus using only a subset based on the model selection results. For managers hoping to use these models, we wanted to quantify the impact of using only a subset of covariates, which may be logistically and fiscally easier to collect compared to all model covariates.

In order to test how well these models would predict error rates for other surveys, we performed a leave-one-out cross validation by leaving out all naïve surveys for a particular reach in a particular year. We fit the global model to that subset of the data, performed the model selection procedure described above, and recorded model averaged predictions of omission, commission and net error rates. We also predicted true number of redds for each naïve survey conducted in that particular reach, using the model averaged predictions of the full suite of candidate models. We did this sequentially, leaving out each reach / year combination. We then assessed the error model's performances by calculating the mean bias in predicted error rates, the root mean square error (RMSE) of predicted error rates, and the 95% coverage probability (how often the true error rate fell within the 95% confidence interval of the predicted rate). We also assessed two methods of using error rates to estimate the true number of redds: one that corrected counts by using both predicted commission and omission rates, and one that corrected counts by using only the predicted net error rate. This assessment also consisted of examining the mean predicted bias, the RMSE and the 95% coverage probability of the predicted number of redds. All statistical analyses were conducted using R software (R Core Team, 2015).
2.6 Estimating spawner abundance using redd survey and sex data

The error models described above provide a method to estimate the true number of redds from typical redd surveys. However, to estimate spawner abundance, we also needed an estimate of how many spawners per redd were present. Spawners per redd is a combination of how many redds a typical female salmon builds, and the sex ratio of males to females. The sex ratio of the Wenatchee population was estimated at Tumwater Dam and at Wells Dam for the Methow population. The sex of every fish examined was determined using portable ultrasound units. The error associated with sex ratios was based on comparing sex assignments for those fish that were sampled at Tumwater Dam and, subsequently, collected and spawned for hatchery broodstock operations (i.e., sex is confirmed during spawning operations). Using sex ratios as a FPR value also assumes the average number of redds constructed by a female is known. We reexamined this assumption using a larger time series of data than previously reported (2005–2007) by Murdoch, Pearsons, and Maitland (2009). Female spring Chinook Salmon, associated with a redd, were scanned for PIT tags in the Wenatchee Basin between 2005 and 2013. Because each PIT tag has a unique code, we calculated the average (SD) number of redds assigned to a female spring Chinook Salmon.

Redd observer error models developed in 2012 and 2013 were applied to data collected during typical spring Chinook Salmon redd surveys in 2014. Redd observers conducted weekly surveys of the major spawning reaches in both the Wenatchee and Methow watersheds. We used reach specific covariates and model averaged the top models to predict net error. Unfortunately, visibility was not measured in 2014, but it was not among the top models ($\Delta$AICc <2) and based on the comparison of various predictive models, the difference in net error estimates was practically identical if model averaging included all the models or only the top ones. We estimated spring Chinook Salmon redd abundance and variance in major spawning areas by starting with the total number of redds found in each reach in 2014, and then applying Equation (7) and using the delta method (Doob, 1935). We assumed that each reach was independent (i.e., covariance of 0). The net error model was not used for minor spawning tributaries because observer error was assumed to tend toward zero because of the characteristics of minor tributaries: relatively small size (< 10 m width), low discharge (< 5 CFS) and shallow water depth (<0.2 m), which were outside the range of the data set used to develop the net error model. Census redd counts in the minor spawning areas were assumed to be without error and were added to the respective estimated number of redds in each major spawning stream. We converted reach specific estimates of redd abundance to spawner abundance using the sex ratio of the population as a redd expansion factor or FPR values and reach specific carcass data to determine the proportion of hatchery and natural origin spawners (Murdoch et al., 2010). A population spawner estimate was generated from the sum of all reaches using:

$$\text{Number of spawners}_i = \sum_{i=1}^{n} \frac{\bar{x}_i}{y_i} \times \left(1 + \frac{m_i}{f_i}\right)$$

where $n$ is the number of reaches in population $i$, $\bar{x}$ is the estimated number of redds, $y$ is the average number of redds per female, $m$ is the number of males in the sample and $f$ is the number females in the sample. Standard errors of spawner estimates were calculated by again using the delta method, assuming that estimates of total redds, the average number of redds per female and the sex ratio were all independent of one another.

3 RESULTS

Mark-recapture estimates of female spawners were conducted in major spawning reaches (84% of all redds in 2012 and 86% of all redds in 2013) in the upper Wenatchee Basin (WDFW, unpublished data). Census redd counts fell within the 95% confidence interval of the mark-recapture estimate indicating census redd counts represented the true female population size (Table 1), supporting the assumption that census redd counts were the true number of redds available to be counted. While this assumption was only tested in the Wenatchee Basin, we have no compelling reason why it would not also be true in the Methow Basin.

Between 2012 and 2013, 56 naïve observers conducted weekly surveys in 13 study reaches in the Wenatchee and Methow watersheds (Supporting Information). Redd survey experience of naïve observers was lower and more variable (mean = 24.3, $SD = 17.3$) than that of the census surveyors (mean = 31.2, $SD = 15.9$). Owing to a prolonged redd life (mean = 38 d, $SD = 12$), identification of a redd on the first survey after it was constructed was not as important as it is for species with a much shorter redd life (i.e., steelhead). Hence, error rates were calculated over the entire spawning season. The mean ($SD$) commission error rate of redd observers was 0.094 (0.063) and was smaller and less variable than mean redd omission error rate (mean = 0.226, $SD = 0.121$). Because errors of commission were far less common than errors of omission, the mean net error rate (mean = –0.141, $SD = 0.148$) was less than zero (Supporting Information). A majority of naïve redd observers (88%) underestimated (i.e., net error < 0) the number of spring Chinook Salmon redds. The majority of unidentified redds were either overlooked (37%) by observers or were superimposed on an existing redd (32%). Test redds
or those identified as hydraulic features represented 27 and 4% of the unidentified redds, respectively. Overestimates of redd abundance were not only less common (12%), but on average were also of lesser magnitude (7%) than underestimates (−17%). The majority of misidentified redds were due to test digs (71%) and redd superimposition (23%) with only a small proportion of hydraulic features (6%) identified as redds.

4 | MODEL SELECTION AND PERFORMANCE

The CV of thalweg depths (i.e., index of habitat complexity) was the most important covariate across all three error models. As habitat complexity increased, omission rates increased, commission rates decreased, and net error rates shrank. Surveyor effort was also important for omission and commission models, with higher effort leading to lower omission and higher commission rates. However, the effects nearly offset each other, leaving it relatively unimportant for predicting net error. The level of observer experience and observed redd density were important covariates for omission and net error rates, but not for commission errors. Inexperienced observers were more likely to have higher omission rates and lower net error rates, while higher observed redd densities tended to lead to lower omission rates and had a positive effect on net error rates. Increased visibility decreased commission error rates but had negligible effects on omission and net error rates. Discharge had a small and insignificant effect on all error rates (Figure 2 and Table 2).

The different combinations of covariates or which subset of models were used in model averaging made very little difference in error rate predictions. For all three error models, the correlation coefficients between various predictions were all over 0.96, and most were greater than 0.98. In particular, the model averaged predictions from all candidate models and those with a ΔAICc less than two were greater than 0.999.

Based on the leave-one-out cross validation, the models performed best when predicting observer omission, next best for net error, and poorest for commission (Supporting Information). The RMSE was between 6 and 15% for all error rates, and the average bias was very low, so although predictions were not always accurate, they were generally unbiased (Supporting Information).

All observer error models led to better estimates of total redds than assuming surveys were a census of redds. Observers usually undercounted redds by between 2 and 20 redds depending on the watershed and year, while the average prediction bias for the net error model was −0.3 redds across all years and watersheds, and 1.4 redds when using the omission / commission models (Supporting Information). Both methods failed to capture all the appropriate uncertainty, as reflected in the inadequate 95% coverage probabilities or a measure of how frequently the 95% confidence interval contained the true value and should be close to 95%. Across both watersheds and years, the net error method had coverage probability of 52%, while the omission / commission method had a coverage probability of 32%.

| Year | Stream | Census redd count | Mark-recapture estimates |
|------|--------|------------------|-------------------------|
|      |        | Lower 95% | Point | Upper 95% |
| 2012 | Chiwawa | 610        | 522  | 642  | 761  |
|      | Nason  | 353        | 305  | 385  | 464  |
| 2013 | Chiwawa | 526        | 477  | 566  | 655  |
|      | Nason  | 193        | 125  | 175  | 225  |
4.1 Estimating spawner abundance using redd survey and sex data

In 2014, the mean predicted net error rate in the Methow (−0.151) watershed was slightly closer to zero compared to the Wenatchee (−0.175) watershed (Table 3) in part due to higher redd densities and lower habitat complexity (i.e., thalweg CV). Using the net error model to correct for observation bias, we estimated that 101 and 104 redds were not counted during surveys in the Methow and Wenatchee watersheds, respectively, about 10% of the total redds. Estimates of redds within each major spawning stream in the Methow and Wenatchee watersheds had an average CV of 4.9% (Table 3). The CV for natural and hatchery origin spawner estimates in the Wenatchee and Methow were between 2 and 6% (Table 4), below the recommended precision level (CV =15%) for ESA listed populations (Crawford & Rumsey, 2009).

5 DISCUSSION

This study resulted in a relatively simple model which can provide unbiased estimates of redds that can be applied with little extra data required (i.e., thalweg CV). Although formally all covariates should be collected and utilized, the exclusion of visibility, discharge and effort made very little difference in net error predictions. From a practical standpoint, the collection of thalweg CV, experience and redd density data is sufficient to provide a reliable estimate of the observer’s net error. While the heterogeneity of spring Chinook Salmon spawning habitats included in the model was maximized, the study was limited regionally to the Upper Columbia watershed. However, the range of covariates measured within the Upper Columbia covers the measured range

| Population | Stream | Mean estimated net error | Number of redds |
|------------|--------|--------------------------|-----------------|
|            |        |                          | Observed | Estimated | SE  | CV   |
| Methow     | Chewuch| −0.1841                  | 239      | 290.9     | 4.545| 1.6% |
| Methow     |        | −0.08163                 | 763      | 781.2     | 10.68| 1.4% |
| Twisp      |        | −0.215                   | 138      | 169.3     | 3.501| 2.1% |
| Total      |        | −0.1512                  | 1,140    | 1,241     | 12.12| 1%   |
| Wenatchee  | Chiwawa| −0.1247                  | 485      | 536.8     | 8.298| 1.5% |
| Little Wenatchee |        | −0.2874                 | 25       | 34.41     | 2.545| 7.4% |
| Nason      |        | −0.2142                  | 115      | 146.1     | 3.269| 2.2% |
| Wenatchee  |        | −0.1477                  | 23       | 27.07     | 3.773| 13.9%|
| White      |        | −0.2301                  | 26       | 33.62     | 3.296| 9.8% |
| Total      |        | −0.1748                  | 674      | 778.1     | 10.54| 1.4% |

Abbreviations: CV, coefficient of variation; SE, standard error.
in streams across Washington (Entiat, Tucannon), Oregon (John Day, Grande Ronde) and Idaho (Lemhi, South Fork Salmon, Yankee Fork; ISEMP/CHaMP, 2017; Figure 3). This suggests that these error models could be applied to any redd surveys that followed a similar protocol (e.g., a single observer conducting weekly surveys). The utility of the model could be expanded by collecting similar data from less similar types of habitats.

The estimated net error rates for the eight major spawning areas in this study were not highly variable (CV = 8%) and the mean error rate (−0.14) indicates that a substantial proportion of redds were not counted

| Population | Stream | Observed Redds | Estimated Redds (SE) | Redds per female (SE) | Fish per redd (SE) | Natural Spawners (SE) | Hatchery Spawners (SE) |
|------------|--------|----------------|----------------------|-----------------------|--------------------|------------------------|------------------------|
| Methow     | Chewuch| 239            | 291 (4.5)            | 1.02 (0.003)          | 1.806 (0.063)      | 188 (16.6)             | 295 (17.8)             |
| Methow     | Methow | 763            | 781 (10.7)           | 1.02 (0.003)          | 1.806 (0.063)      | 263 (20.1)             | 1,115 (26.5)           |
| Methow     | Twisp  | 138            | 169 (3.5)            | 1.02 (0.003)          | 1.806 (0.063)      | 100 (14.6)             | 154 (14.9)             |
| Methow     | Total  | 1,140          | 1,241 (12.1)         | 1.02 (0.003)          | 1.806 (0.063)      | 550 (29.9)             | 1,564 (35.2)           |
| Wenatchee  | Chiwawa| 485            | 537 (8.3)            | 1.02 (0.003)          | 2.029 (0.241)      | 581 (36.1)             | 485 (32.5)             |
| Wenatchee  | L. wen.| 25             | 34 (2.5)             | 1.02 (0.003)          | 2.029 (0.241)      | 55 (11)                | 14 (8.1)               |
| Wenatchee  | Nason  | 115            | 146 (3.3)            | 1.02 (0.003)          | 2.029 (0.241)      | 206 (18.7)             | 84 (15.2)              |
| Wenatchee  | Wenatchee| 23            | 27 (3.8)             | 1.02 (0.003)          | 2.029 (0.241)      | 42 (9)                 | 12 (5.1)               |
| Wenatchee  | White  | 26             | 34 (3.3)             | 1.02 (0.003)          | 2.029 (0.241)      | 50 (11.4)              | 17 (9.2)               |
| Wenatchee  | Total  | 674            | 778 (10.5)           | 1.02 (0.003)          | 2.029 (0.241)      | 935 (44.5)             | 611 (38.3)             |

**FIGURE 3** Violin plots of model covariates from data in this study with similar covariates collected across the interior Columbia River basin by the Columbia Habitat Monitoring Program (ISEMP/CHaMP, 2017)
rates has also been investigated for bull trout (*Salvelinus confluentus*). Dunham, Rieman, and Davis (2001) measured error with mostly novice observers, while Muhlfeld, Taper, Staples, and Shepard (2006) used highly experienced redd observers (mean experience = 10 years). While redd observer experience was not quantified and incorporated into a modeling framework (e.g., as in this study) or assessed in importance relative to other sources of observation error, both studies concluded experienced redd observers should be used to minimize the variability in error rates. High variability in redd observer experience (CV = 71%) in this study was intended to assist in modeling and assessing the relative influence of experience. Because redd observer experience is very important (i.e., one of top covariates in omission and net error models), quantifying experience is important. Unfortunately, the influence of other covariates on error rates confounded our ability to develop a robust relationship between surveyor experience and error rates. Hence, the optimal number of years of experience is still unknown.

Covariates that were important in influencing spring Chinook error rates were also important for steelhead redd survey error rates (Murdoch et al., 2018). However, important differences in these studies are worth noting. This study only evaluated a redd survey protocol that included a single observer walking downstream. It is possible that if different redd survey protocols were evaluated (e.g., two observers) the covariates and their relative influence would be different as well as the error rates. Lastly, because spring Chinook Salmon redds exhibited such a long life (e.g., 88% of redds still visible at the end of the survey period), error rates were not based on conditions during each weekly survey like steelhead (Murdoch et al., 2018), but rather the entire spawning period.

Error rates using a weekly survey protocol were based on intensive census surveys conducted twice a week. If the frequency of conducting redds counts is so important in reducing bias, then other redd survey protocols (e.g., 3–4 per season or single peak counts) would result in greater bias consistent with the results reported in Chasco et al. (2014). Error rates in minor spawning areas were also not included in the study design (i.e., we assumed no error). Habitat characteristics of small streams were different enough from the major spawning streams such that the methodology was not transferable. While it is possible that redds in these small streams were missed, the number of missed redds, assuming a similar 10% error rate, and their overall influence on the total redd estimate (0.2% = Wenatchee; 1.3% = Methow) would be negligible.

Redd expansion factors are how estimates of redds are converted to spawners. In this study, the sex ratio of the population monitored at a downstream location (i.e., run escape-ment) was generated using portable ultrasounds and later

(Wenatchee = 15%; Methow = 12%). Hence, these results suggest that raw spring Chinook Salmon redd counts could adequately represent trends in abundance, assuming observation bias is relatively constant. However, if the observation bias differs substantially within years or between years due to different redd survey protocols, new or rookie observers or differences in redd densities, then raw redd counts will not appropriately capture trends in abundance. The differences in net error estimates between different tributaries will also mask the true extent of population distribution across its spatial extent (i.e., numbers of spawners returning to each tributary within a population), one of the criteria for delisting salmon (McElhany, Ruckelshaus, Ford, Wainwright, & Bjorkstedt, 2000). Given that redd density and surveyor experience, covariates used to predict net error, are not constant through any data time series, the annual observation bias is also variable. This work represents a method to estimate redd abundance, with uncertainty, which would lead to bias is also variable. This work represents a method to estimate redd abundance, with uncertainty, which would lead to bias is also variable. This work represents a method to estimate redd abundance, with uncertainty, which would lead to bias is also variable. This work represents a method to estimate redd abundance, with uncertainty, which would lead to
adjusted by visually determining the sex at hatchery broodstock spawning operations or PIT tagged carcasses when recaptured on the spawning grounds. However, differential prespawn mortality (based on gender or origin) between the time of sampling and spawning would bias the sex ratio of the spawning population and subsequently the estimates of spawners. We incorporated uncertainty in the sex ratio and the number of redds each female constructs due to the sample size contributing to each variable. We did have the luxury of a data set that tracked how many redds a number of tagged females built, but given the consistency in the average redds per female and the variation around that average for at least 6 years, these numbers may be transferable to other systems. Ideally, sex ratio-based redd expansion factors should be based on the spawning population. Using carcasses to generate the sex ratio would capture the interannual variability often observed, but if raw carcass data is not corrected for sex and sized-based recovery bias, the raw sex ratio will produce biased estimates of spawners (see Murdoch et al., 2010).

Unfortunately, most salmon PVAs or life cycle models are parameterized with spawner data generated from raw or expanded redd counts, use constant redd expansion factors and do not include any measure of uncertainty (e.g., Beamesderfer et al., 1997). When these analyses are then used to estimate the probability of quasi-extinction, there are two potential problems. First, if the bias in redd counts or redd expansion factors is not constant through time, the estimate of population growth rate will be incorrect. Second, without an estimate of observation error, the person conducting a PVA will assign all the annual variation in redds as coming from process error, which contributes directly to the probability of quasi-extinction (observation error does not). Observation error among redd survey protocols is variable (Chasco et al., 2014; Courbois et al., 2008; Liermann, Rawding, Pess, & Glaser, 2015) and covariates that vary through time (e.g., redd density and observer experience) influence error within a given protocol (e.g., this study). It is unclear to what extent this unknown variability in observation error within a single time series has biased population growth curves and ultimately predictions of quasi-extinction risk probability. However, quantifying observation error is the first step to incorporating error into a modeling framework.

ACKNOWLEDGEMENTS

We would like to thank the numerous biologists and technicians who walked many miles of streams in order to complete this study. M. Johnson, J. Jorgensen, and R. Zabel provided much appreciated comments on an earlier draft manuscript. This study was funded with a grant from the Bonneville Power Administration (Project # 2010-034-00). Chelan, Douglas and Grant County public utility districts also provided additional financial support.

CONFLICT OF INTEREST

The authors declare no potential conflict of interest.

AUTHORS CONTRIBUTIONS

All authors were involved in the development of the experimental design. C.H.F and M.S.H led field and data collection; K.S. conducted error modeling; A.R.M. wrote the first draft.

DATA ACCESSIBILITY

Descriptive statistics of study reaches and possible covariates are listed Appendix S1. Summary statistics of observer error rates by watershed are in Appendix S2. Top candidate redd observer models are listed in Appendix S3. Summary statistics comparing observed and predicted error rates can be found in Appendix S4. Plots of observed versus predicted error rates and weighted correlation coefficients are in Appendix S5. Summary statistics comparing estimated redds using net error or omission/commission models are found in Appendix S6. Authors are solely responsible for the content and functionality of these materials. Queries (other than absence of the material) should be directed to the corresponding author.

ETHICS STATEMENT

The tagging and monitoring of fish and redds were approved by the Washington Department of Fish and Wildlife and National Oceanic Atmospheric Administration - Fisheries (West Coast Region) under an Endangered Species Act Section 10 permit.

ORCID

Andrew R. Murdoch https://orcid.org/0000-0003-2482-7689

REFERENCES

Beamesderfer, R. C. P., Schaller, H. A., Zimmerman, M. P., Petrosky, C. E., Langness, O. P., & LaVoy, L. (1997). Spawner-recruit data for spring and summer Chinook Salmon populations in Idaho, Oregon and Washington. In D. R. Marmorek & E. Peters (Eds.), Plan for analyzing and testing hypotheses (PATH): Report of retrospective analysis for fiscal year 1997. Vancouver, Canada: ESSA Technologies Ltd.
Burnham, K. P., & Anderson, D. R. (2002). Model selection and multi-model inference: A practical information-theoretic approach. New York, NY: Springer-Verlag.

Chapman, D., Pevcn, C., Giorgi, A., Hillman T., & Utter F. (1995). Status of spring Chinook in the Mid-Columbia Region, Report for Chelan County Public Utility District. Don Chapman Consultants, Boise, ID.

Chasco, B. E., Ward, E. J., Hesse, J. A., Rabe, C., Kinzer, R., Vogel, J. L., & Orme, R. (2014). Evaluating the accuracy and precision of multiple abundance estimators using stat-space models: A case study for a threatened population of Chinook Salmon in John- son Creek, Idaho. North American Journal of Fisheries Management, 34, 945–954.

Crawford, B. A., & Rumsey, S. (2009). Draft guidance for monitoring and recovery of Salmon and Steelhead, listed under the Federal Endangered Species Act (Idaho, Oregon, and Washington). Portland, OR: NOAA Fisheries Service-Northwest Region.

Doob, J. L. (1935). The limiting distributions of certain statistics. The Annals of Mathematical Statistics, 6, 160–169.

Dunham, J., Rieman, B., & Davis, K. (2001). Sources and magnitude of sampling error in redd counts for bull trout. North American Journal of Fisheries Management, 21, 343–352.

Gallagher, S. P., Hahn, P. K. J., & Johnson, D. H. (2007). Redd counts. In D. H. Johnson (Ed.), Salmonid field protocols handbook: Techniques for assessing status and trends in salmon and trout populations (pp. 197–234). Bethesda, MD: American Fisheries Society.

Hughes, M. S., & Murdoch, A. R. (2017). Spawning habitat of hatchery spring Chinook Salmon and possible mechanisms contributing to lower reproductive success. Transactions of the American Fisheries Society, 146, 1016–1027.

ISEMP/CHaMP. (2017). Integrated Status and Effectiveness Monitoring Program (ISEMP) and Columbia Habitat Monitoring Program (CHaMP) Annual Combined Technical Report, January – December 2016, BPA Projects 2003–017-00 and 2011–006-00.

Liermann, M. C., Rawding, D., Pess, G. R., & Glaser, B. (2015). The spatial distribution of salmon and steelhead redds and optimal sampling design. Canadian Journal of Fisheries and Aquatic Sciences, 72, 434–446.

Madej, M. A. (1999). Temporal and spatial variability in thalweg profiles of a gravel-bed river. Earth Surface Processes and Landforms, 24, 1153–1169.

McElhany, P., Ruckelshaus, M. H., Ford, M. J., Wainwright, T., & Bjorkstedt, E. (2000). Viable salmonid populations and the recovery of evolutionarily significant units. In NOAA technical memorandum NMFS-NWFSC-42. Seattle, WA: U.S. Department of Commerce.

Muhlfeld, C. C., Taper, M. L., Staples, D. F., & Shepard, B. B. (2006). Observer error structure in bull trout redd counts in Montana streams: Implications for inference on true redd numbers. Transactions of the American Fisheries Society, 135, 643–654.

Mullan, J. W., Williams, K. R., Rhodus, G., Hillman, T. W., & McIntyre, J. D. (1992). Production and habitat of salmonids in mid-Columbia River tributary streams. United States Fish and Wildlife Service Monograph 1. Leavenworth, WA: U.S. Fish and Wildlife Service.

Murdoch, A. R., Herring, C., Frady, C. H., See, K., & Jordan, C. E. (2018). Estimating observer error and steelhead redd abundance using a modified Gaussian area-under-the-curve framework. Canadian Journal of Fisheries and Aquatic Sciences, 75, 2149–2158.

Murdoch, A. R., Pearsons, T. N., & Maitland, T. W. (2009). The number of redds constructed per female spring Chinook Salmon in the Wenatchee river basin. North American Journal of Fisheries Management, 29, 441–446.

Murdoch, A. R., Pearsons, T. N., & Maitland, T. W. (2010). Estimating the spawning escapement of hatchery and natural origin spring Chinook salmon using redd and carcass data. North American Journal of Fisheries Management, 30, 361–375.

NWFSC (Northwest Fisheries Science Center). (2015). Status review update for Pacific salmon and steelhead listed under the Endangered Species Act: Pacific Northwest. Seattle, WA: National Marine Fisheries Service, Northwest Fisheries Science Center.

R Core Team. (2015). R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing Retrieved from http://www.R-project.org

Seber, G. A. F. (1973). The estimation of animal abundance and related parameters. London, England: Griffin.

Thruow, R., & McGrath, C. (2010). Evaluating the bias and precision of Chinook salmon redd counts in the middle fork Salmon River basin, Idaho. Annual Report Project Number: 2002–049-00. Fort Collins, CO: U.S. Department of Agriculture, Rocky Mountain Research Station.

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section at the end of this article.

How to cite this article: Murdoch AR, Frady CH, Hughes MS, See K. Estimating population size and observation bias for spring Chinook Salmon. Conservation Science and Practice. 2019;1:e120. https://doi.org/10.1111/csp2.120