# EVALUATING STATISTICAL APPROACHES TO QUANTIFYING JUVENILE CHINOOK SALMON HABITAT IN A REGULATED CALIFORNIA RIVER

M. P. BEAKES[a,e*,†], J. W. MOORE[a,e†], N. RETFORD[a,b], R. BROWN[c], J. E. MERZ[d] AND S. M. SOGARD[b]

[a] *University of California, Santa Cruz, CA 95064, USA*
[b] *National Marine Fisheries Service, Santa Cruz, CA 95060, USA*
[c] *University of California, Davis, CA 95616, USA*
[d] *Cramer Fish Sciences, West Sacramento, CA 95691, USA*
[e] *Simon Fraser University, Earth to Ocean research group, British Columbia V5A 1S6, Canada*

## ABSTRACT

Decisions on managed flow releases in regulated rivers should be informed by the best available science. To do this, resource managers require adequate information regarding the tradeoffs between alternative methodologies. In this study, we quantitatively compare two competing multivariate habitat models for juvenile Chinook salmon (*Oncorhynchus tschawytscha*), a highly valued fish species under serious decline in a large extent of its range. We conducted large-scale snorkel surveys in the American River, California, to obtain a common dataset for model parameterization. We built one habitat model using Akaike Information Criterion analysis and model averaging, 'model G', and a second model by using a standard method of aggregating univariate habitat models, 'model A'. We calculated Cohen's kappa, percent correctly classified, sensitivity, specificity and the area under a receiver operator characteristic to compare the ability of each model to predict juvenile salmon presence and absence. We compared the predicted useable habitat of each model at nine simulated river discharges where usable habitat is equal to the product of a spatial area and the probability of habitat occupancy at that location. Generally, model G maintained greater predictive accuracy with a difference within 10% across the diagnostic statistics. Two key distinctions between models were that model G predicted 17.2% less useable habitat across simulated flows and had 5% fewer false positive classifications than model A. In contrast, model A had a tendency to over predict habitat occupancy and under predict model uncertainty. The largest discrepancy between model predictions occurred at the lowest flows simulated and in the habitats most likely to be occupied by juvenile salmon. This study supports the utility and quantitative framework of Akaike Information Criterion analysis and model averaging in developing habitat models. Copyright © 2012 John Wiley & Sons, Ltd.

KEY WORDS: salmonids; habitat models; River 2D; AIC; model averaging; flow regulation

## INTRODUCTION

With over half of the earth's large river systems currently dammed (Nilsson *et al.*, 2005), there is an increasing need for robust quantitative tools that can predict the impacts of flow regulation on riverine ecosystems (Petts, 2008). Dams can allow water managers to control when and how much water flows downstream. As such, these quantitative tools must be able to provide resource managers with information that accurately reflects the tradeoffs and their uncertainties between resource use and any potential ecological impacts of flow regulation. Understanding the ecological consequences of different flow releases can help balance the needs for energy generation, water storage, flood control, and downstream aquatic communities.

Over the last few decades, numerous competing techniques have been developed to aid managers in predicting how changes in river flow will modify inhabitable space for riverine species downstream of dams (Manel *et al.*, 2001; Ahamadi-Nedushan *et al.*, 2006; Mouton *et al.*, 2010; Dunbar *et al.*, 2011). These statistical techniques range considerably in complexity, but the common aim is to estimate how flow regulation alters physical characteristics of rivers (e.g. velocity and depth) and predict how those changes will impact individual species. For example, Bovee (1986) developed a statistical method where several habitat variables (e.g. depth, water velocity, cover or substrate) are parameterized independently and then combined as univariate model estimates into a composite index of habitat suitability (CSI). The CSI can be constructed by one of several available methods (e.g. geometric mean, arithmetic mean and product). The CSI approach has historically been integrated into a physical habitat simulation (PHABSIM), which is a fundamental component of the instream flow incremental methodology (IFIM). The IFIM and PHABSIM are used to inform

*Correspondence to: Michael P. Beakes, Simon Fraser University, Earth to Ocean research group, 8888 University Dr. Burnaby, British Columbia V5A 1S6, Canada.
E-mail: mbeakes@sfu.ca
†Present address: Simon Fraser University, Earth to Ocean research group, British Columbia V5A 1S6, Canada.

management decisions in regulated river systems, such as setting minimum flow standards and quantifying flow regulation impacts on aquatic habitats (Stalnaker *et al.*, 1995; Waddle, 2001). The CSI approach is still integrated into IFIM and contemporary management decisions and is published in peer-reviewed literature (e.g. Ayllon *et al.*, 2010; Lee *et al.*, 2010; Boavida *et al.*, 2011; Im *et al.*, 2011). However, this method has several key assumptions that have been subject to considerable criticism over the last few decades. Specifically, the CSI approach requires assumptions that each parameter is selected independently by the target species (Bovee, 1986), that each variable is equally important and that the covariance structure among variables is negligible (Mathur *et al.*, 1985; Beecher *et al.*, 2002; Jowett, 2003; Leclerc *et al.*, 2003). Furthermore, the CSI approach often ignores any uncertainty in model predictions (Burgman *et al.*, 2001).

In contrast to the CSI approach, more complex techniques that allow numerous parameters to be estimated simultaneously have been developed in the past decade, such as generalized linear models (GLMs; Guisan *et al.*, 2002). GLMs have only recently been applied in aquatic habitat modelling (Labonne *et al.*, 2003; Ahamadi-Nedushan *et al.*, 2006), but many of the statistical flaws and assumptions of the CSI approach are addressed with this technique; thus, our ability to describe ecological data has greatly improved (Guisan *et al.*, 2002; Ahamadi-Nedushan *et al.*, 2006). In addition, previous research has demonstrated that multivariate methods produce dissimilar predictions of usable habitat in comparison with a CSI method, and multivariate techniques provide a greater amount of information (Vismara *et al.*, 2001). However, resource managers have been slow to adopt new methods, and multivariate techniques have received relatively little attention (Vismara *et al.*, 2001; Dunbar *et al.*, 2011).

The slow progression of new methods into the management community may be a result of inadequate information allowing resource managers to distinguish the tradeoffs between alternative techniques. For example, Manel *et al.* (2001) reviewed 87 articles published in ecological literature between 1989 and 1999 and reported that over 67% of the studies using presence–absence models failed to attempt any kind of model evaluation. In addition, there are relatively few published articles focused on aquatic habitat models that compare competing methodologies on a common data set (Ahamadi-Nedushan *et al.*, 2006). As a consequence, there is a paucity of knowledge to make informed decisions about methodologies for estimating how flow regulation alters aquatic habitat.

The aim of this study is to quantitatively compare two statistical methodologies and examine the potential role of Akaike Information Theory and model averaging (AICc, Akaike, 1974; Burnham and Anderson, 2002) in aquatic habitat modelling. This study builds on comprehensive

reviews of the variety of methods available to resource managers by Manel *et al.* (2001), Ahamadi-Nedushan *et al.* (2006), Mouton *et al.* (2010) and Dunbar *et al.* (2011), and submits a novel application of AICc model averaging for estimating the impacts of flow regulation on habitat for the juvenile life stage of Chinook salmon (*Oncorhynchus tshawytscha*), a highly valued fish species under serious decline in a large extent of its range (Myers *et al.*, 1998). We use AICc analysis and model averaging to construct a multivariate GLM and develop a second habitat model, following the CSI approach, comprised of aggregated univariate models. We compare the models with five diagnostic statistics deemed appropriate for gauging model performance during model development (Mouton *et al.*, 2010). In addition, we include estimates of uncertainty around each model's predictions and extrapolate these predictions under several flow scenarios to gain perspective on the tradeoffs between selecting one model over the other. We hypothesize that the AICc-averaged habitat model will have greater predictive accuracy and provide a more robust and conservative prediction of the relative impact of flow regulation on juvenile salmon habitat. Therefore, results from this study aim to advance an existing foundation for hydrodynamic habitat model development and application (Petts, 2008).

## METHODS

### Study system

This study was conducted in the Lower American River (LAR), which is primarily a snow-fed system, draining approximately $4900 \, km^2$ of the Sierra Nevada Mountains in Northern California. Like other California Central Valley rivers, the American River has been highly modified from its historic state, including flow regulation and diversion, water pollution, gold and gravel mining, hydropower and floodplain development, and the introduction of numerous non-native aquatic species (McEwan, 2001; Williams, 2001; Moyle, 2002). Just downstream of the American River north and south fork confluences, Folsom Dam was completed in 1955, blocking upstream habitat for migratory fishes such as anadromous salmonids. The Bureau of Reclamation currently operates the dam for flood control, water storage and hydroelectric generation. The LAR is defined as the 37 km of unobstructed channel that flows downstream of Nimbus Dam, which is located approximately 11 km downstream from Folsom Dam. This portion of the river still provides spawning and rearing habitat for anadromous steelhead (*Oncorhynchus mykiss*) and Chinook salmon (Yoshiyama *et al.*, 2001). American River fall-run Chinook salmon typically spawns from late September to December, with juvenile rearing from early January to June. Our study

reach is approximately 800 m long and located within the Sacramento city limits, just downstream of the American River Parkway, Sunrise recreation area (Figure 1).

### Fish and habitat surveys

Following methods of Gard (2006), we attained our habitat occupancy data via large-scale snorkel surveys conducted from February to July (2009, 2010) across as many different accessible habitat types as feasible (e.g. riffles, runs and backwater habitat). We conducted the snorkel surveys during daylight along a linear transect from downstream to upstream and marked occupied locations of juvenile Chinook salmon (fork length >40 mm, <95 mm) with flagged weights. Fish sizes were visually estimated to the nearest 5 mm fork length. Regardless of the number of fish at occupied locations, each observation counted as a single statistical unit in our common dataset and analysis. When depth and velocity were too high to allow unaided upstream snorkelling, we used fixed ropes to facilitate our upstream movement. We then measured physical habitat characteristics (i.e. velocity, depth, cover and substrate) in flagged locations. Depth was measured as the distance from the water surface to the riverbed to the nearest centimetre, and velocity was measured to the nearest centimetre per second with a portable electromagnetic velocity flow metre (Marsh-McBirney, Flo-Mate) at 60% depth. Substrate diameter was visually estimated to the nearest centimetre except for particles less than 1 cm, which were classified as sand (0.25 cm) or fine sediment (0.025 cm). At regular intervals along each snorkel transect (10 or 20 m), we randomly selected unoccupied locations using a random number generator, excluding locations less than 1 m from occupied locations, and measured the same suite of habitat characteristics as those measured in occupied locations. These data collection efforts thus resulted in a series of fish presence–absence data with associated habitat attributes.

We mapped dominant habitat cover into polygons using two categories (with and without cover) over the entire study reach. These cover types were similar to those described in Gard (2006). We identified locations with cover when they were within 50 cm of large woody debris (>7.5 cm diameter), tall vegetation (>50 cm above ground), overhanging vegetation (<50 cm from water's surface), large boulders (>17.5 cm diameter), undercut banks, large bedrock crevasses or combinations of these cover types. Areas without cover included characteristics such as small vegetation, small substrate (<17.5 cm diameter) or filamentous algae. We used tools in Arc Map to spatially join cover attributes from delineated polygons to every node, or point, in the hydrodynamic model.
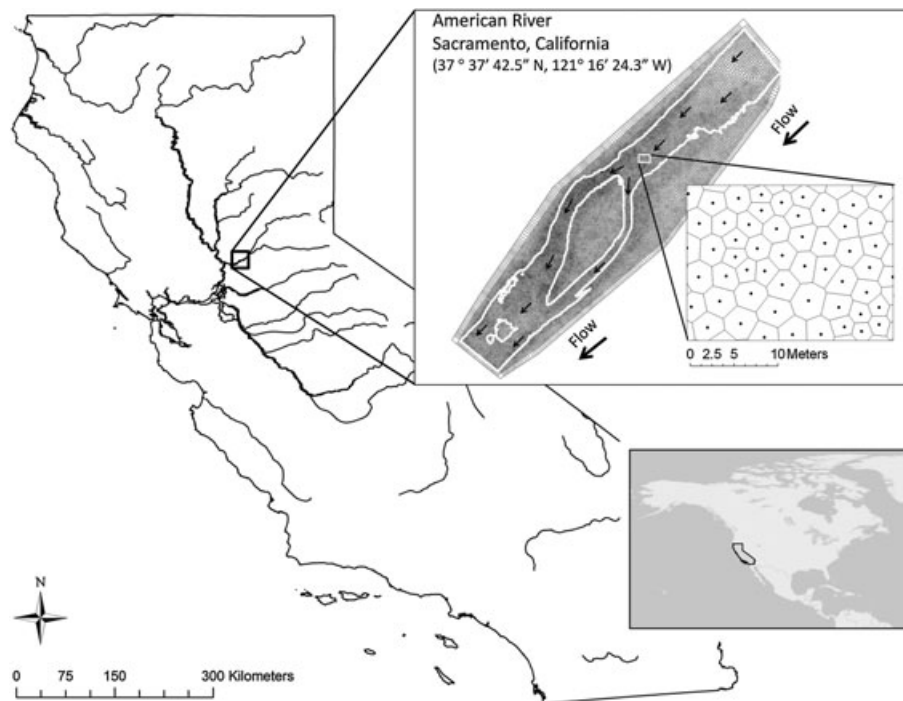


Figure 1. Map of study reach in the American River, California. Thiessen polygons (inset) encompassing the nodes generated in River 2D were used to estimate the area around each point for integration of the Hydrodynamic and habitat model estimates. Waters edge from the lowest (33.98 m³ s⁻¹) discharge model outlined in white and the direction of flow indicated by bold black arrows.

## Competing habitat models

We analysed the binary occupied and unoccupied snorkel survey data via polynomial GLM with a logit link function (i.e. multiple logistic regression) in program R (R development core team, 2011), which estimates the probability of habitat occupancy as a function of the predictive habitat variables. Logistic regression has been used for similar modelling exercises in numerous studies and is arguably the most appropriate for presence–absence data (Ahamadi-Nedushan *et al.*, 2006).

We used two competing methods to develop habitat models that predict the probability of habitat occupancy by our study species. In the first method, we estimated the probability of habitat occupancy for depth, velocity, substrate and cover independently of one another. We assumed that the logistic relationship for velocity, depth and substrate was parabolic and thus assigned a quadratic term to each variable (e.g. $velocity + velocity^2$). In addition, we tested all possible interactions among velocity, depth and cover, of which none were significant predictors of habitat occupancy ($p > 0.05$) and were subsequently excluded from further analysis. We selected statistically significant univariate habitat occupancy models and aggregated them by taking the geometric mean of the model predictions for each corresponding variable at a point in space. This method is comparable with the CSI approach, supported by the US Fish and Wildlife Service standards for habitat suitability index model development (USFWS, 1981), recommended in contemporary habitat modelling software (e.g. River 2D) and applied in recent field studies (e.g. Ayllon *et al.*, 2010). We estimated upper and lower confidence limits for this model by adding 1.96 standard errors, estimated with the 'predict' function in program R, to each point-specific probability of habitat occupancy. We refer to this model as 'model A' (for aggregated) hereafter.

We developed a second model with information theory via AICc and model averaging corrected for small sample sizes (AICc, Akaike, 1974; Hurvich and Tsai, 1989; Burnham and Anderson, 2002). We compared the most complicated model with all possible model combinations that did not include interaction terms to rank them in order of model parsimony. A difference greater than four AICc units between models can be interpreted as evidence for model superiority (Burnham and Anderson, 2002). For each model and corresponding AICc score, we calculated an AICc weight, which is an estimate of the relative support for each model across all the models compared. The sum of all model weights equals 1; thus, an AICc weight of 0.25 is analogous to having 25% of the relative support across models. We averaged model coefficients across the 95% confidence model set (summed weight, Burnham and Anderson, 2002) because the AICc weight of the 'top' model was <0.9 (as recommended by Grueber *et al.*, 2011).

Specifically, we averaged the parameter coefficients of the full model proportionally to the coefficients in the 95% confidence model set based on the AICc weight and relative support of each model. The main function of model averaging was to avoid losing information explained by any of the competing models. As such, the parameter coefficients of the averaged model represented a weighted average of the coefficients of all candidate models based on their relative support, or AICc weight (Burnham and Anderson, 2002). By model averaging, we allowed for comparably supported models to influence the parameter coefficients of the global model and less supported models to have limited effects on the global model. We estimated upper and lower confidence limits for this model by adding 1.96 standard errors to each point-specific probability of habitat occupancy. We refer to the AICc averaged model as 'model G' (for global) hereafter.

## 2D hydrodynamic model

We used a two-dimensional depth-averaged hydrodynamic model (River 2D), following methods described in Steffler and Blackburn (2002), to estimate point-specific values of depth and velocity at different flow rates within our study reach. River 2D is based on the St. Venant shallow water equations for shallow flows in natural streams, in conservation form (Steffler and Blackburn, 2002). The key parameter inputs include a digital elevation model (DEM) of the study reach, estimates of riverbed roughness height and measured water surface elevations (WSELs) at the modelled discharge. The modelling process includes the development of a triangulated irregular network between points ('nodes') that are overlaid on top of the DEM at a user specified density. By spatial linear interpolation of the DEM, each node is assigned *x*, *y* and *z* coordinates and an estimate of bed roughness height. The user inputs the WSEL at the upstream and downstream boundaries along with the corresponding discharge. The River 2D software then solves a hydraulic algorithm using Manning's N and the conservation of mass and momentum at each node. The end result is an estimate of velocity and depth, for a specified discharge, at each node in the study reach. For greater detail in the modelling procedure and model output, see Steffler and Blackburn (2002).

We constructed our study site DEM with topographic data that were previously collected on the LAR and enriched this data set with additional topographic surveys (2009, 2010). All additional topographic data were collected with a Topcon brand survey-grade RTK GPS in NAD (1983) State Plane, California Zone II, FIPS 0402 (Feet). We removed erroneous topographic points with spatial tools in ArcGIS (version 9.3). The average topographic survey density for the study reach was $0.13\,points\,m^{-2}$, and the hydrodynamic model node

density was consistent between models and comparable with the topographic point density at 0.15 (nodes m$^{-2}$).

We mapped median riverbed substrate ($D_{50}$) over the entire study reach to characterize bed roughness in the 2D hydrodynamic model. We encompassed homogeneous regions of substrate with a large polygon, delineating the boundaries with the RTK GPS. In highly heterogeneous areas, we took point measurements and developed Thiessen polygons to encompass these points in ArcGIS. Using the equation from Thompson and Campbell (1979), we estimated bed roughness height ($k_s$) as $k_s = 4.5 D_{50}$. We used spatial join functions in ArcGIS to assign bed roughness height to each topographic point in the consolidated topography dataset.

We measured WSELs at four discharges (49.84, 55.39, 111.63 and 131.74 m$^3$ s$^{-1}$) and extrapolated WSELs for discharges of 33.98, 70.79, 87.78, 144.42 and 169.90 m$^3$ s$^{-1}$. We considered these flows to be within a conservative range of interpolation and extrapolation from our measured WSELs and discharges. These flows were representative of the mean daily flow frequency distribution post dam construction during salmonid rearing months, February to July. We then completed the modelling procedure in River 2D for each flow level.

We gauged the River 2D model performance on the basis of three criteria. First, we determined if a stable hydrodynamic solution was reached by evaluating the change in model variable estimates between the final model iterations. Specifically, the difference of the square root of sum of squares for all hydrodynamic model variables between the final model iterations needed to be less than 0.1% (Steffler and Blackburn, 2002). Second, simulated outflow (m$^3$ s$^{-1}$) was within 1% of the measured inflow (m$^3$ s$^{-1}$) for all discharges modelled. The third criteria is the absence of supercritical flow, or low Froude (ratio of inertial force/gravitation force) throughout the modelled reach; in natural systems, Froude numbers rarely exceed 1.0 (Grant, 1997). The resulting hydrodynamic model was combined with each of the habitat models described previously to develop estimates of usable habitat at different discharges.

### Model performance

To compare the performance of the competing habitat models (model A versus model G), we used program R (R development core team, 2011) and the 'PresenceAbsence' library (Freeman and Moisen, 2008) to calculate optimized thresholds for models A and G. We set the threshold value where Kappa was maximized for both models and used this threshold value to estimate Kappa and three additional model performance statistics. Each statistic is a measure of the capacity to accurately discriminate the correct outcome of our habitat occupancy data, where probabilities that exceeded the threshold would be classified as occupied (positive) and those below

the threshold would be classified as absent (negative). We evaluated model performance using Cohen's Kappa statistic, percent correctly classified (PCC), sensitivity, specificity and the area under a receiver operator characteristic (AUC). The Kappa statistic is a measure of all possible outcomes of presence or absence that are predicted correctly, after accounting for chance predictions; it is generally accepted as a conservative and standardized metric for comparing the predictive accuracy of binary models regardless of their statistical algorithm (Manel *et al.*, 2001). PCC compares the proportion of outcomes correctly classified. In this application, sensitivity represents the proportion of true positives correctly identified, and specificity is the proportion of true negatives correctly identified, where 1-specificity is the proportion of false positives. Kappa, PCC, sensitivity and specificity are all threshold dependent statistics, so we also included AUC as a threshold independent statistic for model diagnostics. AUC is a measure of model accuracy across all potential thresholds in binary models. In our study, an AUC value of 1 is equivalent to perfect model agreement between observed and predicted habitat occupancy outcomes. An AUC value of 0.5 is approximately equal to random predictions of habitat occupancy.

### Model extrapolation and usable habitat

To be useful in management, these models must be able to identify areas most likely to be occupied by a target species and to predict how the total area that is likely to be occupied will change as river flow is changed. To do this, we calculated the point-specific predictions from each habitat model and extrapolated those predictions over space for several flow scenarios. We constructed three quantitative comparisons between models. First, we plotted all point-specific probabilities of each model's predictions against one another to determine if there were systematic differences between competing models. Second, we calculated and plotted the absolute difference in point-specific model predictions for several flows to identify spatial disparities among flow scenarios. Third, we estimated the total usable habitat predicted by each model and included estimates of uncertainty around those predictions.

To extrapolate habitat model predictions across space, we first developed Thiessen polygons in ArcGIS for each node in all nine hydrodynamic models (one for each discharge level) to estimate an area around each node (Figure 1). We used the estimated velocity, depth and cover at each node across the study reach, for all nine simulated discharges, as independent parameter inputs for both habitat models mentioned previously and predicted a probability of habitat occupancy for each node. By multiplying this probability by the area encompassing each node, we calculated the usable habitat for each polygon and then summed the resulting values across the entire study reach to quantify the total usable habitat for

each discharge. However, we restricted extrapolation to velocities $<120\,cm\,s^{-1}$ and depths $<113\,cm$ to avoid extrapolating model predictions beyond the physical conditions observed in our habitat data. By multiplying the upper and lower confidence limits of each predicted probability by the area encompassing each node, we calculated the minimum and maximum estimates of usable habitat for each polygon and again summed the resulting values across the entire study reach for each discharge to estimate uncertainty around usable habitat predictions. Our aim was to identify any systematic difference in model predictions across space and under different river flows, thereby providing a context for determining the tradeoffs of using one model or the other in river management.

## RESULTS

### Habitat occupancy data

Our snorkel surveys covered a total area of approximately $182\,000\,m^2$ along a total of approximately 22.5 river kilometres, and we observed a mean fish density of 0.005 juvenile Chinook $m^{-2}$. The mean stream temperature we observed during the dates of our snorkel surveys was $12.9\,°C$ ($\pm 1.5\,°C$ S.D., United States Geologic Survey Station ID. 11446500, Fair Oaks Ave, California). We observed 88 habitat locations occupied by juvenile Chinook within our target size class and 391 unoccupied habitat locations. These data provided a common dataset for habitat model development and testing.

Table I. Model A coefficients, standard errors and $p$-values for each independently fit polynomial logistic regression. In all analysis, 'Cover' is categorical and unitless. Specifically, 'Cover' is defined as the presence or absence of one or more cover types (e.g. large wood, overhanging vegetation). Except for cover, we assumed that the correlation between habitat occupancy and our predictor variables was non-linear and thus included a quadratic term with each parameter

| Parameter | Coefficient | Standard error | $p$ |
|---|---|---|---|
| Intercept | −1.7393 | 0.2380 | <0.001 |
| Velocity | 0.0290 | 0.0143 | 0.043 |
| Velocity$^2$ | −0.0004 | 0.0002 | 0.020 |
| Intercept | −2.1877 | 0.4523 | <0.001 |
| Depth | 0.0582 | 0.0254 | 0.022 |
| Depth$^2$ | −0.0009 | 0.0003 | 0.010 |
| Intercept | −1.1098 | 0.3266 | <0.001 |
| Substrate | −0.1133 | 0.0846 | 0.180 |
| Substrate$^2$ | 0.0060 | 0.0046 | 0.187 |
| Intercept | −1.5921 | 0.1311 | <0.001 |
| Cover | 0.6324 | 0.3066 | 0.039 |

### Model A

Independent factors of velocity, depth and cover were significant (Table I, $p < 0.05$) predictors of habitat occupancy. There were no significant interactions between any variables, and they were excluded from further analysis. In addition, substrate was not significantly correlated with habitat occupancy (Table I, $p > 0.05$) so it was excluded from model A. However, large substrate ($>17.5\,cm$ diameter) was included as a type of cover and thus included in the cover variable. The predicted optimal depth ($33.8\,cm$) and velocity ($36.7\,cm\,s^{-1}$) maintained a probability of habitat occupancy of 23.1% and 23.0%, respectively (Figure 2). In the absence of cover, the predicted probability of habitat occupancy was 16.9%, which increased to 27.7% in the presence of cover (Figure 2).
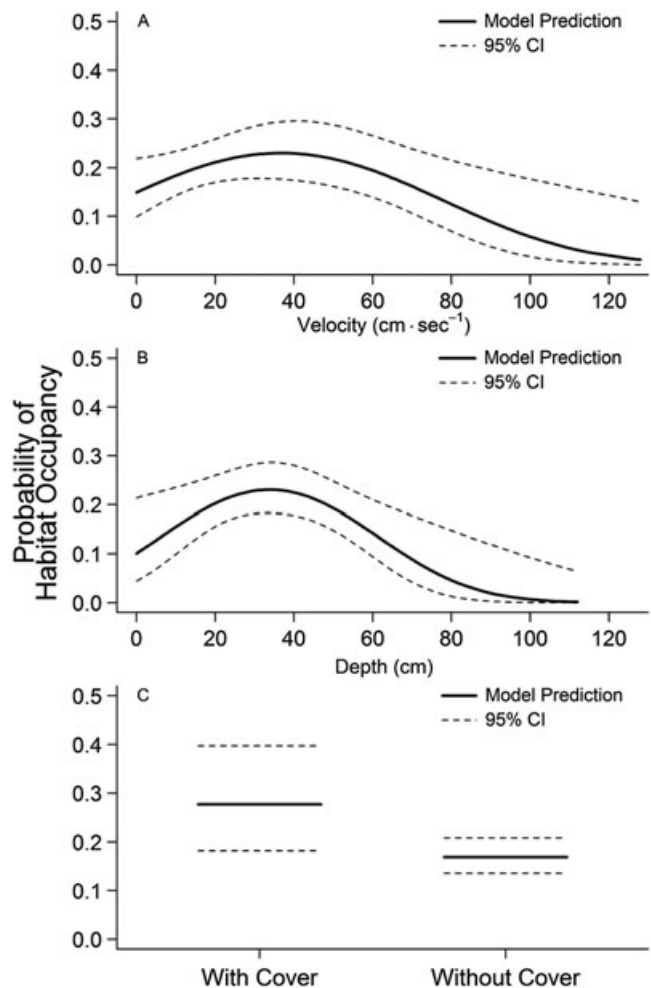


Figure 2. Univariate logistic regression (solid line) plotted as the predicted probability of habitat occupancy for (A) velocity, (B) depth and (C) cover. The fine dashed lines indicate the 95% confidence intervals (CI) of model predictions, estimated with the predict function in program R.

*Model G*

There were six candidate models that were comparably supported by AICc analysis ($\Delta$AICc < 4). The top-supported model, with an AICc weight ($w_i$) of 0.34 or 34% of the relative support, was velocity + velocity$^2$ + depth + depth$^2$ + cover (Table II). All but one of the top six models ($\Delta$AICc 4) were multivariate, with a collective 81% of the total support, indicating that a combination of predictor variables was far more supported than the univariate models. In total, multivariate models received 88% of the total support within the 95% candidate model set (Table II).

We adjusted the coefficient estimates in the global model for further analysis on the basis of the AICc weight of the 95% confidence model set (Table III), thereby accounting for the relative support of each model. Depth was the most strongly supported variable in AICc analysis, with relative variable support of 95%. Cover and velocity were also strongly supported variables at 72% and 62%, respectively. The inclusion of substrate as a predictor variable had the least support at 13%. The maximum predicted probability of habitat occupancy for model G was 34.9%, with velocity at $41.2\,\mathrm{cm\,s}^{-1}$, with depth at 32.7 cm, with substrate at 20.32 cm and with cover present. The probability of habitat occupancy increased with increased substrate size. However, the probability of habitat occupancy changed by less than 2% at the smallest and largest substrate sizes in our model. The presence of cover increased the probability of habitat occupancy by up to 9.5% when depth and velocity were at their optimal level (Figure 3). To illustrate the general correlation between the continuous variables in model G, we predicted the probability of habitat occupancy when depth was held constant at 2.7, 32.7 or 62.7, substrate constant at 7.62 cm, with and without cover (Figure 3).

*Model performance*

The optimum threshold values for models A and G were similar and estimated at 0.201 and 0.231, respectively. We

Table II. AICc 95% candidate model set and corresponding AICc score and AICc weight ($W_i$). A change greater than four AICc units ($\Delta$AICc) is evidence of model superiority. The AICc weight is a proportional measure representing the relative support estimated with AICc analysis for each competing model

| Model rank | Model parameters | AICc | $\Delta$AICc | $W_i$ |
|---|---|---|---|---|
| 1 | $V + V^2 + D + D^2 + C$ | 449.86 | 0.00 | 0.34 |
| 2 | $D + D^2 + C$ | 450.64 | 0.78 | 0.23 |
| 3 | $V + V^2 + D + D^2$ | 451.71 | 1.85 | 0.13 |
| 4 | $D + D^2$ | 452.09 | 2.23 | 0.11 |
| 5 | $V + V^2 + D + D^2 + S + S^2 + C$ | 453.45 | 3.59 | 0.06 |
| 6 | $V + V^2 + C$ | 453.55 | 3.69 | 0.05 |
| 7 | $D + D^2 + S + S^2 + C$ | 454.12 | 4.26 | 0.04 |
| 8 | $V + V^2 + D + D^2 + S + S^2$ | 454.44 | 4.58 | 0.03 |

V, velocity; D, depth; S, substrate; C, cover.

Table III. Model averaged coefficients and standard errors for model G. We cannot calculate a *p*-value for each of our parameters using model averaging and information theory so it is not provided. The coefficients are calculated as a weighted average of coefficients from each model in our candidate model set

| Parameter | Averaged coefficient | Standard error |
|---|---|---|
| Intercept | −2.2724 | 0.4876 |
| Velocity | 0.0293 | 0.0157 |
| Velocity$^2$ | −0.0004 | 0.0002 |
| Depth | 0.0511 | 0.0268 |
| Depth$^2$ | −0.0008 | 0.0003 |
| Substrate | −0.0727 | 0.0927 |
| Substrate$^2$ | 0.0042 | 0.0049 |
| Cover | 0.6316 | 0.3263 |

estimated each performance statistic using the optimized thresholds and our presence and absence dataset. Model G had the highest Kappa statistic and correctly predicted 20.3% of all possible presence and absence data, adjusted for correct predictions by chance. In contrast, model A had a Kappa statistic of 18.9%. Model G also had a higher PCC statistic compared with model A at 75.2% and 72%, respectively. Model G predicted 37.5% of the true positives (sensitivity) correctly and 83.6% of the true negatives correctly (specificity), with a false positive classification of 16.4%. In contrast, model A predicted 43.2% proportion of the true positives (sensitivity) correctly and 78.5% of the true negatives correctly (specificity), with a false positive classification of 21.5%. The AUC value, a threshold independent statistic, for models G and A was 0.649 and 0.647, respectively, indicating that both models maintained fair predictive ability above random chance.

*2D hydrodynamic model*

All hydrodynamic solutions reached a stable solution, indicating little change in the River 2D spatial algorithm solving for depth and velocity between solution iterations. The average end solution change was 0.18%, with minimum and maximum solution changes of 0.02% and 0.38%, respectively. The mean difference between the measured inflow and modelled outflows across the nine levels of discharge was 0.07%, with minimum and maximum differences of 0.03% and 0.10%, respectively. Within wetted regions, the average area with a Froude number $\geq 1$, or supercritical flow, was $11.77\,\mathrm{m}^2$ with minimum and maximum estimates of 0.00 and $26.00\,\mathrm{m}^2$, respectively, or 0.00% and 0.04% of the wetted area. The model projections for velocity, depth and inundated area were qualitatively consistent with field observations and intuitive predictions, where all three variables increased positively with discharge. We could not obtain several modelled discharges to converge on a more
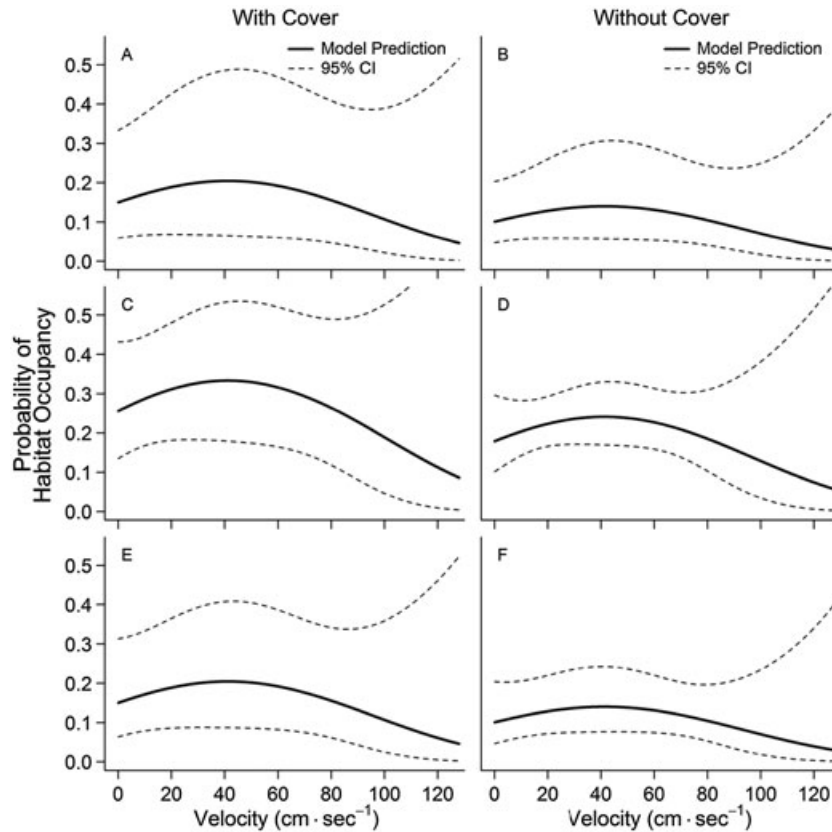
Figure 3. Multivariate logistic regression (solid line) plotted as the predicted probability of habitat occupancy for velocity (*x* axis), at depths of (A and B) 2.5 cm, (C and D) 32.5 cm and (E and F) 62.5 cm, (A, C and E) with cover and (B, D and F) without cover. Substrate size was held constant at 7.62 cm for (A–F) all model iterations. The fine dashed lines indicate the 95% confidence intervals (CI) of model predictions, estimated with the predict function in program R.

stable solution and below our threshold of 0.001 even after following the prescribed QA/QC procedures in Steffler and Blackburn (2002). One potential, and likely, explanation is that one or more shallow nodes were oscillating between wet and dry status through different iterations of the hydrodynamic solution, thereby increasing the final solution change (Steffler and Blackburn, 2002). We carefully examined the entire study reach for all modelled flows and could not find any sign of hydraulic anomalies (i.e. super critical flow and erroneous flow direction). The difference between measured inflow and modelled outflows was far below our threshold (<1%), and at least 99.96% of the modelled space maintained physically natural velocities (Froude <1.0). As such, we are confident that each discharge modelled was accurately estimating the depth and velocity across the study reach.

Modelled discharge altered total wetted area, where the total wetted area for flows of 33.98, 111.63 and 169.90 $m^3 s^{-1}$ was 63 501.5, 78 437.3 and 90 965.9 $m^2$, respectively. Increasing flow from 33.98 to 111.63 $m^3 s^{-1}$ resulted in a net increase of wetted area by 14 935 $m^2$ or 19.0%. Increasing flow from 111.63 to 169.90 $m^3 s^{-1}$ resulted in a net increase of wetted

area of 12 528.6 $m^2$ or 13.8%. At higher flows, the wetted surface area increased by roughly 10.5% per additional cubic metre of water compared with the lower flows.

*Comparing model predictions and calculating usable habitat*

Over the three quantitative comparisons between competing models, we found clear systematic differences between model predictions. Model A generally estimated a higher probability of habitat occupancy for a point in space; however, the relationship between point-specific model predictions was non-linear and at the highest probabilities of habitat occupancy model G predicted higher probability (Figure 4). The greatest differences in model predictions occurred at intermediate and higher probabilities of habitat occupancy (Figure 4). In general, there was very little agreement between point-specific model predictions, and agreement only occurred where the probability of habitat occupancy was approximately 18%.

To examine model comparisons in space, we calculated the difference between the point-specific predictions of each
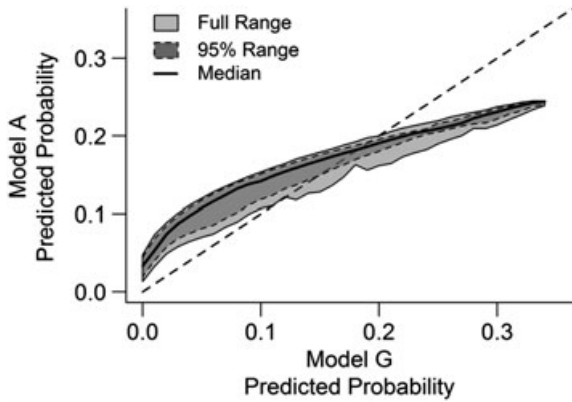
Figure 4. Point-specific estimates for model G (*x* axis) and model A (*y* axis) for all nine simulated flows. Model predictions are equal where they intersect with the dashed line.

model for three of the nine flows we modelled (33.98, 111.63 and 169.90 m$^3$ s$^{-1}$). Disparities in model predictions covered the greatest area at the lowest flows we modelled (Figure 5) and decreased as flows increased; however, discrepancies in model predictions remained high in areas that were most likely to be occupied by juvenile Chinook (Figure 5). Specifically, model disagreement remained high in the stream margins and over the floodplain.

We calculated total usable habitat across all nine flows simulated for models G and A (Figure 6). The relationship between usable habitat and discharge was non-linear differed between models. On average, model G predicted 17.25% less habitat than model A across the flows we simulated. Importantly, there was a high degree of uncertainty around the estimates of usable habitat, and model G had higher uncertainty than model A. At low flows, model G predicted 20.6% less usable habitat than model A; however, this difference decreased to 14.7% at higher flows. In addition, each model predicted a different change in habitat resulting from a change in flow indicating that extrapolation of either model over time would produce dissimilar habitat-flow, associations (Figure 6).

## DISCUSSION

The results of our diagnostic statistics indicate that model G correctly predicted a greater proportion of the outcomes in our habitat occupancy data compared with model A, commonly used in habitat modelling. However, the competing models were somewhat comparable in overall predictive accuracy and maintained slight to fair predictive capacities based on the Kappa statistic and AUC value. As well, the total usable habitat estimated by each model was similar given the degree of uncertainty around the model predictions. Despite these similarities, there were several important differences.
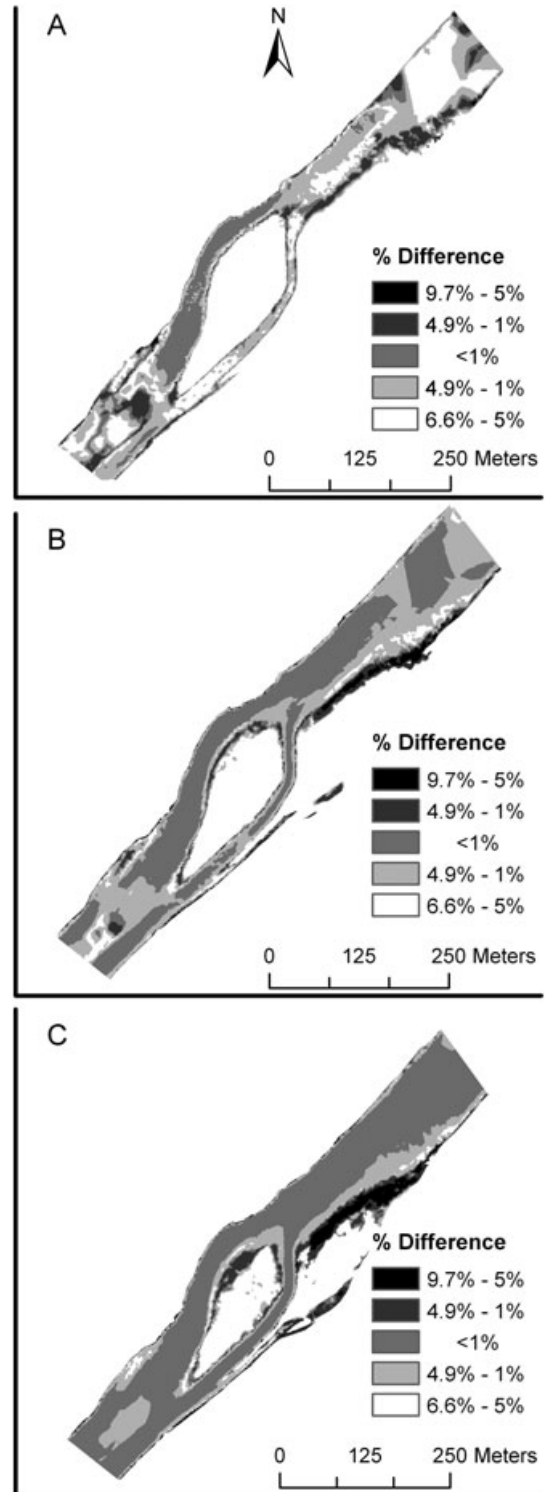


Figure 5. Estimated percent difference in model predictions plotted for the study reach at three simulated flows: (A) 33.98, (B) 111.63 and (C) 169.90 m$^3$ s$^{-1}$. Dark grey shades indicate space where model G estimated a higher probability of habitat occupancy than model A. Light grey shades indicate space where model A estimated a higher probability of habitat occupancy than model G.
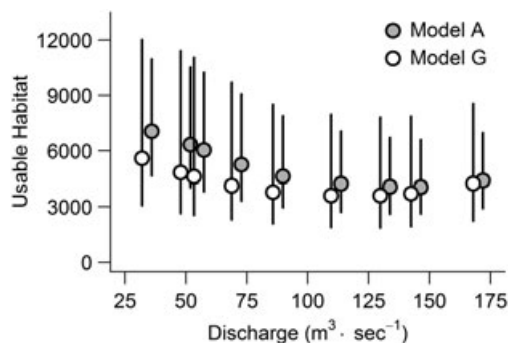
Figure 6. Estimated usable habitat based on predictions from competing hydrodynamic habitat models, A (grey circles) versus G (white circles). Error bars represent the minimum and maximum predicted usable habitat based on the upper and lower confidence limits of each respective model. X axis is slightly jittered ($\pm 2\,\mathrm{m}^3\,\mathrm{s}^{-1}$) to avoid overlap.

First, the models differed in the proportion of false positives and false negatives. Model A had a higher proportion of correctly identified occupied locations but at the cost of an increased proportion of falsely identified unoccupied locations. As a result, model A had a greater tendency to overestimate habitat occupancy, whereas model G tended to underestimate occupancy. Second, the difference between model predictions was greatest at the lowest flows we modelled and in the physical locations most likely to be occupied by juvenile Chinook salmon. Third, the amount of uncertainty around model predictions was much greater for model G than that for model A.

The relative consequence of using the aggregated (i.e. CSI method; model A) versus the global model (i.e. GLM; model G) in a management scenario is largely dependent on the differences between the model predictions. For example, each model had a clear bias between under-predicting and over-predicting habitat occupancy. If the management goal is to minimize habitat lost by flow regulation, then we posit it is more important to design a habitat model that has the lowest false positive rate. In other words, the best predictive model is one that accurately identifies habitat that will not be occupied by the target species. In this modelling exercise, our results support the conclusion that model G more accurately identified unoccupied locations. In this context, resource managers would have a better estimate of flow regulation effects using model G. In addition, the greatest differences between model predictions occurred in the stream margins and floodplain, which are known to be important rearing habitats for juvenile salmon (Sommer et al., 2001; Beechie et al., 2005). As such, our results indicate that model G had a higher overall predictive ability and therefore provides resource managers with a better tool for predicting how flow regulation will impact those important habitats. In total, the overall accuracy between the models was similar; however, the global

model (G) is better suited for management applications and predicting the impact of flow regulation on rearing habitat for juvenile salmon.

The disparity in the amount of uncertainty around model predictions is one of the most important differences between the competing models. As suggested in previous research, by aggregating univariate models, the modeller must assume each parameter is independent, and thus, the error is also independent (Mathur et al., 1985; Beecher et al., 2002; Jowett, 2003; Leclerc et al., 2003). As such, taking the geometric mean of multiple univariate model predictions does not correctly propagate the uncertainty between model parameters in the aggregated prediction. In this study, we posit that model A considerably underestimated the error around model predictions while simultaneously overestimating habitat occupancy. It is quite possible that this flaw would magnify when additional parameters are added and model complexity increases. For habitat models to be useful management tools to guide flow releases, it is imperative that they accurately reflect the uncertainty in parameter estimation. In doing so, resource managers would be better equipped to identify potential consequences arising from flow regulation within the full range of uncertainty in model predictions.

Some of the uncertainty in each model was due to low salmon densities and the high frequency of unoccupied habitat locations in our data. It is possible that the high proportion of unoccupied locations was a result of low salmon returns in the LAR during the period of time this study was conducted (Carlson and Satterthwaite, 2011). The lower the salmon population, the more difficult it is to parameterize the types of models we used in this study. As numerous salmon populations are currently depressed (Nehlsen et al., 1991; Gustafson et al., 2007), this is a concern and possible problem for future modelling in other watersheds. Our results lend further support for a more conservative statistical approach that more accurately incorporates uncertainty, such as AICc model averaging.

There were several candidate multivariate models with similar AICc scores in this study, but generally, multivariate models far outperformed the univariate models in predicting our observed patterns of occupancy by juvenile salmon. These results indicate that habitat occupancy is controlled by a combination of variables. Here, our study is limited to several key physical habitat variables that have repeatedly been reported to affect densities of aquatic species, including Chinook salmon, specifically water depth (e.g. Geist et al., 2000; Guay et al., 2000; Kynard et al., 2000; Beecher et al., 2002), velocity (e.g. Peeters and Gardeniers, 1998; Geist et al., 2000; Kynard et al., 2000; Mallet et al., 2000), substrate size (e.g. Knapp and Preisler, 1999; Vadas and Orth, 2001) and physical cover (e.g. Vadas and Orth, 2001; Gard 2006). However, we recognize there are a number of additional habitat characteristics that are hypothesized to be important

for functional salmonid rearing habitat, such as prey availability, stream temperature and dissolved oxygen (Hill and Grossman, 1993; Braaten *et al.*, 1997; Malcolm *et al.*, 2003; Hayes *et al.*, 2007). Furthermore, it is likely that interactions between parameters (e.g. temperature and velocity) will strongly influence the probability of habitat occupancy. Within the AICc framework we have presented, additional parameters and interactions can be easily integrated into predictive models and their relative importance balanced by model averaging.

We propose that the use of AICc and model averaging may be a valuable tool in producing objective and robust models of physical habitat for juvenile Chinook salmon or other species. Our results support the use of multivariate models over a CSI-based method for predicting the effects of habitat variables on the probability of habitat occupancy. We encourage the use of Cohen's kappa and other diagnostic statistics in model comparison and fully considering the context and consequences of model selection. Future research and management practices should carefully examine the role of multivariate dynamics and consider the implications of excluding them. Fresh water is, and will increasingly be, a limited resource throughout many areas of the world (Poff *et al.*, 2003; Richter *et al.*, 2003). Wise management of aquatic resources within highly regulated river systems is required to support ecosystem services and maintain quality of life for the people relying on such services (Arthington *et al.*, 2006). Therefore, the decision on how to best manage flow releases for downstream organisms is critical and needs to be informed by the best available science. This study illustrates the importance of analytic approaches in developing habitat models that support effective management of flow in regulated river systems.

## REFERENCES

Ahamadi-Nedushan B, St-Hilaire A, Berube M, Robichaud E, Thiemonge N, Bobee B. 2006. A review of statistical methods for the evaluation of aquatic habitat suitability for instream flow assessment. *River Research and Applications* **22**: 503–523.

Akaike H. 1974. A new look at the statistical model identification. *IEEE Transaction on Automatic Control* **19**: 716–723.

Ayllon D, Almodovar A, Nicola GG, Elvira B. 2010. Modelling brown trout spatial requirements through physical habitat simulations. *River Research and Applications* **26**: 1090–1102.

Arthington AH, Bunn SE, Poff NL, Naiman RJ. 2006. The challenge of providing environmental flow rules to sustain river ecosystems. *Ecological Applications* **16**: 1311–1318.

Beecher HA, Caldwell BA, DeMond SB. 2002. Evaluation of depth and velocity preferences of juvenile coho salmon in Washington streams. *North American Journal of Fisheries Management* **22**: 785–795.

Beechie TJ, Liermann M, Beamer EM, Henderson R. 2005. A classification of habitat types in a large river and their use by juvenile salmonids. *Transactions of the American Fisheries Society* **134**: 717–729.

Boavida I, Santos JM, Cortes RV, Pinheiro AN, Ferreira MT. 2011. Assessment of instream structures for habitat improvement for two critically endangered fish species. *Aquatic Ecology* **45**: 113–124.

Bovee KD. 1986. Development and evaluation of habitat suitability criteria for use in the instream flow incremental methodology. *U.S. Fish and Wildlife Service Biological Report* **86**: 1–235.

Braaten PJ, Dey PD, Annear TC. 1997. Development and evaluation of bioenergetic-based habitat suitability criteria for trout. *Regulated Rivers: Research & Management* **13**: 345–356.

Burgman MA, Breininger DR, Duncan BW, Ferson S. 2001. Setting reliability bounds on habitat suitability indices. *Ecological Applications* **11**: 70–78.

Burnham KP, Anderson DR. 2002. *Model Selection and Multimodel Inference: A Practical-Theoretic Approach*, 2nd ed. Springer-Verlag: New York.

Carlson SM, Satterthwaite WH. 2011. Weakened portfolio effect in a collapsed salmon population complex. *Canadian Journal of Fisheries and Aquatic Sciences* **68**: 1579–1589.

Dunbar MJ, Alfredsen K, Harby A. 2011. Hydraulic-habitat modelling for setting environmental river flow needs for salmonids. *Fisheries Ecology and Management*. http://dx.doi.org/10.1111/j.1365-2400.2011.00825.x

Freeman EA, Moisen G. 2008. PresenceAbsence: an R package for presence absence analysis. *Journal of Statistical Software* **23**: 1–31.

Gard M. 2006. Modeling changes in salmon spawning and rearing habitat associated with river channel restoration. *International Journal of River Basin Management* **4**: 201–211.

Geist DR, Jones J, Murray CJ, Dauble DD. 2000. Suitability criteria analyzed at the spatial scale of redd clusters improved estimates of fall Chinook salmon (*Oncorhynchus tshawytscha*) spawning habitat use in the Hanford Reach, Columbia River. *Canadian Journal of Fisheries and Aquatic Sciences* **57**: 1636–1646.

Grant GE. 1997. Critical flow constrains flow hydraulics in mobile-bed streams: a new hypothesis. *Water Resources Research* **33**: 349–358.

Grueber CE, Nakagawa S, Laws RJ, Jamieson IG. 2011. Multimodel inference in ecology and evolution: challenges and solutions. *Journal of Evolutionary Biology* **24**: 699–711.

Guay JC, Boisclair D, Rioux D, Leclerc M, Lapointe M, Legendre P. 2000. Development and validation of numerical habitat models for juveniles of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences* **57**: 2065–2075.

Guisan A, Edwards TC, Hastie T. 2002. Generalized linear and generalized additive models in studies of species distributions: setting the scene. *Ecological Modelling* **157**: 89–100.

Gustafson RG, Waples RS, Myers JM, Weitkamp LA, Bryant GJ, Johnson OW, Hard JJ. 2007. Pacific salmon extinctions: quantifying lost and remaining diversity. *Conservation Biology* **21**: 1009–1020.

Hayes JW, Hughes NF, Kelly LH. 2007. Process-based modeling of invertebrate drift transport, net energy intake and reach carrying capacity for drift-feeding salmonids. *Ecological Modelling* **207**: 171–188.

Hill J, Grossman GD. 1993. An energetic model of microhabitat use for rainbow trout and rosyside dace. *Ecology* **74**: 685–698.

Hurvich CM, Tsai CL. 1989. Regression and time series model selection in small samples. *Biometrika* **79**: 297–307.

Im D, Kang H, Kim K, Choi S. 2011. Changes of river morphology and physical fish habitat following weir removal. *Ecological Engineering* **37**: 883–892.

Jowett IG. 2003. Hydraulic constraints on habitat suitability for benthic invertebrates in gravel-bed river. *River Research and Applications* **19**: 495–507.

Knapp RA, Preisler HK. 1999. Is it possible to predict habitat use by spawning salmonids? A test using California golden trout (*Oncorhynchus mykiss aguabonita*). *Canadian Journal of Fish and Aquatic Sciences* **56**: 1576–1584.

Kynard B, Horgan M, Kieffer M. 2000. Habitats used by shortnose sturgeon in two Massachusetts rivers, with notes on estuarine Atlantic sturgeon: a hierarchical approach. *Transactions of the American Fisheries Society* **129**: 487–503.

Labonne J, Allouche S, Gaudin P. 2003. Use of a generalised linear model to test habitat preferences: the example of *Zingel asper*, an endemic endangered percid of the River Rhone. *Freshwater Biology* **48**: 687–697.

Leclerc M, St-Hilaire A, Bechara JA. 2003. State-of-the-art and perspectives of habitat modeling for determining conservation flows. *Canadian Water Resources Journal* **28**: 135–151.

Lee JH, Kil JT, Jeong S. 2010. Evaluation of physical fish habitat quality enhancement designs in urban streams using a 2D hydrodynamic model. *Ecological Engineering* **36**: 1251–1259.

Malcolm IA, Youngson AF, Soulsby C. 2003. Survival of salmonid eggs in a degraded gravel-bed stream: effects of groundwater-surface water inter-actions. *River Research and Applications* **19**: 303–316.

Mallet JP, Lamouroux P, Sagnes P, Persat H. 2000. Habitat preferences of European grayling in a medium size stream, the Ain river, France. *Journal of Fish Biology* **56**: 1312–1326.

Manel S, Williams HC, Ormerod SJ. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* **38**: 921–931.

Mathur D, Bason WH, Purdy Jr ED, Silver CA. 1985. A critique of the instream flow incremental methodology. *Canadian Journal of Fisheries and Aquatic Sciences* **42**: 825–831.

McEwan DR. 2001. Central Valley steelhead. *California Department of Fish and Game, Fish Bulletin* **179**: 1–43.

Mouton AM, De Baets B, Goethals PLM. 2010. Ecological relevance of performance criteria for species distribution models. *Ecological Modelling* **221**: 1995–2002.

Moyle PB. 2002. *Inland Fishes of California, Revised and Expanded*. University of California Press: Berkeley, California, USA.

Myers JM, Kope RG, Bryant GJ, Teel D, Lierheimer LJ, Wainwright TC, Grant WS, Waknitz FW, Neely K, Lindley ST, Waples RS. 1998. Status review of Chinook salmon from Washington, Idaho, Oregon, and California. *United States Department of Commerce. National Oceanic Atmospheric Administration Technical. Memo* NMFS-NWFSC-35.

Nehlsen W, Williams JE, Lichatowich JA. 1991. Pacific salmon at the crossroads: stocks at risk from California, Oregon, Idaho, and Washington. *Fisheries* **16**: 121–145.

Nilsson C, Reidy CA, Dynesius M, Revenga C. 2005. Fragmentation and flow regulation of the world's large river systems. *Science* **308**: 405–408.

Peeters ETHM, Gardeniers JJP. 1998. Logistic regression as a tool for defining habitat requirements of two common gammarids. *Freshwater Biology* **39**: 605–615.

Petts GE. 2008. Instream flow science for sustainable river management. *Journal of the American Water Resources Association* **45**:1071–1086.

Poff NL, Allan JD, Palmer MA, Hart DD, Richter BD, Arthington AH, Rogers KH, Meyer JL, Stanford JA. 2003. River flows and water wars: emerging science for environmental decision-making. *Frontiers in Ecology and the Environment* **1**, 298–306.

R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Richter BD, Mathews R, Harrison DL, Wigington R. 2003. Ecologically sustainable water management: managing river flows for ecological integrity. *Ecological Applications* **13**: 206–224.

Sommer TR, Nobriga ML, Harrell WC, Batham W, Kimmerer WJ. 2001. Floodplain rearing of juvenile Chinook salmon: evidence of enhanced growth and survival. *Canadian Journal of Fisheries and Aquatic Sciences* **58**: 325–333.

Stalnaker CB, Lamb BL, Henriksen J, Bovee K, Bartholow J. 1995. The instream flow incremental methodology: a primer for IFIM. Washington, DC: U.S. Geological Survey Biological Report 29. 45 p.

Steffler P, Blackburn J. 2002. River 2D: two-dimensional depth averaged model of river hydrodynamics and fish habitat. *Introduction to depth averaged modeling and user's manual*. University of Alberta, Edmonton, Alberta. http://www.river2d.ualberta.ca/download.htm

Thompson SM, Campbell PL. 1979. Hydraulics of a large channel paved with boulders. *Journal of Hydraulic Research* **17**: 341–354.

United States Fish and Wildlife Service. 1981. Standards for the develop-ment of habitat suitability index models. 103 ESM. Division of Ecological Services, Department of the Interior, Washington, DC.

Vadas RL, Orth DJ. 2001. Habitat suitability models for stream fish guilds: do the standard methods work? *Transactions of the American Fisheries Society* **130**: 217–235.

Vismara R, Azzellina A, Bosi R, Crosa G, Gentili G. 2001. Habitat suitability curves for brown trout (*Salmo trutta fario* L.) in the river Adda, Northern Italy: comparing univariate and multivariate approaches. *Regulated Rivers: Research & Management* **17**: 37–50.

Waddle TJ. 2001. PHABSIM for Windows: User's Manual and Exercises: Fort Collins, CO, U.S. Geological Survey, 288 p.

Williams JG. 2001. Chinook salmon in the lower American River, California's largest urban stream. In *Contributions to the Biology of Central Valley Salmonids*, Brown RL (ed). California Department of Fish and Game, Fish Bulletin No. 179.2: Sacramento; 1–38.

Yoshiyama RM, Gerstung ER, Fisher FW, Moyle PB. 2001. Historical and present distribution of Chinook salmon in the Central Valley. In *Contribu-tions to the Biology of Central Valley Salmonids*, Brown R (ed). California Department of Fish and Game Fish Bulletin 179: Sacramento, CA; 71–176.