



Estimating abundance of endangered fish by eliminating bias from non-constant detectability

David R. Stewart^{1,*}, Matthew J. Butler¹, Grant Harris¹, Lacreacia A. Johnson², William R. Radke³

¹US Fish and Wildlife Service, Division of Biological Sciences, Albuquerque, NM 87103, USA

²US Fish and Wildlife Service, Sonoran and Chihuahuan Deserts, 12661 East Broadway, Tuscon, AZ 85748, USA

³US Fish and Wildlife Service, San Bernardino National Wildlife Refuge, PO Box 3509, Douglas, AZ 85608, USA

ABSTRACT: Worldwide, approximately half of all freshwater fish are threatened with extinction or lack sufficient data to classify their conservation status. We focused on 3 such species endemic to southeastern Arizona, USA, and Sonora, Mexico: Yaqui topminnow *Poeciliopsis occidentalis sonoriensis*, Yaqui chub *Gila purpurae*, and beautiful shiner *Cyprinella formosa*. These species, like many others, require accurate estimates and trends of abundance to characterize their conservation status. Hence, sampling must be designed appropriately. We used historical data to determine precision and minimum number of traps necessary to estimate abundance with relative precision ≤ 25 . Next, we examined alternative trap soak times to improve sampling efficiency. We then incorporated variables influencing detectability to produce unbiased abundance estimates. Finally, we simulated how ignoring biases from heterogeneous detectability affects abundance indices and the ability to detect changes in abundance. We found catch estimated from the historical design, which assumed constant detectability, had much variability, thereby requiring ~ 40 traps pond^{-1} . A 4 h soak duration increased catch for 2 species and detection probability for all. Detectability increased while abundance decreased with water temperature (all 3 species). Detection of beautiful shiner declined as abundance increased with depth. Our simulations indicated that if detectability varies but is assumed constant (i.e. ignoring detection probability), the probability of finding population change when none occurred (i.e. Type I error) is 50%, and the probability of detecting true population change (i.e. power) declines to ~ 0.75 . Ignoring variable detectability in count data can misrepresent species status, habitat relationships, and ultimately mislead conservation and stewardship of endangered species.

KEY WORDS: Endangered species · Río Yaqui fishes · Bayesian mixture models · Monitoring

INTRODUCTION

The abundance of rare and endangered species is notoriously difficult to estimate (Magnuson et al. 1994, Jerde et al. 2011), yet such information remains vital for evaluating species conservation status, assaying their trends through time, and evaluating the efficacy of protection and management efforts toward their recovery. Given the challenges and impetus for data describing population sizes, a common temp-

tation is to acquire sample data without accounting for critically important analytical considerations (i.e. detection probability), which may require additional sampling time and effort to address (Stewart & Long 2016). Doing so provides data quicker, cheaper, and is logistically easier. The risk lies in providing data that can be wrong, and in the worst case, inadvertently applied to misinform conservation plans and strategies. Clearly, such an outcome counters the objectives that conservation programs are engineered to address.

*Corresponding author: david_stewart@fws.gov

© Outside the USA, the US Government 2017. Open Access under Creative Commons by Attribution Licence. Use, distribution and reproduction are unrestricted. Authors and original publication must be credited.

In southern Arizona (USA), 7 fish species are endemic to the Río Yaqui watershed and remain at and around San Bernardino and Leslie Canyon Nation Wildlife Refuges. Of these, the Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* and the Yaqui chub *Gila purpurae* are federally listed as endangered, while the beautiful shiner *Cyprinella formosa* and the Yaqui catfish *Ictalurus pricei* are federally listed as threatened (USFWS 1994). Our focus was on the first 3 species. These fish are relegated to spring-fed ponds and streams. From 2004 to 2014, Federal, State and multiple conservation organizations annually monitored populations of these fish to insure their persistence and evaluate their recovery. For instance, Yaqui chub and Yaqui topminnow will be downlisted to threatened, and beautiful shiner delisted, when self-sustaining populations become established and secure on San Bernardino and Leslie Canyon Nation Wildlife Refuge lands for at least 10 yr (USFWS 1994). However, historical monitoring efforts were based on a standardized sampling design that does not provide a bias-corrected estimate of abundance (i.e. it assumes constant detection probability). This situation is problematic, as other studies have demonstrated that using simple catch indices may inappropriately index the status or abundance of any species, much less inform recovery criteria of endangered species (Hyde & Simons 2001, Pollock et al. 2002).

Although standardization of sampling designs is encouraged, the efficacy of this approach is still contingent on counts (i.e. samples) having a constant and proportional relationship with abundance such that a direct linear relationship exists between the index and population size (Anderson 2001, Pollock et al. 2002). A major problem with studies that rely on techniques trying to control for factors that affect detectability (through standardization) is that few variables affecting detection probability can be controlled, and such variables usually do not influence the abundance and detection processes. Thus, fluctuations observed over the course of monitoring with such techniques may be nothing more than changes in factors affecting measurement error (detection probability; Stewart & Long 2016). Hence, observed trends or differences between treatments may stem from changes in detectability, not changes in true abundance. Therefore, it is important to attain abundance estimates corrected for detection probability by developing sampling designs that can effectively monitor changes in abundance and detection probability over time. When variation in

detection probability is not accounted for as part of the estimation process, the power to detect change is ultimately masked (Hyde & Simons 2001, Pollock et al. 2002).

The use of historical survey data is a valuable source of information for developing future sampling designs. For example, the use of information from historical surveys can be used to gauge the efficiency and shortcomings of the previous design, measure precision, and quantify the degree of inter-annual variability along with measurement issues related to variance components (Bonar & Hubert 2002). Aspects of the old versus a new sampling protocol should be evaluated using an experimental framework to determine the effects of measurement error on the utility of the index to generalize to the population's abundance (Kéry & Royle 2010). If the sampling objectives are to infer changes in abundance over time, then the design of a statistically valid survey that advances beyond the shortcomings of the existing survey is essential to ensure that inferences can be characterized to the population of interest (Maunder & Punt 2004, Wagner et al. 2009).

The goal of this study was to evaluate existing data and investigate components of an alternative sampling design that could be used to improve monitoring and management of populations of rare, endangered fish species. Our objectives were to (1) determine minimum effective number of trap deployments, (2) optimize soak times for traps, (3) develop and evaluate models that incorporate detection processes into abundance estimates and identify which site-specific correlates effect detection and abundance, and (4) evaluate the effects of variable detection probability on using raw counts to monitor temporal trends.

MATERIALS AND METHODS

Study area

We conducted our study at San Bernardino National Wildlife Refuge (NWR; 9.34 km²) and nearby private properties located in southeastern Arizona (31.34° N, 109.25° W), located on the US–Mexico border east of Douglas, Arizona. San Bernardino NWR was established in 1982 to provide conservation and management to Río Yaqui fishes endemic to southeastern Arizona and Sonora, Mexico. The Río Yaqui fishes historically occupied streams, ciénegas (spring-fed marshes), pools, and spring-fed ponds supported by flowing artesian wells (Minckley 1985).

Evaluation of the historical survey design

Fish were sampled in 12 spring-fed ponds using unbaited Gee funnel minnow traps (3.18 mm mesh, 25 mm openings) during October and November 2005 to 2014. Ponds were 28 to 8813 m² in size and their depth ranged from 0.4 to ~4 m. Most ponds were relatively shallow, clear, and dominated by submergent aquatic vegetation such as spiny naiad *Najas minor*, sago pondweed *Potamogeton pectinatus*, coontail *Ceratophyllum demersum*, and horned pondweed *Zannichellia palustris*. A fixed number of traps were used to sample each pond annually, ranging from 8 to 30 (depending on pond size). Traps were set following a proportional stratified random sampling design based on vegetation and depth surveys conducted 1 wk prior to sampling. For example, if 50% of the pond was dominated by cattail, then 50% of the traps were set randomly in cattail and 50% randomly out of cattail. Target species for this project included Yaqui chub *Gila purpurae* and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis*. The beautiful shiner *Cyprinella formosa* was excluded from this analysis because of insufficient temporal data. For both Yaqui chub and Yaqui topminnow, data from a pond were included when at least 5 individuals of that species were sampled from the respective pond during the study years. The final number of ponds included in the analysis was 10 for both species.

Evaluation of an alternative survey design

We selected 12 ponds for sampling using 12 unbaited Gee funnel minnow traps (3.18 mm mesh, 25 mm openings) per pond during November 2015. Each pond was sampled for approximately 7 h, split into 2 sample periods. Each sample period was assigned a soak duration (2 or 4 h) with an approximate 1 h rest period between them. The 2 h soak duration was based on the methodology from the historical survey design, and we doubled the amount of soak time to determine if this additional time improved survey results. We used a crossover sampling design where the soak duration sequence (2 then 4 h, or 4 then 2 h) was assigned randomly and to an equal number of sites within each pond. A crossover design requires fewer samples, avoids confounding spatial and temporal effects with soak duration, and allows carryover effects (i.e. the sequence of treatments influencing results) to be removed as part of the estimation process (Kuehl 2000). Each pond was strati-

fied according to water depth (deep water: >60 cm; shallow water: ≤60 cm). Trap density was randomly and equally divided between the 2 depth strata so that an equal number of traps were set randomly in both shallow and deep depths within each pond. Both soak treatments were administered equally in each period and depth stratum such that the treatment comparisons were independent of comparisons between periods. Within each stratum, 3 traps were assigned the treatment sequence '2 h then 4 h' and the other 3 traps were assigned the '4 h then 2 h' treatment sequence. All fish were removed from each trap and processed after the first soak duration. Traps were then re-set at the same location and allowed to soak using the opposite soak duration (after a 1 h rest period). At each site, we measured trap depth (m), Secchi depth (cm), water temperature, and visually estimated vegetation cover around each trap. We also recorded whether traps were set closer to the inlet or outlet of the pond to assess spatial structuring (Stewart et al. 2015).

Data analysis

Objective 1: Minimum effective number of traps

We calculated the total number of fish trapped and the relative standard error (RSE) from the historical survey data set for each target species (Yaqui chub and topminnow) by pond and sample year, to determine the degree of precision in catch indices (Stewart & Long 2012). For each pond, we determined precision and the minimum effective number of traps needed to attain $RSE \leq 25$ (RSE25) using data from each pond and target species as part of a stochastic resampling procedure (Stewart & Long 2012). We chose a target level of RSE25 based on recommendations by Robson & Regier (1964) and Hardin & Conner (1992). The resampling technique randomly sampled 2 traps from the empirical data for each pond and species 3000 times with replacement. Next, the proportion of the 3000 samples achieving an $RSE \leq 25$ was calculated (Stewart & Long 2012). If the proportion was <0.80, then the number of traps was increased from 2 to 3 and continued to increase by an additional trap until the proportion of the 3000 samples from the resampling routine achieving the desired precision (RSE25) was ≥0.80 (Dumont & Schlechte 2004, Stewart & Long 2012).

To guide future sampling efforts, we evaluated the relationship between the projected number of traps needed to attain a precision of RSE25 and catch sta-

tistics (mean number of fish captured and coefficient of variation [CV]) calculated from the historical survey data using log-linear mixed effect models. We assumed that the projected number of trap deployments followed a Poisson distribution, and evaluated separate variance components of random pond and year effects that were incorporated into the model as an exchangeable term: $N \sim (0, \sigma)$. Only 1 random effect (pond or year) was used in the final model and was selected based on Akaike's information criterion (AIC). Catch and CV were not included together in the same model due to collinearity, and we explored relationships with each variable. Analyses were conducted using the 'lme4' package in R (R Core Team 2015). The package 'lmerTest' was used to calculate F - and p -values using the Kenward-Rogers approximation to determine denominator degrees of freedom (Kuznetsova et al. 2013).

Objective 2: Soak time

We used a crossover sampling design to determine if soak duration (2 vs. 4 h) influenced the number of fish caught (i.e. catch) of the target species (Yaqui chub, Yaqui topminnow and beautiful shiner). Since catch can vary by trap location, pond, soak sequence, and sampling time, we incorporated these terms using indicator variables. First, we used $\log_{10}(x + 1)$ to normalize the catch rate data (y). We then developed a crossover statistical design to account for treatment sequence (2 to 4 h, 4 to 2 h) based on the interaction effect between the sampling event (i.e. period) and the treatment ($\lambda_{c[k-1]}$). This approach represents a carryover effect using the linear mixed model:

$$Y_{ij(h)k} = \mu + \alpha_i + \beta_j + \rho_{ij(h)} + \gamma_k + \tau_{d(ik)} + \lambda_{c(k-1)} + \varepsilon_{ij(h)k} \quad (1)$$

where α_i is the i^{th} treatment sequence effect, β_j is the j^{th} species effect, $\rho_{ij(h)}$ is the random effect for the j^{th} site (i.e. trap location) within the h^{th} pond of the i^{th} treatment sequence, γ_k is the k^{th} period effect, $\tau_{d(ik)}$ is the direct (d) effect of the treatment administered in period k of sequence group i (since soak treatments varied randomly by site and period) and $\varepsilon_{ij(h)k}$ is the random experimental error for the j^{th} site within the h^{th} pond of the k^{th} period (Kuehl 2000). Therefore, this model accounts for the variability in catch due to the location of the traps within a pond, the variability between ponds, the order of soak times, and when sampling took place (period). Analysis was conducted using the package 'lme4', which was followed by a Tukey's honestly significant difference (HSD) test in instances of significance ($p \leq 0.05$). The package

'lmerTest' was used to determine F - and p -values while correcting the denominator degrees of freedom using the Kenward-Rogers approximation (Kuznetsova et al. 2013).

Objective 3: Site-specific correlates of detection and abundance

To quantify detection and true local abundance in a pond, we developed a mixture model in a Bayesian framework to link the subset of individuals collected from temporally replicated samples to a latent abundance state of animals present at a site within each population (i.e. pond). Our model differs from the N -mixture model presented by Royle (2004) in that our model estimates local abundance for multiple populations, inherently accounts for spatial autocorrelation and allows for overdispersion (i.e. mean \neq variance). The model presented herein builds on a multi-species mixture model presented by Stewart & Long (2016), which was developed to account for the added complexity of modeling multiple species instead of the single-species Binomial mixture framework. We consider an experimental design where the observed elements of the model represent a sequence of replicated counts of unmarked individuals, y , at each site $i = 1, 2, \dots, R$ within $k = 1, \dots, K$ populations (i.e. ponds) for a total of $t = 1, \dots, T$ primary sampling occasions. Therefore, the observed data, y_{itk} , can be denoted by $Y = \{y_{itk}; i = 1, 2, \dots, R; t = 1, \dots, T; k = 1, \dots, K\}$ and is regarded as a binomial outcome $h(y_{itk} | N_{ik}, p_{itk})$. This outcome is conditional on the unknown total number of individuals available for sampling, N_{ik} , at location i within the k^{th} population (i.e. pond), and p_{itk} , which is the probability of detecting an individual at the i^{th} site at the t^{th} sampling occasion within the k^{th} population (i.e. pond):

$$y_{itk} | N_{ik}, p_{itk} \sim \text{Bin}(N_{ik}, p_{itk}) \quad (2)$$

Since the abundance, N_{ik} , at a site varies, we specified our model to address overdispersion by specifying the process model for N_{ik} to be marginal to a hierarchical element, ε_i . This procedure accounts for variation at a site within a population by integrating this dispersion parameter into the likelihood of a Poisson process as a random effect to account for extra-Poisson variation (i.e. overdispersion; Stewart & Long 2016). This framework results in a marginally distributed multi-population negative binomial mixture model by considering $f(N_{ik} | \gamma_{ik}, \varepsilon_i)$, $\varepsilon_i \sim \text{gamma}(\theta, \theta)$ and $\theta > 0$, which results in a probability distribution (P) for N_{ik} as:

$$P(N_{ik} = n_{ik} | E_{ik}, \lambda_{ik}, \theta) = \frac{\Gamma(n_{ik} + \theta)}{\Gamma(n_{ik} + 1)\Gamma(\theta)} \left(\frac{\lambda_{ik}}{\lambda_{ik} + \theta} \right)^{n_{ik}} \left(\frac{\theta}{\lambda_{ik} + \theta} \right)^\theta \quad (3)$$

where θ is the parameter of extra-Poisson variation in abundance among sites within each population, with large values of θ consistent to variability similar to the Poisson distribution. As $\theta \rightarrow \infty$, the distribution of N_{ik} converges to a Poisson random variable, where the level of overdispersion (θ) is assumed to be the same in all K ponds.

We specified the parameter model to relate covariates to the detection probability through a logistic link function, as:

$$\text{logit}(p_{itk}) = \gamma_{0k} + \sum_{v=1}^w \gamma_{v,k} x_{v,itk} \quad (4)$$

and mean abundance through a log link function, as:

$$\text{log}(\lambda_{ik}) = \alpha_{0k} + \sum_{v=1}^w \alpha_{v,k} x_{v,i} + \delta_k \quad (5)$$

where x_v are predictors $v = 1, 2, \dots, w$ such as water temperature, Secchi depth, water depth, percent submergent aquatic vegetation, location near the pond inlet, and soak duration (2 vs. 4 h) measured at a site, i (and also time, t) in population k (i.e. pond). The γ 's and α 's are the fixed effect coefficients, and δ is a random effect or exchangeable error term that is $N \sim (0, \sigma^2)$ and specifies the variation in mean abundance among ponds. Given data compiled in a pond-by-site-by-survey count matrix, we let $Q = \{Q_\gamma, Q_\alpha, Q_\varepsilon\}$ denote the model structures for the vector of covariates for p_{itk} and λ_{ik} and the hierarchical dispersion parameter $\varepsilon = \{\varepsilon_1, \dots, \varepsilon_R\}$, respectively. Therefore, we denoted the likelihood function \mathcal{L} to integrate out N_{ik} to form the marginal distribution of observing y as

$$\mathcal{L}(\mathbf{Y} | Q, \gamma, \alpha, \varepsilon) = \prod_{i=1}^R \prod_{t=1}^T \prod_{k=1}^K [y_{itk} | N_{ik}, Q_\gamma, \gamma] [N_{ik} | Q_\alpha, \alpha, Q_\varepsilon, \varepsilon_i] \quad (6)$$

where $[y_{itk} | N_{ik}, \gamma, Q_\gamma]$ denotes the conditional distribution of y_{itk} given the parameters, and let $y = (y_{it1}, \dots, y_{itk})$ denote the matrix of observed data.

We fit hierarchical models using covariates standardized with mean zero and standard deviation of one, and only predictor variables that had a correlation coefficient (r) < 0.60 were used in the same model to reduce intercorrelation due to potential multicollinearity (Dormann et al. 2013). All predictor variables could be included in the same models except Secchi depth and water depth, as they were correlated. Models were implemented using jagsUI 1.3.1 to access JAGS 3.4.0 using Markov chain Monte Carlo (MCMC) algorithms to generate posterior dis-

tributions of the parameters (Plummer 2003, Kellner 2015). Uninformative, independent priors were chosen for the model parameters. Specifically, we considered $N(\mu = 0, \sigma = 0.01)$ for all regression parameters (the intercept and slope coefficients) and a $\text{Unif}(0, \dots, 10)$ prior for θ . Models were fit using 3 parallel chains simulated for 150 000 iterations with a burn-in of the first 50 000 iterations. To assess convergence, we used the Gelman-Rubin diagnostic convergence statistics, examination of chain histories and the posterior density plots (Gelman & Rubin 1992).

Objective 4: Variable detection probability mask trends in raw counts

We used simulation to estimate how ignoring biases from heterogeneous detection affects the utility of abundance indices (the actual result and inferences gained from it). Specifically, we examined (1) detection of change in the number of fish captured when a change 'has not' occurred (Type I error rate) and (2) detection of change in the number of fish captured when a change 'has' occurred (statistical power).

We simulated the number of fish captured (abundance index) for 2 yr and simulated a population change (100% increase) during the second year. In pre-population change years, we simulated the number of fish captured from 6 to 120 sampling sites from 3000 multi-year datasets (i.e. replications) using random draws from a negative binomial distribution:

$$\text{catch}_i \sim \text{NBin}(\lambda p_y, k) \quad (7)$$

where λp_y is the expected catch at site $i = 6, \dots, 120$, p_y is the probability of detection in year y , k is the overdispersion parameter that allows for extra-Poisson variation, and i is the site-specific number of fish available for capture (Royle 2004). We simulated extra-Poisson variation by using a negative binomial distribution since many ecological scenarios result in overdispersed count data (Lindén & Mäntyniemi 2011). We set λ and variation in p_y to reflect our species-specific data. We based λ on the average estimated species-specific abundance estimates, and variation in p_y was simulated as a random draw from a beta distribution and parameterized (a, b) using estimates of species-specific detection probability determined from the Bayesian mixture models. In post-population change years, catch was simulated where site-specific abundance λ was doubled, $\text{catch}_i \sim \text{NBin}(2\lambda p_y, k)$ (Hangsleben et al. 2013).

We compared species-specific estimates among annual (1 yr) sampling schedules pre- and post-

population change using the historical sampling design, which did not account for detection probability and assumes no variation in p_y . We calculated statistical power and Type I error rate using a 2-tailed t -test with equal variance, by comparing simulated catch from pre- and post-change years (Hangsleben et al. 2013). Statistical power was calculated as the probability that a 100% increase in population size in post-change years could be detected. Type I error rate was calculated as the probability that aberrations in population change could be detected in post-change years when a 0% increase was simulated.

RESULTS

Minimum effective number of traps

Observed catch of Yaqui chub *Gila purpurae* and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* were highly variable among ponds and years (Fig. 1). Median catch among sampling sites for Yaqui chub was 162 (range: 3 to 832) and 178 (range: 12 to 1132) for Yaqui topminnow among ponds and years. Similarly, CV ranged from 0.56 to 4.75 and RSE ranged from 0.14 to 1.15 for both species. Annual sampling rarely achieved the median RSE25. Our resampling analysis indicated that ~40 traps pond⁻¹ would be needed to acquire a RSE25 and effectively sample both species (Fig. 1). Less than 25 traps pond⁻¹ were needed to reach an RSE25 at least 40% of the time. Importantly, the amount of samples required changes over time given variability in catch.

The best approximating variance structure in the log-linear mixed effect models included the random year effect (Table 1). Therefore, the mixed effects models reported herein included a random year effect term. The model incorporating CV outperformed the model with catch (Table 1, Fig. 2). The parameter estimates suggested that an increase in CV would result in an increase in the number of traps needed to effectively sample these species (Fig. 2). For example, for each increase of 1.00 in the CV of fish captured, 83 and 76% more traps

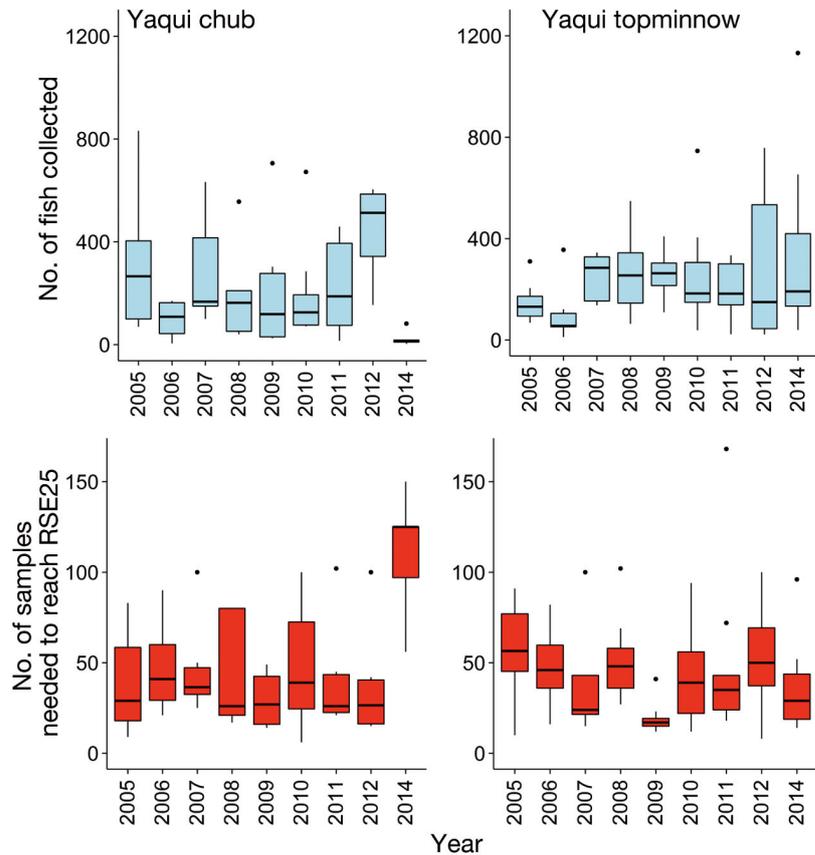


Fig. 1. Box plots of numbers of fish captured and sampling effort needed to achieve a relative standard error of 25% (RSE25) using historical data from 2005 to 2014 for Yaqui chub *Gila purpurae* and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* from 10 spring-fed ponds. Boxes are 25th and 75th percentile; black horizontal line: median; black dot: outlier; whiskers: $Q \pm 1.5 \times IQR$ (interquartile range)

would be required to reach an RSE25 for Yaqui chub and Yaqui topminnow (Fig. 2). Relationships between the estimated number of traps and CV of the sample were significant for both species ($p = 0.01$), with high statistical power (>0.96) and good

Table 1. Log-linear mixed-effect regression models relating the required number of traps to Yaqui chub *Gila purpurae* and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* catch statistics. AIC: Akaike's information criterion; CV: coefficient of variation

Species/Model	AIC	Pseudo Power R ²
Yaqui chub		
$\log(\text{traps}) = 2.71 + 0.63 \times \text{CV}$	593.86	0.76
$\log(\text{traps}) = 4.34 - 0.003 \times \text{catch}$	989.16	0.43
Yaqui topminnow		
$\log(\text{traps}) = 2.81 + 0.58 \times \text{CV}$	823.97	0.67
$\log(\text{traps}) = 4.08 - 0.001 \times \text{catch}$	1524.5	0.13

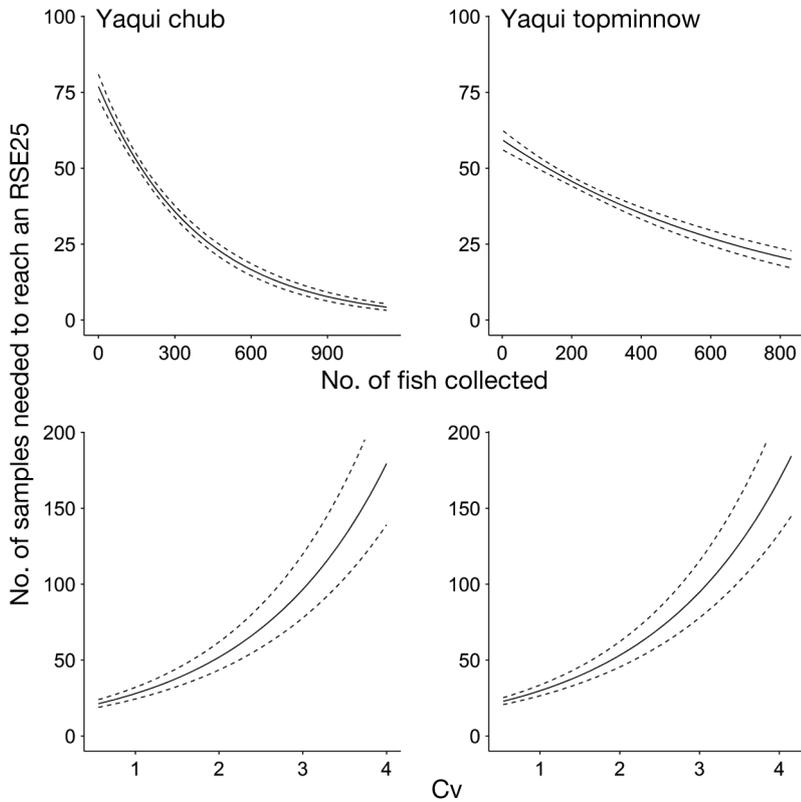


Fig. 2. Number of Gee funnel minnow traps needed to achieve a relative standard error of 25% (RSE25) for Yaqui chub *Gila purpurea* and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* given numbers of fish captured and coefficient of variation. Dashed lines are 95% confidence intervals

model fit (pseudo $R^2 > 0.67$). This model predicted that 39 traps pond^{-1} on average were required to achieve RSE25 for both species, given a mean CV of 1.50.

Parameter estimates for the model incorporating catch indicated that the estimated number of traps was positively related to catch of both species (Table 1, Fig. 2). For each increase of 100 in the number of fish captured, 25 and 12% fewer traps would be required to reach an RSE25 for Yaqui chub and Yaqui topminnow, respectively. Similar to the resampling analysis, this model predicted that 45 traps pond^{-1} on average were required to achieve RSE25 for both species, given a mean catch of 200 fish.

Soak time

A total of 9723 fish were collected from 24 trap deployments from the 2 trapping periods in 12 ponds. The fish catch was dominated by Yaqui topminnow ($n = 9021$, 93% of total catch). Yaqui chub ($n = 460$, 5% of total catch) and beautiful shiner (*Cyprinella formosa*; $n = 242$, 2% of total catch) were not as abundant (Table 2). Overall, traps set for 4 h caught more fish (mean = 23.55, SD = 64.59, $n = 5652$) compared to traps set for 2 h (mean = 16.69, SD = 58.73, $n = 4,071$; $F_{2,317.7} = 7.25$, $p = 0.01$) based on a species \times duration interaction (Table 2). Tests for carryover ($F_{1,317.7} = 0.23$, $p > 0.05$), period ($F_{1,317.7} = 1.24$, $p > 0.05$), and sequence

for 2 h (mean = 16.69, SD = 58.73, $n = 4,071$; $F_{2,317.7} = 7.25$, $p = 0.01$) based on a species \times duration interaction (Table 2). Tests for carryover ($F_{1,317.7} = 0.23$, $p > 0.05$), period ($F_{1,317.7} = 1.24$, $p > 0.05$), and sequence

Table 2. Number of beautiful shiner *Cyprinella formosa*, Yaqui chub *Gila purpurea*, and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* captured with Gee funnel minnow traps set for 2 or 4 h in 12 spring-fed ponds. -: species did not reside in pond

Pond	Beautiful shiner			Yaqui chub			Yaqui topminnow		
	2 h	4 h	Total	2 h	4 h	Total	2 h	4 h	Total
Ben New Pond	1	0	1	0	0	0	234	254	487
Brasher Pond	-	-	-	-	-	-	635	880	1515
Cobble Pond	12	95	107	-	-	-	337	161	498
Hay North Pond	-	-	-	6	1	7	1349	1967	3316
Hay South Pond	-	-	-	-	-	-	99	184	283
Horse Tank	-	-	-	119	251	370	-	-	-
Lower Twin Pond	6	1	7	20	27	47	360	188	548
Magoffin Pond	1	6	7	0	1	1	11	529	540
Minckley North	-	-	-	0	1	1	417	322	739
Minckley South	-	-	-	0	0	0	249	314	563
Tennis Court	-	-	-	11	24	35	-	-	-
Twin Pond	11	109	120	0	0	0	192	340	532
Totals	31	211	242	157	304	461	3883	5138	9021

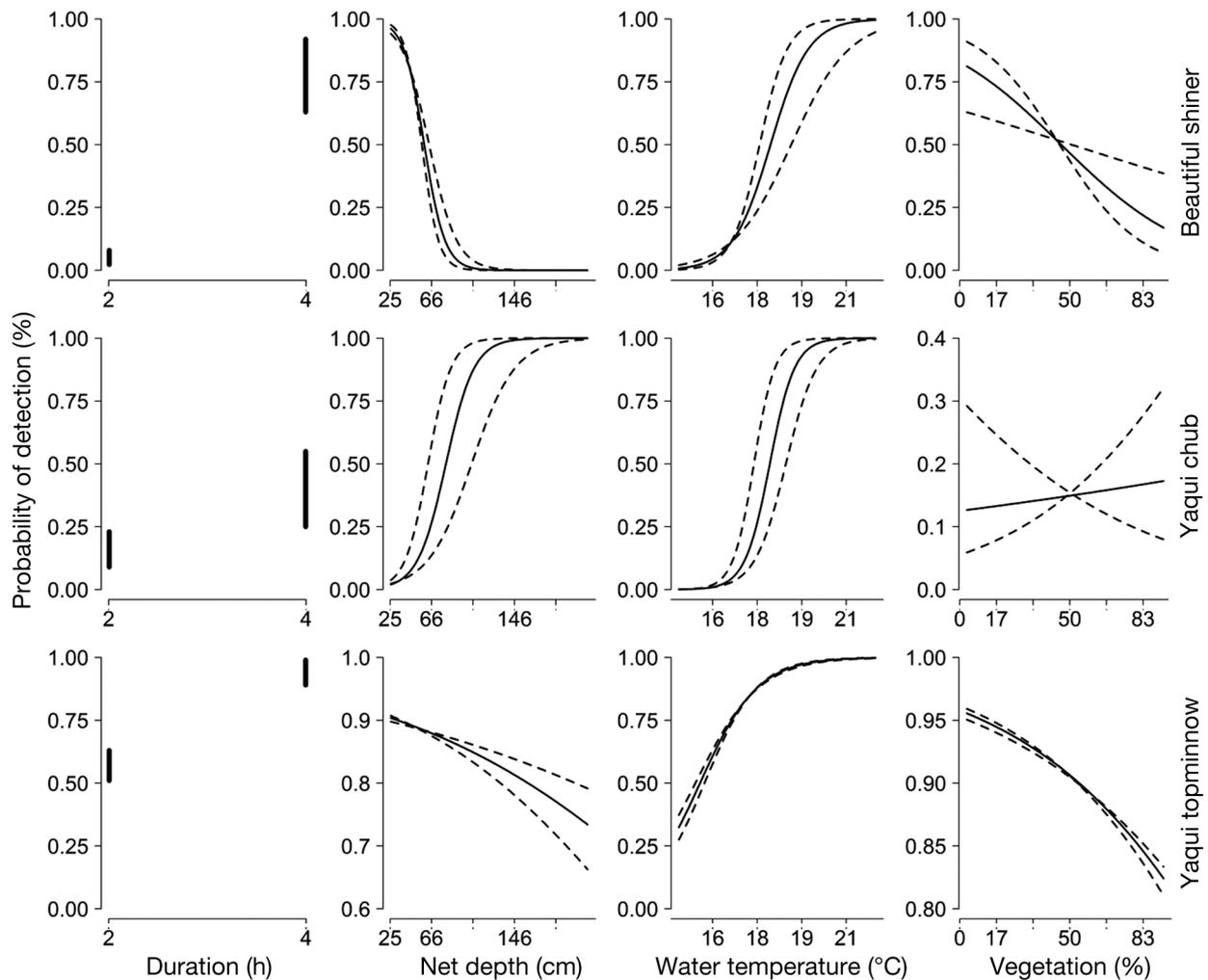


Fig. 3. Estimated mean conditional detection probabilities ($p|N$) in relation to soak duration (2 vs. 4 h), trap depth (cm), water temperature ($^{\circ}\text{C}$), and submergent aquatic vegetation (%) for beautiful shiner *Cyprinella formosa*, Yaqui chub *Gila purpurea*, and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis*. Middle lines: posterior means; dashed lines: 95% credible intervals

effects ($F_{1,122} = 1.25$, $p > 0.05$) were not significant. Traps set for 4 h caught more beautiful shiner (Tukey's HSD = 2.70, $p < 0.01$) and Yaqui topminnow (Tukey's HSD = 2.63, $p < 0.01$) compared to traps set for 2 h. Catch of Yaqui chub did not significantly differ between the 2 and 4 h trap sets (Tukey's HSD = 1.86, $p > 0.05$).

Site-specific correlates of detection and abundance

Estimates of detection probability and total 'local' abundance (sum of the site-specific bias-corrected estimates of abundance at a site) ranged from 0.06 to 0.25 and 4 to 1019 for beautiful shiner, 0.14 to 0.83 and 0 to 1182 for Yaqui chub, and 0.05 to 0.34 and 832 to 16447 for Yaqui topminnow. The single-

species hierarchical Bayesian mixture models relating site-specific environmental correlates to detection and abundance indicated that Río Yaqui fish distributions were affected by habitat factors. The best overall model included duration (2 vs. 4 h), net depth, water temperature, percent submergent aquatic vegetation, and location relative to the inlet or outlet of a pond (Figs. 3 & 4).

For all 3 species, the 4 h soak duration increased detection probability. Detection rose by ~75% for the beautiful shiner, and approximately doubled for the Yaqui chub and Yaqui topminnow (Fig. 3). These species also displayed greater abundance at the inlet (Fig. 4): beautiful shiner were approximately 10 times more abundant at inlets, Yaqui topminnow were 200 to 600 times more abundant, and Yaqui chub were ~1.5 times more abundant (Fig. 4). Fish

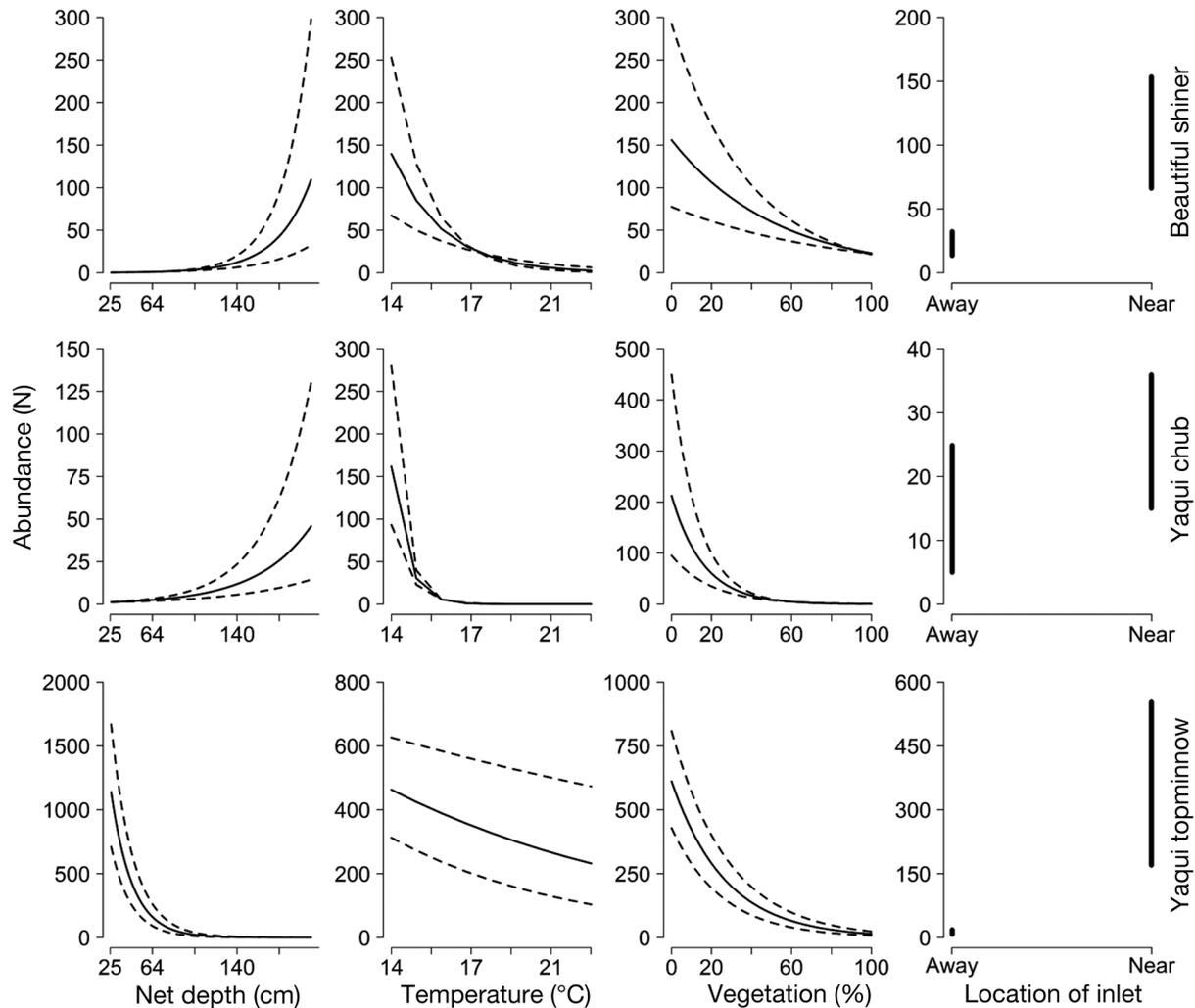


Fig. 4. Estimated mean site-specific abundance estimates, $f(N|\lambda, \epsilon)$, in relation to trap depth (cm), water temperature ($^{\circ}\text{C}$), submergent aquatic vegetation (%), and location of nets relative to the inlet of the sampled spring-fed ponds for beautiful shiner *Cyprinella formosa*, Yaqui chub *Gila purpurea*, and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis*. Middle lines: posterior means; dashed lines: 95% credible intervals

detectability increased while abundance decreased with increasing water temperature (Figs. 3 & 4). Detectability increased from 0 to 25% at 16 $^{\circ}\text{C}$ to 100% by 19 $^{\circ}\text{C}$. Meanwhile, abundance of all species fell by 70% over this same temperature span. Vegetation increases were associated with declines in detection and abundance for 2 fish (Figs. 3 & 4). The estimated detection probabilities for beautiful shiner and Yaqui topminnow at 10% submergent aquatic vegetation were 78% (95% CI = 62, 94) and 94% (95% CI = 93, 95); while estimated detection probability at 80% submergent aquatic vegetation was 26% (95% CI = 22, 48) and 85% (95% CI = 84, 86). Vegetation did not appear to influence detectability of Yaqui chub, but as percent vegetation rose, the abundance of Yaqui chub declined. At 10% sub-

mergent aquatic vegetation, Yaqui chub abundance was 175 (95% CI = 100, 450) and at 80% vegetation, abundance was 0.

As net depth increased, detection for beautiful shiner declined but abundance increased. For example, estimated detection probabilities for beautiful shiner at 25 cm were 98% (95% CI = 97, 100), while abundance was essentially 0 (Figs. 3 & 4). Meanwhile, estimated detectability at 500 cm was 1% (95% CI = 0, 2) and mean abundance 100 (95% CI = 25, 300). For Yaqui chub, as net depth increased, detectability and abundance increased. Estimated detection probabilities for Yaqui chub at 25 cm were 2% (95% CI = 1, 3) while detection at 500 cm was 99% (95% CI = 98, 100). Mean abundance increased from 0 at 25 cm to 50 (95% CI = 15, 125) at 500 cm.

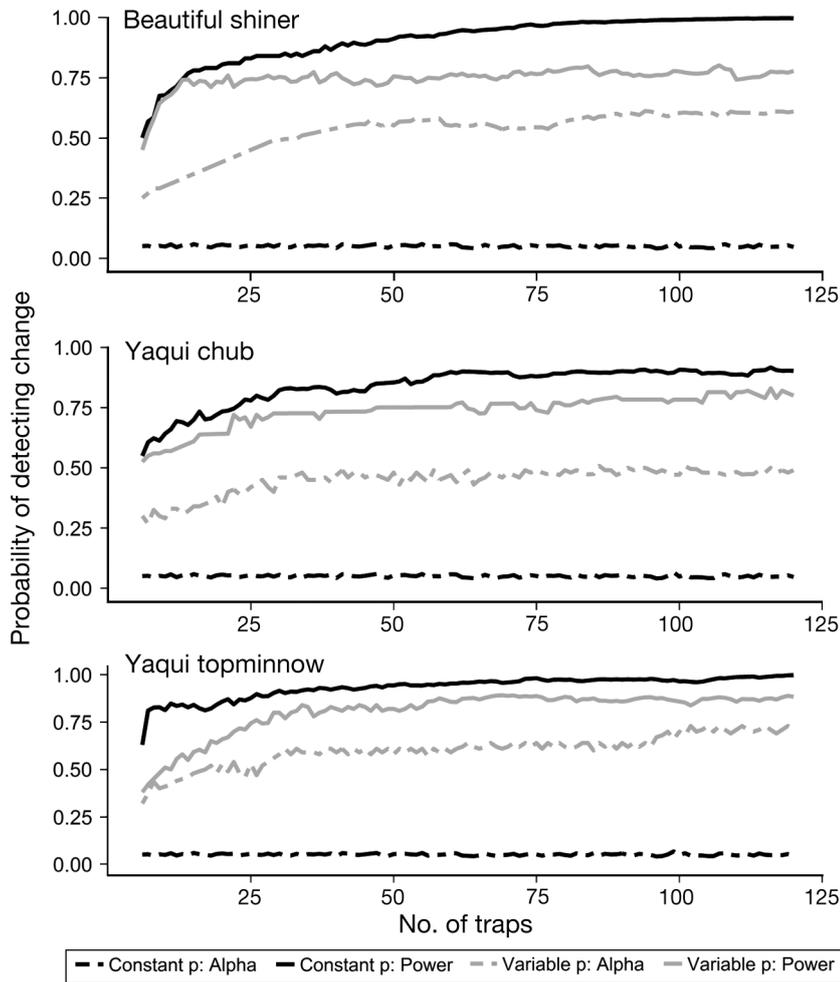


Fig. 5. Simulated statistical power and Type I error rates to determine true and false changes in catch data with constant (solid black line: statistical power; dashed black line: Type I error rate) and variable (solid gray line: statistical power; dashed gray line: Type I error rate) probability of detection estimates for annual, 2 yr, and 5 yr sampling intervals for beautiful shiner *Cyprinella formosa*, Yaqui chub *Gila purpurea*, and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis*

Detection probability and abundance declined with net depth for the Yaqui topminnow. The estimated detection probability at 25 cm was 90% (95% CI = 89, 91) with mean abundance of 1100 (95% CI = 1600, 600). At 500 cm, detection probability was 76% (95% CI = 69, 83) with mean abundance of 0.

Variable detection probability masks trends in raw counts

We simulated the effects of constant and heterogeneous detection probability to determine the utility of raw catch indices to express temporal trends in beau-

tiful shiner, Yaqui chub, and Yaqui topminnow abundance (Fig. 5). Annual variability in detection probability was simulated for each species using shape parameters (beautiful shiner: $\alpha = 2.11$ and $\beta = 15.08$; Yaqui chub: $\alpha = 1.08$ and $\beta = 1.88$; Yaqui topminnow: $\alpha = 2.06$ and $\beta = 12.41$) estimated from a beta distribution using the model-based species-specific detection estimates (Table 2). Our approach analyzed these 2 sets of data, with one dataset generated assuming constant detection and the other incorporating variable detection.

If detection probability is constant, then there exists a 5% chance of detecting and reporting a population change when no change occurred (Type 1 error, $\alpha = 0.05$; Fig. 5 dashed black line). However, when detection probability varies but is assumed constant, the chance of detecting a population change when none occurred rises to 50% ($\alpha = 0.05$, Fig. 5 dashed gray line). If detection probability is constant, the power to detect a population change when a change did occur is >0.9 . However, when detectability varies—and again is assumed constant—the power to detect a population change, when a change did occur, declines to approximately 0.75. When variable detectability is occurring, the probability of detecting a change in abundance when no change in population size was simulated (i.e. spurious change) was highest for beautiful shiner (0.72) and Yaqui topminnow (0.80). Our results also identified that the probability of detecting a spurious change increased when sampling with more traps, because as the number of samples increases, the precision increases. Without properly addressing variable detectability, estimates of abundance can become precise, but remain biased and misleading.

DISCUSSION

The International Union for the Conservation of Nature (IUCN) lists 3120 freshwater fishes, of which 1147 (37%) are threatened with extinction (IUCN

2016). Status assessments for these species relies on information describing geographical restrictions, declines in habitat or species occurrence, and observed or suspected population declines. These factors can operate singly or in combination. Another 431 (14%) freshwater fish remain data deficient, whereby information to assess these species remains inadequate. Robust quantitative data are needed to identify threats and develop appropriate protocols that can accurately index the status of these fish.

Our study provides a credible approach for estimating abundance and trend of endangered freshwater fish, to establish their status and generate a robust monitoring program. We used historical data to inform improvements in sampling efforts and developed a technique accounting for factors affecting detection probability. Doing so ensures that any changes in abundance that are measured reflect actual changes in abundance and not changes caused by extenuating factors (Pollock et al. 2002, Stewart & Long 2016). We addressed sample size and sampling efficiency considerations, exemplified how to model detectability to produce unbiased abundance estimates, and reported the consequences of ignoring biases from heterogeneous detectability.

Fish monitoring efforts are often designed such that larger bodies of water receive greater survey effort; this is a typical standard sampling recommendation (Bonar et al. 2009). Importantly, however, Stewart & Long (2012) identified that survey effort was not related to the size of lakes. Instead, survey effort should relate to site-specific heterogeneity in abundance. Our analysis using the results from the resampling routine corroborated these findings, and identified that the number of traps should not be fixed on an annual basis (Figs. 1 & 2). We also identified a relationship between the estimated number of traps and the CV of the sample (Table 1, Fig. 2). This information is useful for predicting the number of traps needed in a given year based on *a priori* sampling events (similar to Stewart & Long 2012). This prediction addresses the temporal heterogeneity in the sampling design by basing survey effort on heterogeneity and not the size of pond. This approach also increases the precision of monitoring efforts, to better characterize how populations of these rare fishes change over time.

Many freshwater fish monitoring efforts rely on a fixed number of samples per year (Bonar & Hubert 2002). For example, the Río Yaqui fish monitoring program has relied on 8 to 30 samples pond⁻¹, where the number was fixed through time and dependent upon pond size. However, catch estimates rarely

(11%) achieved the RSE25 threshold. Our analyses suggested ~40 traps pond⁻¹, on average, were needed to achieve RSE25 (Fig. 1). Hence, long-term fish monitoring efforts should not assume desired precision is being achieved with current levels of sampling. Instead, biologists should employ a simulation technique such as ours to determine the number of samples needed. In our situation, imagine a pilot effort using 10 traps and collecting 300 Yaqui chub in a pond. Given the equation in Table 1, approximately 31 traps (samples) would be required to attain a RSE25 (Fig. 2). However, abundance is a less desirable measure, because 9 traps may have captured 0 fish and 1 trap 300, generating high CV. Were CV high, say ~3.2, then ~113 traps would be required (Table 1, Fig. 2). Hence, while catch and CV inform sampling, practitioners should base sampling amount on the variability in catch, irrespective of the numbers caught. Otherwise, years of monitoring may be expended without achieving useful results.

When monitoring a community of sympatric fishes, monitoring programs strive to develop efficient sampling that effectively samples all species of interest. To improve Río Yaqui fish monitoring efforts, we evaluated and compared how 2 soak durations (i.e. set time) affected species-specific detection probabilities and abundance, by accounting for site level and temporal variability. Effects of the 4 h set time improved detection probability and abundance estimates of these endangered fishes, especially the historically less abundant beautiful shiner. Thus, just by altering soak time, we increased detection probability of beautiful shiner by 43 to 70%. Yaqui chub detection and abundance did not change significantly with soak time. Allowing traps to set for a longer period ensured that rarer, historically less-abundant and sympatric species like the beautiful shiner were monitored effectively.

No sampling design, standardized or not, can control for all factors influencing abundance and detection probability (Peterson et al. 2004, Price & Peterson 2010). Hence, well-designed monitoring efforts should account for detectability in estimates of abundance. Though many techniques exist that account for detectability (e.g. distance sampling, mark-resight, etc.; Williams et al. 2002), few are appropriate or efficient for small-bodied fishes. One technique, temporally replicated sampling in conjunction with mixture modeling (e.g. Royle & Nichols 2003, Royle 2004, Stewart & Long 2016) holds promise. Here, we developed a class of Bayesian hierarchical binomial mixture models that resulted in site-specific abundances and detection probabilities for

multiple populations (i.e. ponds). In addition, we demonstrated through the utility of a random effect that overdispersion in the data can be accounted for by estimating true abundance marginal to a hierarchical element. Importantly, our model accounts for extra sources of variation by accounting for variation from secondary sampling units within a primary sampling unit, and variation by the primary unit using exchangeable error terms. These contributions provide foundational work to develop spatial dependence models using spatial basis functions that our dataset did not allow. The outcome is unique and useful because most mixture models consider a single population. Often, more than one primary sampling unit (e.g. pond, field, grid, or management unit) is being sampled. As exemplified here, information can be leveraged from these multiple populations to (1) improve the predictive performance of the model to provide a means to estimate population-specific parameters, (2) provide inference to management actions or habitat, and (3) account for variation in both the secondary and primary sampling units using exchangeable error terms. Our approach illustrates a technique to derive a marginal probability distribution to account for extra-Poisson variation (i.e. overdispersion) in the data by including a hierarchical element into the Poisson likelihood.

We also examined site-specific correlates of detection and abundance using mixture modeling. Our results identified that detection probabilities varied considerably by species and ponds. This outcome is consistent with similar studies (Kennard et al. 2006, Price & Peterson 2010, Stewart & Long 2016). For example, we estimated that detection probability was much greater for Yaqui chub (~40%) than for Yaqui topminnow (~14%) and beautiful shiner (~14%), likely due to habitat factors affecting capture efficiency. We also estimated that fish were more difficult to detect in areas having a high percentage of submergent aquatic vegetation. This outcome was likely due to submergent aquatic vegetation enveloping the funnel of the Gee funnel minnow traps and preventing capture of fish.

Regardless of species, our results suggest that detection probability is not constant and a likely source of heterogeneity is habitat. These results highlight the inherent problems with monitoring programs that control for habitat factors using a standardized approach while relying on a fixed number of traps annually without considering how temporal and spatial variation in detection probability may affect precision and inference. Thus, the temporal variation in catch begs the question, 'Does temporal variability in

catch imply a change in population size or varying conditional detection probabilities?' Our work answers that question, showing that annual and site variation in catch statistics is a function of changing detectability.

Including these effects in the analysis enabled monitoring these species with a high degree of confidence. Indeed, this approach provided robust abundance estimates for these species in each pond (Table 3). While many details of habitat requirements of these 3 Río Yaqui fish species are unknown, our models indicate that the abundances of each species was strongly related to water depth and temperature, percent submergent aquatic vegetation, and trap location within a pond. Peak abundances of beautiful shiner and Yaqui chub occurred in deep water (>140 cm) with little submergent aquatic vegetation (<20%). Yaqui topminnows were predominately found in shallow water (<70 cm) and in open areas of the pond that contained little submergent aquatic vegetation (<60%). In developing a management strategy, the obvious first step is to consider the interaction between species and habitats, but also the interaction between the endangered Yaqui chub and threatened beautiful shiner that both occur in deeper, cooler water with less vegetation. The degree to which competition and predation affect these 3 species is unknown.

Lastly, we examined how variability in detection probability may mask trends in raw counts. Via simulation, we found that using simple catch to index abundance of small bodied fishes like the Río Yaqui fishes is limited and would likely result in spurious conclusions about population change. In our scenario of population doubling, we found the likelihood of committing Type I error rose to 50%, and statistical power declined to 0.75 given variable detectability. In scenarios of less abundance, power would decline even more. Hence, ignoring variable detectability in count data by assuming it constant can misrepresent species status, species-habitat relationships, and ultimately mislead conservation and stewardship of endangered species. In our situation, as depth increased (>70 cm), the probability of detecting beautiful shiner decreased. Contemporary analysis omitting detection would suggest that this species was more abundant in shallow water. Since our analysis accounted for detection, the inverse is true. Captures of the beautiful shiner were reduced in deeper water, where they are more abundant. Similarly, fish detection increased while abundance decreased with increasing water temperature. This example is not unique; Hangsleben et al. (2013) found that annual variability in capture

Table 3. Number captured, and mean probabilities of detection (p; 95 % CI) and site-specific abundance (N; 95 % CI) of beautiful shiner *Cyprinella formosa*, Yaqui chub *Gila purpurea*, and Yaqui topminnow *Poeciliopsis occidentalis sonoriensis* captured with Gee funnel minnow traps from 12 spring-fed ponds

Pond	Beautiful shiner			Yaqui chub			Yaqui topminnow		
	Catch	p	N	Catch	p	N	Catch	p	N
Ben New Pond	1	0.18 (0.09, 0.26)	4 (1, 17)	0	0.49 (0.05, 0.94)	0	487	0.05 (0.04, 0.06)	4937 (4112, 5552)
Brasher Pond	–	–	–	–	–	–	1 515	0.16 (0.12, 0.21)	4794 (3699, 6526)
Cobble Pond	107	0.07 (0.02, 0.12)	877 (469, 1232)	–	–	–	498	0.19 (0.11, 0.27)	1397 (925, 2364)
Hay North Pond	–	–	–	7	0.14 (0.05, 0.42)	39 (10, 99)	3316	0.1 (0.07, 0.31)	16447 (11870, 23724)
Hay South Pond	–	–	–	–	–	–	283	0.19 (0.07, 0.31)	832 (458, 1922)
Horse Tank	–	–	–	370	0.17 (0.08, 0.26)	1,182 (725, 2157)	–	–	–
Lower Twin Pond	7	0.25 (0.12, 0.39)	27 (7, 81)	47	0.83 (0.68, 0.93)	28 (27, 33)	548	0.16 (0.12, 0.21)	4794 (3698, 6526)
Magoffin Pond	7	0.13 (0.05, 0.37)	41 (11, 97)	1	0.5 (0.05, 0.95)	0	540	0.05 (0.04, 0.06)	5747 (5123, 6293)
Minckley North	–	–	–	1	0.33 (0.05, 0.91)	4 (1, 17)	739	0.34 (0.28, 0.40)	1104 (943, 1313)
Minckley South	–	–	–	0	0.5 (0.05, 0.95)	0	563	0.17 (0.09, 0.14)	1793 (1081, 3327)
Tennis Court	–	–	–	35	0.14 (0.05, 0.36)	170 (51, 377)	–	–	–
Twin Pond	120	0.06 (0.02, 0.10)	1019 (591, 1386)	0	0.5 (0.05, 0.95)	0	532	0.11 (0.05, 0.21)	2941 (1286, 5398)

efficiency of boat electrofishing, calculated from the capture of marked and unmarked individuals, coincided with reduced confidence in the utility of catch per unit effort of large-bodied fishes in large lentic systems. Thus, these and other environmental factors affect detection probability of other gear types, seasons, and environments, thus affecting the practitioners' ability to effectively draw inference to habitat relationships and trends in abundance if detection probability is not accounted for in the sampling design. Our study illustrates the effects of egregiously assuming constant detection probability when indexing abundance and trends of threatened and endangered fish species. Often species recovery plans are tied to abundances of species to describe status to ensure survival. Indeed the species recovery plan of the Yaqui topminnow, Yaqui chub, and beautiful shiner requires monitoring to determine if the species is secure for at least 10 yr (USFWS 1994). Similarly, monitoring programs of other endangered species like the desert pupfish *Cyprinodon macularius* and Kendall Warm Springs dace *Rhinichthys osculus thermalis* calculate an abundance estimate to inform species re-

covery plans and associated management actions related to water manipulations, habitat enhancement, and translocations (USFWS 1993, 2015). Since these monitoring programs rely on relative abundance and assume constant detection probability (Gryska 1996, Gryska & Hubert 1995, 1997, Martin & Saiki 2005), their ability to monitor true abundance and habitat associations is poor. In our example, it is impossible to accurately index the annual status of the endangered Yaqui topminnow, Yaqui chub, and beautiful shiner if we rely on simple counts uncorrected for detection probability. Using simple count in the beautiful shiner example would imply that this species was more abundant in shallow water—misinforming management. Elsewhere, simple catch statistics cloud understanding of tuna catch (Maunder et al. 2006), or are improperly used to elucidate habitat relationships of endangered Spanish toothcarp *Aphanius iberus* (Casas et al. 2011), and fish inhabiting mangroves (Bologna 2014). As our results suggest, remediating these problems requires accounting for variability in capture efficiency to produce robust abundance trends and species–habitat relationships.

SYNTHESIS

Many studies present methodologies to account for detection bias (Royle & Nichols 2003, Royle 2004, Stewart & Long 2016). Implementing these methods into survey designs by conservation and management agencies to inform abundance of common species, much less threatened and endangered species, remains lacking. In our experience, biologists often discount the importance of incorporating detection probability in animal surveys. Instead, attention is misplaced on simple counts to index species abundance and trends (assuming constant detection probability).

Our study highlights the importance of incorporating detection probability when developing and implementing surveys to monitor fish and wildlife populations, and the errors that can manifest when it is not. First, developing an adaptive sampling design (not relying on a fixed number of sites) to sample annually should be used for sampling highly fluctuating populations. This point was evidenced by our analyses, as we could not identify a fixed number of sites to use annually, given the high variation in temporal catch. This outcome is likely a result of varying detection probability and site-level covariates that remained unmeasured, and therefore unaccounted for. On average, however, the amount of samples was ~20 more than used historically. This result highlights the importance of pursuing power analyses to ensure sampling effort matches precision goals. Second, using an experimental design allowed us to formally evaluate 2 different soak times and the effects of environmental factors on detection probability. This practice enabled us to optimize the sampling strategy for the rare members of this fish community. Third, mixture models that integrate out the effects of heterogeneous detection probabilities should be used to estimate true abundance of fishes. Our simulations provide empirical support for incorporating detection probability when sampling animals, especially those that are notoriously difficult to sample because of their *r*-selected life history (Price & Peterson 2010, Stewart & Long 2016).

This information is important to consider when assessing status of endangered species. The novel techniques we illustrate can inform new surveys using existing data, test components of alternative sampling designs, and employ these data via simulations to determine the effects of egregious miscalculations, especially when biologists rely on simple count to inform conservation of rare and threatened fish or other animals. Conservationists must under-

stand how conservation objectives may be affected by the methods used to inform them. Only then will individuals and agencies embark on appropriate means to assess species populations to better inform their decisions and aid species recovery. Hence, biological programs must account for changes in detection probability in the estimation process to effectively and accurately measure abundance of species. Failure to account for detection probability can systematically bias measures of animal population status, reduce the ability to identify if a true change in population size occurred, and increase the probability (Type I error rate) of concluding that the population size changed when in fact it did not. Clearly, such outcomes are detrimental to the conservation and stewardship of rare and endangered species. Equally clear is that such outcomes are avoidable given proper project planning and forethought.

Acknowledgements. We thank Chris Lohrengel, Robert Hastings, and Matthew Christensen for assistance with data collection. Funding and additional support was provided by the US Fish and Wildlife Service. Any use of trade, firm, or product names is for descriptive purposes only and does not imply endorsement by the US Government. The findings and conclusions are those of the authors and do not necessarily represent the views of the US Fish and Wildlife Service.

LITERATURE CITED

- Anderson DR (2001) In my opinion: the need to get the basics right in wildlife field studies. *Wildl Soc Bull* 29: 1294–1297
- ✦ Bologna PAX (2014) Mangrove loss leads to fish hyperutilization of seagrass beds in a UNESCO biosphere reserve. *J Fish Biol* 84:1620–1625
- ✦ Bonar SA, Hubert WA (2002) Standard sampling of inland fish: benefits, challenges, and a call for action. *Fisheries* (Bethesda, MD) 27:10–16
- Bonar SA, Hubert WA, Willis DW (2009) Standard methods for sampling North American freshwater fishes. American Fisheries Society, Bethesda, MD
- ✦ Casas JJ, Sanchez-Oliver JS, Sanz A, Furne M and others (2011) The paradox of conservation of an endangered fish species in a Mediterranean region under agricultural intensification. *Biol Conserv* 144:253–262
- ✦ Dormann CF, Elith J, Bacher S, Buchmann C and others (2013) Collinearity: a review of methods to deal with it and a simulation study evaluating their performance. *Ecography* 36:27–46
- Dudley RK, Platania SP, White GC (2015) Rio Grande silvery minnow population monitoring program results from February to December 2014. Middle Rio Grande Endangered Species Collaborative Program. Albuquerque, NM
- ✦ Dumont SC, Schlechte JW (2004) Use of resampling to evaluate a simple random sampling design for general monitoring of fishes in Texas reservoirs. *N Am J Fish Manage* 24:408–416
- ✦ Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Stat Sci* 7:457–511

- Gryska AD (1996) Development of population monitoring protocols and description of several life history aspects of Kendall Warm Springs dace *Rhinichthys osculus thermalis*. MSc thesis, University of Wyoming, Laramie, WY
- Gryska AD, Hubert WA (1995) Development of population monitoring techniques and description of life history parameters for the Kendall Warm Springs dace. Progress report submitted to the US Fish and Wildlife Service, Cooperative Agreement No: 14-48-0009-94-1542. US Fish and Wildlife Service, Cheyenne, WY
- Gryska AD, Hubert WA (1997) Observations on the reproduction, sources of mortality and diet of the Kendall Warm Springs dace. *Great Basin Nat* 57:338–342
- Hangsleben MA, Allen MS, Gwinn DC (2013) Evaluation of electrofishing catch per unit effort for indexing fish abundance in Florida lakes. *Trans Am Fish Soc* 142: 247–256
- Hardin S, Conner LL (1992) Variability of electrofishing crew efficiency, and sampling requirements for estimating reliable catch rates. *N Am J Fish Manage* 12:612–617
- Hyde EJ, Simons TR (2001) Sampling plethodontid salamanders: sources of variability. *J Wildl Manage* 65:624–632
- IUCN (2016) The IUCN Red List of Threatened Species, version 2016-1. www.iucnredlist.org (accessed 31 August 2016)
- Jerde CL, Mahon AR, Chadderton WL, Lodge DM (2011) 'Sight-unseen' detection of rare aquatic species using environmental DNA. *Conserv Biol* 4:150–157
- Kellner KF (2015) jagsUI: a wrapper around 'rjags' to streamline 'JAGS' analyses. R package version 1.3.1
- Kennard MJ, Pusey BJ, Harch BD, Dore E, Arthington AH (2006) Estimating local stream fish assemblage attributes: sampling effort and efficiency at two spatial scales. *Mar Freshw Res* 57:635–653
- Kéry M, Royle JA (2010) Hierarchical modelling and estimation of abundance and population trends in metapopulation designs. *J Anim Ecol* 79:453–461
- Kuehl RO (2000) Design of experiments: statistical principles of research design and analysis, 2nd edn. Duxbury Press, Pacific Grove, CA
- Kuznetsova A, Brockhoff PB, Christensen RHB (2013) lmerTest: tests for random and fixed effects for linear mixed effect models (lmer objects of lme4 package). R package version 2.0-0. <http://cran.r-project.org/web/packages/lmerTest/index.html> (accessed 15 Jan 2016)
- Lindén A, Mäntyniemi S (2011) Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology* 92:1414–1421
- Magnuson JJ, Benson BJ, McLain AS (1994) Insights on species richness and turnover from long-term ecological research: fishes in north temperate lakes. *Am Zool* 34: 437–451
- Martin BA, Saiki MK (2005) Relation of desert pupfish abundance to selected environmental variables in natural and manmade habitats in the Salton Sea basin. *Environ Biol Fishes* 73:97–107
- Maunder MN, Punt AE (2004) Standardizing catch and effort data: a review of recent approaches. *Fish Res* 70: 141–159
- Maunder MN, Sibert JR, Fonteneau A, Hampton J, Kleiber P, Harley SJ (2006) Interpreting catch per unit effort data to assess the status of individual stocks and communities. *ICES J Mar Sci* 63:1373–1385
- Minckley WL (1985) Native fishes and natural aquatic habitats of US Fish and Wildlife Service Region II, west of the Continental Divide. Final report for USFWS-AZSU. Arizona State University, Tempe, AZ
- Peterson JT, Thurow RF, Guzevich JW (2004) An evaluation of multipass electrofishing for estimating the abundance of stream-dwelling salmonids. *Trans Am Fish Soc* 133: 462–475
- Plummer M (2003) JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Hornik K, Leisch F, Zeileis A (eds) Proceedings of the 3rd international workshop on distributed statistical computing, 20–22 March, Vienna
- Pollock KH, Simons TR, Farnsworth GL, Bailey LL, Sauer JR (2002) Large scale wildlife monitoring studies: statistical methods for design and analysis. *Environmetrics* 13: 105–119
- Price AL, Peterson JT (2010) Estimation and modeling of electrofishing capture efficiency for fishes in wadeable warmwater streams. *N Am J Fish Manage* 30:481–498
- R Core Team (2015) R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna
- Robson DS, Regier HA (1964) Sample size in Peterson mark-recapture experiments. *Trans Am Fish Soc* 93:215–226
- Royle JA (2004) *N*-mixture models for estimating population size from spatially replicated counts. *Biometrics* 60: 108–115
- Royle JA, Nichols JD (2003) Estimating abundance from repeated presence-absence data or point counts. *Ecology* 84:777–790
- Stewart DR, Long JM (2012) Precision of channel catfish catch estimates using hoop nets in larger Oklahoma reservoirs. *N Am J Fish Manage* 32:1108–1112
- Stewart DR, Long JM (2016) Using hierarchical Bayesian multi-species mixture models to estimate tandem hoop-net based habitat associations and detection probabilities of fishes in reservoirs. *Trans Am Fish Soc* 145: 450–461
- Stewart DR, Long JM, Shoup DE (2015) Spatial structuring within a reservoir fish population: implications for management. *Mar Freshw Res* 66:202–212
- USFWS (US Fish and Wildlife Service) (1993) Desert pupfish recovery plan. US Fish and Wildlife Service, Phoenix, AZ
- USFWS (1994) Yaqui fishes recovery plan. US Fish and Wildlife Service, Albuquerque, NM
- USFWS (2015) Recovery plan for the Kendall Warm Springs dace (*Rhinichthys osculus thermalis*). Revision: original approved July 12, 1982. US Fish and Wildlife Service, Cheyenne, WY
- Wagner T, Vandergoot CS, Tyson J (2009) Evaluating the power to detect temporal trends in fishery-independent surveys: a case study based on gill nets set in the Ohio waters of Lake Erie for walleyes. *N Am J Fish Manage* 29:805–816
- Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations: modeling, estimation and decision making. Academic Press, San Diego, CA