Quantile Regression Forest Models to Estimate Habitat Capacity for Spring-Summer Chinook and Steelhead

Project 2003-017-00

Report covers work performed under BPA contracts #74633 and #77840
Report was completed under BPA contract #77840

December 2017 – November 2018

Kevin See, Michael Ackerman, Richie Carmichael, Braden Lott, Chris Beasley
Biomark – Applied Biological Services
Boise, Idaho

Report Created: November 2018

This report was funded by the Bonneville Power Administration (BPA), U.S. Department of Energy, as part of BPA's program to protect, mitigate, and enhance fish and wildlife affected by the development and operation of hydroelectric facilities on the Columbia River and its tributaries. The views in this report are the author's and do not necessarily represent the views of BPA.

LIST OF TABLES

LIST OF FIGURES

SUPPLEMENTAL TABLES

SUPPLEMENTAL FIGURES

ABSTRACT

The decline of anadromous Pacific salmonids (*Oncorhynchus* spp.) across the Pacific Northwest, USA has prompted numerous actions aimed at reversing that trend. These actions are often categorized into the four ‘H’s’ – harvest modification, habitat rehabilitation, hydroelectric operations, and hatchery practices. Problematically, substantial uncertainty remains regarding the degree of change to salmon populations that can be exerted across and within these categories, and what combination of changes might most cost-effectively and sustainably reduce mortality and recover depleted populations. Recently released ESA delisting criteria have identified adult escapement targets at the population scale and have provided a quantitative metric useful for evaluating the magnitude of survival improvements (across life stages) required. In this report, we focus on one of the ‘H’s’ – habitat rehabilitation, and these abundance targets provide a benchmark against which habitat rehabilitation actions can be measured. We describe development of and improvements to a novel approach to estimate the capacity of habitat to support Chinook salmon and steelhead juveniles (during summer and winter months) and redds in wadable streams in the interior Columbia River Basin. Our model leverages quantile regression and random forest methods and uses paired measurements of fish abundance/density and habitat characteristics from sites across the Columbia River Basin. The models provide estimates of available habitat capacity for a given life stage (e.g., summer parr, winter presmolt, redds), for two different species (Chinook salmon and steelhead), and at any given scale (e.g., watershed, population). Available capacity estimates can then be compared to estimates of life stage specific abundance necessary to achieve delisting or recovery goals. Doing so provides a quantitative means to elucidate the relative amount of habitat rehabilitation necessary to provide sufficient habitat (quantity and quality) for recovery. Our approach is entirely empirical, allowing fish-habitat relationships to emerge from the data, even if they are non-linear in nature (as many ecological relationships are). Here, we provide life stage specific capacity estimates and maps for the Lemhi and Upper Grande Ronde rivers that provide an example of outputs from our model. We have validated the QRF estimates of Chinook summer parr with spawner-recruit curves from a variety of watersheds, and found them to match up very well, despite being based on entirely different data. Additionally, QRF predictions of capacity can be built on habitat sampling conducted over a handful of years (or a single year with enough effort), whereas spawner-recruit curves, while often considered a gold standard for estimating capacity, require many years of data with plenty of contrast to be considered valid. Carrying capacity models based on QRF and habitat data, like those presented here, provide managers a framework to guide the identification, prioritization, and development of habitat rehabilitation actions to recover salmon populations. For perhaps the first time, the necessity of tributary habitat restoration actions can be demonstrated, and the magnitude of required change can be placed in context with the other ‘H’s’.

This report should be cited as follows:

See, K., M.W. Ackerman, R. Carmichael, B. Lott, and C. Beasley. 2018. Quantile Regression Forest Models to Estimate Habitat Capacity for Spring-Summer Chinook and Steelhead. December 2017 – November 2018, BPA Project 2003-017-00, pp:1-37

INTRODUCTION

The decline of anadromous Pacific salmonids (*Oncorhynchus* spp.) across the Pacific Northwest, USA has prompted numerous actions aimed at reversing that trend. These actions are often categorized into four 'H's' – harvest modification, habitat rehabilitation, hydroelectric operations, and hatchery practices. Problematically, substantial uncertainty remains regarding the degree of change to salmon populations that can be exerted across and within these categories, and what combination of changes might most cost-effectively and sustainably reduce mortality and recover depleted populations. Recently released 'delisting' criteria (NOAA 2016) identified adult escapement targets at the population scale, providing a quantitative metric useful for evaluating the magnitude of survival improvements (across life-stages) required. These abundance targets provide a benchmark against which habitat rehabilitation actions can be measured. In this report, we describe development of and improvements to a novel approach for estimating life-stage specific habitat capacity that can be used to quantitatively identify the magnitude of tributary habitat restoration to support Endangered Species Act (ESA) delisting. For perhaps the first time, the necessity of tributary habitat restoration actions can be demonstrated, and the magnitude of required change can be placed in context with the other 'H's'.

Pacific salmon species have experienced large declines in abundance throughout much of their range (Good et al. 2005). Declines can be partially attributed to lost or altered habitat, and thus, efforts to recover depleted salmon populations are replete with efforts to rehabilitate habitat used during freshwater life-stages. Specifically, restoring salmonid carrying capacity through tributary rehabilitation actions has been identified as a key component of recovery efforts for salmon and steelhead (*Oncorhynchus mykiss*) in the Pacific Northwest, USA. Reductions in population productivity have been partially attributed to decreases in the quantity and quality of tributary habitat, and therefore, efforts have included both increasing and improving available habitat for spawning adults and rearing juveniles. The goal of habitat restoration is often to restore habitat quantity/quality to historic levels to increase the capacity of the habitat to support spawning adults and rearing juveniles.

Carrying capacity is the maximal abundance or load the habitat can support for a given species and life-stage given current resources. Unfortunately, estimating habitat carrying capacity, both historic and contemporary, for various life-stages of Pacific salmon, as well as identifying important habitat characteristics that influence capacity, has been an ongoing but necessary challenge. Reliable methods to better understand fish-habitat relationships as well as to estimate capacity are necessary to identify those salmon and steelhead life-stages that are limited by habitat capacity. Such methods will provide tools to better direct and prioritize habitat rehabilitation efforts.

Within fisheries research and management, it has long been recognized that biotic and abiotic factors limit productivity within and across life-stages. For this report, we define carrying capacity as the maximum number of individuals that can be supported given the quantity and quality of habitat available at a given life-stage. We assume that higher observed relative fish densities within a given life stage are a function of habitat quantity and quality. Furthermore, we assume that observed fish density is a poor predictor of habitat capacity owing to both a paucity of individuals and the existence of unmeasured biotic or abiotic variables that may serve to limit capacity. To address this, we have developed a model to estimate the carrying capacity of wadeable streams to support spawning and rearing spring-summer run Chinook salmon (*O. tshawytscha*; hereafter Chinook salmon) and summer run steelhead (*O. mykiss*; hereafter steelhead) using quantile regression forest models (QRF; Meinshausen 2006). The model uses paired measurements of fish abundance/density and habitat characteristics. QRF models combine the flexibility of random forest models (Breiman 2001) with the ability of quantile regression to extract relationships from any quantile of the data (Cade and Noon 2003), not just the mean (e.g., as in linear regression).

Quantile regression forests share many of the benefits of random forest models, such as the ability to capture non-linear relationships between the response (e.g., fish density) and the predictor (e.g., habitat metrics) variables, and naturally incorporate interactions between the predictor variables, two common features of ecological datasets (Liaw and Wiener 2002). Random forest models have also been shown to outperform more standard parametric models in predicting fish-habitat relationships in other contexts (Knudby et al. 2010) and can work with untransformed data while being robust to outliers (Breiman 2001; Prassad et al. 2006). Quantile regression models have been used in a variety of ecological systems to estimate the effect of limiting factors (Terrell et al. 1996; Cade and Noon 2003). A random forest model consists of a number (e.g., 1000) of classification and regression trees, which each provide a prediction of fish density provided a given set of habitat characteristics, resulting in a distribution of predictions. The mean of those predictions is analogous to predictions from a linear regression model, but other quantiles of that predictive distribution can also be examined, such as extreme upper quantiles as a proxy of habitat capacity.

In this report, we describe the development and implementation of QRF models to better elicit fish-habitat relationships and to predict habitat rearing capacity at the site scale for juvenile and adult Chinook salmon and steelhead using paired fish and habitat data. The juvenile models pertains to juveniles rearing in wadable streams during both summer (parr) and winter (presmolt) months; the adult model is to elicit fish-habitat relationships for spawning areas and predict habitat capacity to support redds. Fish and habitat data used here to populate the QRF models are available from eight watersheds within the interior Columbia River Basin. The habitat data are from the Columbia Habitat Monitoring Program (CHaMP; https://www.champmonitoring.org). Fish and habitat data were paired at CHaMP sites (200 – 500 m) where fish survey data were available. Importantly, the QRF model places no constraints on possible fish-habitat relationships; instead, relationships are estimated from the data regardless of being positive, negative, linear, non-linear, etc. Based on the observed fish-habitat relationships, we can then predict capacity at all CHaMP sites using measurements of the same habitat covariates that were used to populate the model. In summary, our objectives include:

1. Identify measured habitat characteristics that are most strongly associated with observed Chinook salmon and steelhead parr and redd abundance/density.
2. Use paired fish and habitat measurements to elicit fish-habitat relationships for those habitat characteristics identified as important for determining parr or redd abundance/density.
3. Predict contemporary habitat carrying capacity at all sites where important habitat characteristics are measured (i.e., CHaMP sites within the Columbia River Basin).
4. Extrapolate capacity predictions at CHaMP sites across larger scales (e.g., watershed, population) using globally available attribute data to estimate Chinook salmon and steelhead parr and redd capacity at those larger scales. Here, we provide extrapolated capacity predictions across the Lemhi River, ID and upper Grande Ronde River, OR.
5. Where possible, validate estimates of carrying capacity from QRF across multiple watershed using independent estimates of capacity (e.g., from spawner-recruit relationships).

For summer parr capacity models, we predict capacity at the reach (200 to 500 m) scale. For overwintering presmolt capacity, we modeled capacity at the channel unit scale, but then combined channel units up to the reach scale. For the redd model, we predict capacity at a slightly larger reach (1 river kilometer [rkm]) scale. In doing so, we elicit data-driven fish-habitat relationships from the data. Moreover, we describe a method to extrapolate capacity estimates to larger spatial scales (e.g., basin, population).

Estimates of available habitat capacity for a given life stage (e.g., summer parr, winter presmolt, redds) at any given scale (e.g., watershed, population), can then be compared to estimates of life stage specific abundance necessary to achieve delisting or recovery goals. Doing so provides a quantitative means to elucidate the relative amount of habitat rehabilitation needed to provide sufficient habitat (quantity and quality) for recovery. Carrying capacity models based on QRF and habitat data, like those presented here, provide managers a framework to guide the identification, prioritization, and development of habitat rehabilitation actions to recover salmon populations.

METHODS

Study Site

Fish and habitat data used in the QRF models were collected from eleven watersheds within the interior Columbia River Basin. The eleven watersheds include: Asotin, Entiat, Grande Ronde (upper), John Day, Lemhi, Methow, Minam (tributary of Grande Ronde), Secesh, Tucannon, Wenatchee and Yankee Fork. Juvenile fish and redd data collected at CHaMP survey sites were graciously provided by a number of collaborators and projects and included the Integrated Status and Effectiveness Monitoring Program (ISEMP).

Data

Habitat Data

The habitat data were collected by CHaMP (ISEMP/CHaMP 2017) and were downloaded from https://www.champmonitoring.org. CHaMP sites are 200 – 500 m reaches within wadeable streams across select watersheds within the interior Columbia River Basin and were selected based on a spatially balanced Generalized Random Tesselation Stratified (GRTS) sample selection algorithm (Stevens and Olsen 1999, 2004). Habitat data within CHaMP sites are collected using the CHaMP protocol (CHaMP 2016) which calls for field data collection during the low-flow period, typically from June through October. CHaMP habitat data include, but are not limited to, measurements describing channel complexity, channel units, disturbance, fish cover, large woody debris, riparian cover, size (depth, width, discharge), substrate, temperature, and water quality.

Temperature data collected using instream temperature loggers were only available for a small portion of CHaMP survey sites over appropriate time intervals. Therefore, modeled temperature data (McNyset et al. 2015) was provided by South Fork Research, Inc. Modeled temperature data summarizing the mean of 8-day means and the maximum of 8-day means for CHaMP sites during summer months (August and September) were available for the years 2011 – 2014.

Juvenile Fish Data

Juvenile fish surveys were conducted for Chinook salmon and juvenile steelhead during the summer and winter low-flow seasons at many of the same sites that were surveyed for habitat using the CHaMP protocol. Juvenile fish data included in this analysis was collected during summers of 2011-2014, and in the winter of 2017-2018. Fish survey methods to estimate juvenile abundance included mark-recapture, three-pass removal, two-pass removal, and single-pass electrofishing, as well as snorkeling. Survey data were used to estimate juvenile abundance at all sites where data were available. Mark-recapture estimates used Chapman's modified Lincoln-Peterson estimator (Chapman 1951) and were deemed valid if they met the criteria described in Robson and Regier (1964). Three-pass removal estimates used the Carle-Strub estimator (Carle and Strub 1978), following advice from Hedger et al. (2013). Two-pass removal estimates used the estimator described by Seber (2002). These abundance estimates were made using the *removal* function from the *FSA* package (Ogle 2017) or the *closed.bc* function from the *Rcapture* package (Rivest and Baillargeon 2014) in R software (R Core Team 2015). Snorkel counts were transformed to abundance estimates using a correction factor derived from sites where paired snorkel-electrofishing data were available. For sites with invalid estimates (e.g., zero recaptures) or that were sampled with a single electrofishing pass, a ratio estimator was developed to estimate the probability of capture. The ratio estimator was based on a binomial generalized linear model that was developed separately for each basin and each sampling agency. Possible covariates included the number of fish captured on the first pass, year, site length, and Julian day. After fitting all possible models with those covariates to data with valid abundance estimates, the model with the lowest AICc for each basin and sampling agency was chosen and used to predict abundances based on the number of fish caught on the first pass and any other covariates. Summer sampling and data collection was conducted at the site scale, whereas winter sampling was conducted at the channel unit scale, primarily using snorkel surveys.

Juvenile abundance estimates at all sites were translated into linear fish densities (parr/m) for the summer, and areal densities (parr/m^2) in the winter, and density estimates were paired with the associated CHaMP habitat data. For sites that were sampled in multiple years, only the fish and habitat data from the year with the highest observed fish density was retained to avoid possible pseudo-replication.

Redd Data

Chinook salmon and steelhead redd data were graciously provided by the Idaho Department of Fish and Game, Nez Perce Tribe Department of Fisheries Resources, Oregon Department of Fish and Wildlife, U.S. Fish and Wildlife Service, and the Washington Department of Fish and Wildlife and span the years 1995 to 2016. Redd data were available for the following CHaMP watersheds: Asotin, Entiat, John Day, Lemhi, Methow, Minam, Secesh, Tucannon, upper Grande Ronde, Wenatchee and Yankee Fork.

To pair the redd and habitat data, the number of redds that occurred within a 1 river kilometer (rkm) buffer of the central point (i.e., x-site) for each CHaMP site were tallied. The latitude and longitude of each CHaMP site and each redd were snapped to a route in ArcGIS and the number of redds that occurred within 500 m upstream or downstream of each CHaMP site for each year in which redds were observed were counted and transformed into linear densities (redds/m). For each CHaMP site, we identified the year in which the maximum number of redds were observed because we are ultimately interested in redd capacity, and therefore used the highest observed redd density at each CHaMP site.

Habitat Covariate Selection

A crucial step in developing a QRF model to predict habitat capacity to support juveniles or redds is selecting the habitat covariates to include in the model. Random forest models naturally incorporate interactions between correlated covariates, which is essential since nearly all CHaMP habitat variables are correlated to one degree or another. However, we aimed to avoid including redundant variables (i.e., variables that measure similar aspects of the habitat). Finally, including too many habitat covariates can result in overfitting of the model (e.g., including as many covariates as data points).

The CHaMP protocol produces more than 100 metrics describing the quantity and quality of fish habitat for each survey site, as well as number of metrics at the channel unit scale. To decide which habitat metrics to use in each QRF model, we considered first the association between the habitat metric and observed parr and redd densities and second the correlation among habitat metrics. We used the Maximal Information-Based Nonparametric Exploration (MINE) class of statistics (Reshef et al. 2011) to determine those habitat characteristics most highly associated with observed juvenile or redd densities. MINE statistics were employed using the R package *minerva* (Albanese et al. 2013). Within the MINE statistics, we used the maximal information coefficient (MIC) to measure the strength of the linear or non-linear association between fish density and each habitat characteristic (Reshef et al. 2011). The MIC value was used to inform decisions on which habitat covariates to include in the model. Habitat metrics were grouped into broad categories that include channel unit, complexity, cover, disturbance, riparian, size, substrate, temperature, water quality, and woody debris. Within each category, metrics were ranked according to their MIC value. Our strategy was to select one or two variables with high MIC scores within each category so that covariates describe different aspects of the spawning and juvenile rearing habitat. Additionally, we measured pairwise correlations among all habitat metrics and attempted to avoid covariates that were highly correlated and include covariates that describe potentially meaningful fish-habitat relationships. Table 1 provides a summary of habitat covariates used in each of the QRF models.

QRF Model Fitting

In total, six QRF models were fit including combinations of two species (Chinook salmon and steelhead) and three life stages (summer parr, winter presmolt, redd). Each of the QRF models were fit using the selected habitat covariates and using the *quantregForest* function from the *quantregForest* package (Meinshausen 2016) in R software (R Core Team 2015). The individual predictions from each tree, viewed collectively, describe the entire distribution of the predicted response. Therefore, the random forest model can be used in the same way as other quantile regression methods to predict any quantile of the response. The 90th quantile of the predicted distribution was used as a proxy for habitat carrying capacity. We chose to use the 90th quantile, instead of something higher, to avoid using predictions that are aimed at the very upper tails of observed fish density, which may be influenced by sampling issues.

Chinook salmon and steelhead summer parr, winter presmolt, and redd densities and associated habitat data were paired by site and year; this habitat data contained some missing values. Within each dataset, any site visit with more than three missing covariates were dropped from the dataset; the remaining missing habitat values were imputed using the *missForest* R package (Stekhoven and Buehlmann 2012; Stekhoven 2013).

After model fitting, each QRF model can then be used to predict Chinook salmon or steelhead summer parr, winter presmolt, or redd capacity using measurements of the habitat covariates used to fit the model. In our case, this includes all sites in the Columbia River Basin surveyed using the CHaMP protocol (CHaMP 2016). For CHaMP sites surveyed in multiple years, we first calculated the mean among years prior to making predictions. The 90th quantile of the random forest predictions of juvenile or redd densities were then used as a proxy for habitat carrying capacity. For overwintering capacity, QRF predictions were made at the channel unit scale, and then combined to estimate capacity at the CHaMP site scale.

Model Extrapolation

The QRF model enables predictions of habitat capacity for Chinook salmon and steelhead juveniles (summer and winter rearing) and redds at all 200 – 500 m CHaMP sites within the interior Columbia River Basin. However, fisheries biologists are typically more often interested in capacity at larger scales (e.g., watershed or population). To accomplish this, we developed an extrapolation model that uses globally available attribute (GAA) data to scale CHaMP site predictions to a series of larger scales; often corresponding to the entirety of tributary habitat utilized by a given population. The GAA data used in the extrapolation model were available from a list of Generalized Random Tesselation Stratified (Stevens and Olsen 1999, 2004) master sample sites that CHaMP sites were initially selected from. The extrapolation model uses a multiple linear regression model that incorporates the design weights of CHaMP sites using the *svyglm* function from the *survey* package (Lumley 2004, 2016) in R software (R Core Team 2015).

We extrapolated both linear and areal capacity estimates from CHaMP sites to all master sample sites, using the natural log of the CHaMP site capacity predictions as the response variable, and only CHaMP sites within the domain of spring/summer Chinook salmon or summer steelhead were used to fit the models. The domain for any given watershed was either determined by StreamNet http://www.streamnet.org or using expert opinion from local biologists. For each response (linear or areal juvenile or redd density), two models were developed, one for sites within CHaMP watersheds (using watershed as a covariate), and one for all other sites (i.e., sites outside of CHaMP watersheds. In total, there were potentially twenty-four extrapolation models (3 life stages x 2 responses [linear or areal] x 2 species x 2 locations [CHaMP vs. non-CHaMP]). To summarize capacity at larger scales, the mean linear capacity (i.e., juveniles/m or redds/m) of the master sample points within a given spatial scale was determined and then multiplied by the total length of the stream within the domain to estimate total capacity.

Model Validation

Our goal is to validate each of the QRF and extrapolation models by providing a comparison of habitat capacity from the models to independent estimates of capacity (e.g., using spawner-recruit or historic redd data). However, to date we have only been able to validate models for Chinook salmon summer parr capacity. Time-series spawner-recruit data for steelhead and historic estimates of redd abundance for Chinook salmon and steelhead are sparse for watersheds and populations in the interior Columbia River Basin. As a result, we have only performed model validation for the Chinook salmon parr capacity model.

Estimates of total Chinook salmon parr capacity have been made for a number of watersheds or populations throughout the interior Columbia River Basin. Spawner-recruit data from several watersheds were compiled to validate the extrapolated QRF estimates of Chinook salmon parr capacity. Some watersheds had direct estimates of parr, while some had estimates of smolts from rotary screw traps (RST). For watersheds with only smolt estimates, estimates of parr were calculated using estimates of over-winter survival from those watersheds. A series of spawner-recruit functions were then fit to this data including Beverton-Holt, Ricker, and hockey stick. Each of these functions also provide an estimate of parr capacity; parr capacity estimates from each of the spawner-recruit functions were then compared with QRF estimates of Chinook salmon parr capacity (

Table 2, Figure 1).

RESULTS

Habitat Covariate Selection

We categorized 150 habitat measurements collected using the CHaMP habitat protocol (CHaMP 2016) into 11 habitat groups. For each model an MIC value was calculated for each habitat covariate based on the strength of association between the habitat covariate and the response variable (fish or redd density). Covariates were then ranked within each habitat group and we selected one or two covariates within each habitat group taking into consideration their MIC rank and number of missing values. Our strategy was to 1) consider pairwise correlations among habitat covariates to minimize redundant covariates measuring similar aspects of habitat, and 2) select covariates that describe habitat characteristics likely important towards spawning or rearing.

We focused on each life-stage in turn, examining the MIC statistics of each habitat covariates for both Chinook and steelhead. We chose to use the same suite of habitat covariates for both species within the same life-stage. We selected between eight and fourteen metrics to use in each life-stage. Table 1 shows the CHaMP habitat covariates used to fit each of the QRF models.

Table 1. Habitat metrics and descriptions of metrics included in each of the QRF capacity models. Numbers indicate where each metric ranked in relative importance for each model. Dashed indicate a metric was not used for a given model.

Metric Category	Metric	Description	Chinook			Steelhead		
			Sum. Juv.	Win. Juv.	Redd	Sum. Juv.	Win. Juv.	Redd
Channel Unit	Channel Unit Frequency	Number of channel units per 100 meters.	8	2	–	12	3	–
Channel Unit	Fast Turbulent Frequency	Number of Fast Water Turbulent channel units per 100 meters.	–	–	13	–	–	6
Channel Unit	Fast Turbulent Percent	Percent of wetted area identified as Fast Water Turbulent channel units.	–	–	11	–	–	8
Channel Unit	Tier1	Tier 1 channel unit type.	–	8	–	–	8	–
Complexity	Sinuosity	Ratio of the thalweg length to the straight line distance between the start and end points of the thalweg.	–	4	–	–	6	–
Complexity	Thalweg Depth CV	Coefficient of Variation (CV) of thalweg depths at a site.	9	–	–	7	–	–
Complexity	Wetted Width To Depth Ratio Avg	Average width to depth ratio of the wetted channel measured from cross-sections. Depths represent an average of depths along each cross-section.	4	–	12	5	–	2
Complexity	Wetted Width To Depth Ratio CV	Retired. Coefficient of Variation of wetted width to depth ratios derived from cross-sections.	–	–	9	–	–	14
Cover	Fish Cover: LW	Percent of wetted area that has woody debris as fish cover.	–	6	–	–	4	–
Cover	Fish Cover: Total	Percent of wetted area with the following types of cover: aquatic vegetation, artificial, woody debris, and terrestrial vegetation.	5	–	–	9	–	–
Land Classification	Disturbance Index	Disturbance index that includes measures of % urban, % agricultural, % impervious surface and road density (Whittier et al. 2011).	14	–	6	4	–	3
Land Classification	Natural PC 1	A natural index that describes watershed slope, precipitation, growing season (growing degree day), and low-gradient streams (Whittier et al. 2011).	–	–	3	–	–	7
Riparian	Riparian Cover: Ground	Percent of groundcover vegetation.	6	–	–	14	–	–
Riparian	Riparian Cover: No Canopy	Percent of riparian canopy devoid of vegetation.	–	–	8	–	–	13
Size	Discharge	The sum of station discharge across all stations. Station discharge is calculated as depth x velocity x station increment for all stations except first and last. Station discharge for first	–	–	7	–	–	1

Metric Category	Metric	Description	Chinook Sum. Juv.	Chinook Win. Juv.	Chinook Redd	Steelhead Sum. Juv.	Steelhead Win. Juv.	Steelhead Redd
		and last station is 0.5 x station width x depth x velocity.						
Size	Discharge - Fish	Discharge at time of fish survey	–	1	–	–	1	–
Size	Gradient	Site water surface gradient is calculated as the difference between the top of site (upstream) and bottom of site (downstream) water surface elevations divided by thalweg length.	–	–	2	–	–	4
Size	Thalweg Depth Avg	Average thalweg depth of the wetted channel.	1	–	–	2	–	–
Size	Thalweg Exit Depth	Depth of the thalweg at the downstream edge of the channel unit.	–	3	–	–	2	–
Substrate	Substrate < 6mm	Average percentage of pool tail substrates comprised of sediment <6 mm.	11	–	–	13	–	–
Substrate	Substrate Est: Boulders	Percent of boulders (256-4000 mm) within the wetted site area.	–	–	4	–	–	5
Substrate	Substrate Est: Coarse and Fine Gravel	Percent of coarse and fine gravel (2-64 mm) within the wetted site area.	10	7	5	10	5	11
Substrate	Substrate: D50	Diameter of the 50th percentile particle derived from pebble counts.	12	5	–	8	7	–
Temperature	Max7dAM	Highest 7-day average of daily maximum (7dAM) value between July 15th - August 31st.	7	–	10	6	–	12
Temperature	Summer Hourly Average Temp	Average of all hourly temperature measurements collected July 15th - August 31st.	2	–	–	3	–	–
Water Quality	Conductivity	Measure of the concentration of ionized materials in water, or the ability of water to conduct electrical current.	3	–	1	1	–	9
Wood	Large Wood Frequency: Wetted	Number of large wood pieces per 100 meters within the wetted channel.	13	–	14	11	–	10

QRF Model Fitting

Quantile regression forest models were fit for each of the species and life stages using the chosen habitat covariates (Table 1) and the *quantregForest* package (Meinshausen 2016) in R (R Core Team 2015). Table 1 provides the relative importance of each habitat covariate included in each of the QRF models, after model fit. Additionally, QRF models allow one to visually examine the marginal effect of each habitat covariate on the quantile of interest using partial dependence plots. These plots show the marginal effect of changing a single covariate on the response variable while maintaining all other covariates at their mean values (see Supplemental Figure 1 - Supplemental Figure 6).

Model Extrapolation

After model fitting, QRF models can be used to predict capacity for a given species and life stage at all CHaMP sites within the interior Columbia River Basin, using the 90th quantile of the predicted distribution as a proxy for carrying capacity. CHaMP site carrying capacity predictions from each of the six QRF models (Chinook/steelhead summer parr, overwinter presmolt, and redds) were extrapolated to larger scales (e.g., watershed, population) using GAA covariate data. In this report, we provide capacity extrapolation results for both Chinook salmon and steelhead, for summer parr, overwintering juveniles, and redds, for select watersheds (Lemhi and Upper Grande Ronde) within the Columbia River Basin (see Supplementary Tables and Figures). In the Supplementary Tables and Figures, we provide estimates of habitat capacity, by life stage and species, for tributaries and mainstem habitats in the Lemhi and Upper Grande Ronde rivers. Moreover, we provide maps to visualize predictions of parr and redd capacity at all master sample points in these watersheds using the extrapolation model. These capacity tables and maps provide an example of outputs available from our current QRF and extrapolation models; estimates for other watersheds or whole populations are available upon request.

Model Validation

Estimates of Chinook salmon parr capacity from the QRF and extrapolation models compared well with independent estimates from spawner-recruit data (

Table 2, Figure 1, Figure 2). QRF estimates had overlapping confidence intervals with one or more of the Beverton-Holt, Ricker, or hockey stick model estimates in each of the eight locations where comparisons were made. The spawner-recruit curves may have under-represented uncertainty, as estimates of redds, spawners and parr were assumed to be accurate. This would increase the confidence intervals around spawner-recruit estimates and increase overlap among estimates. Correlations between parr capacity estimates between the QRF model and spawner-recruit models ranged from 0.635 (hockey stick) to 0.708 (Ricker) (Figure 2). This favorable comparison provides strong validation, as the spawner-recruit estimates of Chinook salmon parr capacity were developed from completely independent datasets and using entirely different methods compared to the QRF model.

Table 2. Estimates of parr capacity in selected watershed from both spawner-recruit data (Beverton-Holt, Ricker, hockey stick) and from extrapolated estimates of parr capacity from the quantile regression forest (QRF) model. Numbers in parentheses are coefficient of variation. ODFW = Oregon Department of Fish and Wildlife; WDFW = Washington Department of Fish and Wildlife; IDFG = Idaho Department of Fish and Game; RST = Rotary Screw Trap; QCI = Quantitative Consultants, Inc.

Watershed	Years	Adult Data	Adult Source	Parr Data	Parr Source	Beverton-Holt	Ricker	Hockey Stick	QRF
Catherine Creek	20	Spawners	ODFW	RST	ODFW	135,387 (0.27)	103,021 (0.14)	99,921 (0.21)	119,769 (0.1)
Chiwawa R.	20	Spawners	WDFW	Parr Surveys	BioAnalysts	248,586 (0.24)	166,139 (0.15)	174,216 (0.18)	329,891 (0.05)
Hayden Creek	7	Redds	IDFG	RST	IDFG/QCI	45,111 (0.43)	35,562 (0.21)	35,949 (0.31)	46,637 (0.33)

Watershed	Years	Adult Data	Adult Source	Parr Data	Parr Source	Beverton-Holt	Ricker	Hockey Stick	QRF
Lostine River	17	Redds	ODFW	RST	ODFW	196,260 (0.24)	146,982 (0.16)	144,415 (0.2)	197,560 (0.13)
Minam River	14	Spawners	ODFW	RST	ODFW	–	484,810 (1.44)	662,806 (1.73)	504,892 (0.02)
Pahsimeroi River	18	Redds	IDFG	RST	IDFG	213,017 (1.3)	83,089 (0.9)	110,305 (1.05)	108,283 (0.22)
Upper Grande Ronde River	8	Spawners	ODFW	RST	ODFW	171,609 (0.39)	168,137 (0.3)	127,052 (0.32)	132,940 (0.16)
Upper Lemhi	22	Redds	IDFG	RST	IDFG/QCI	112,746 (0.78)	64,303 (0.5)	72,470 (0.64)	94,419 (0.19)

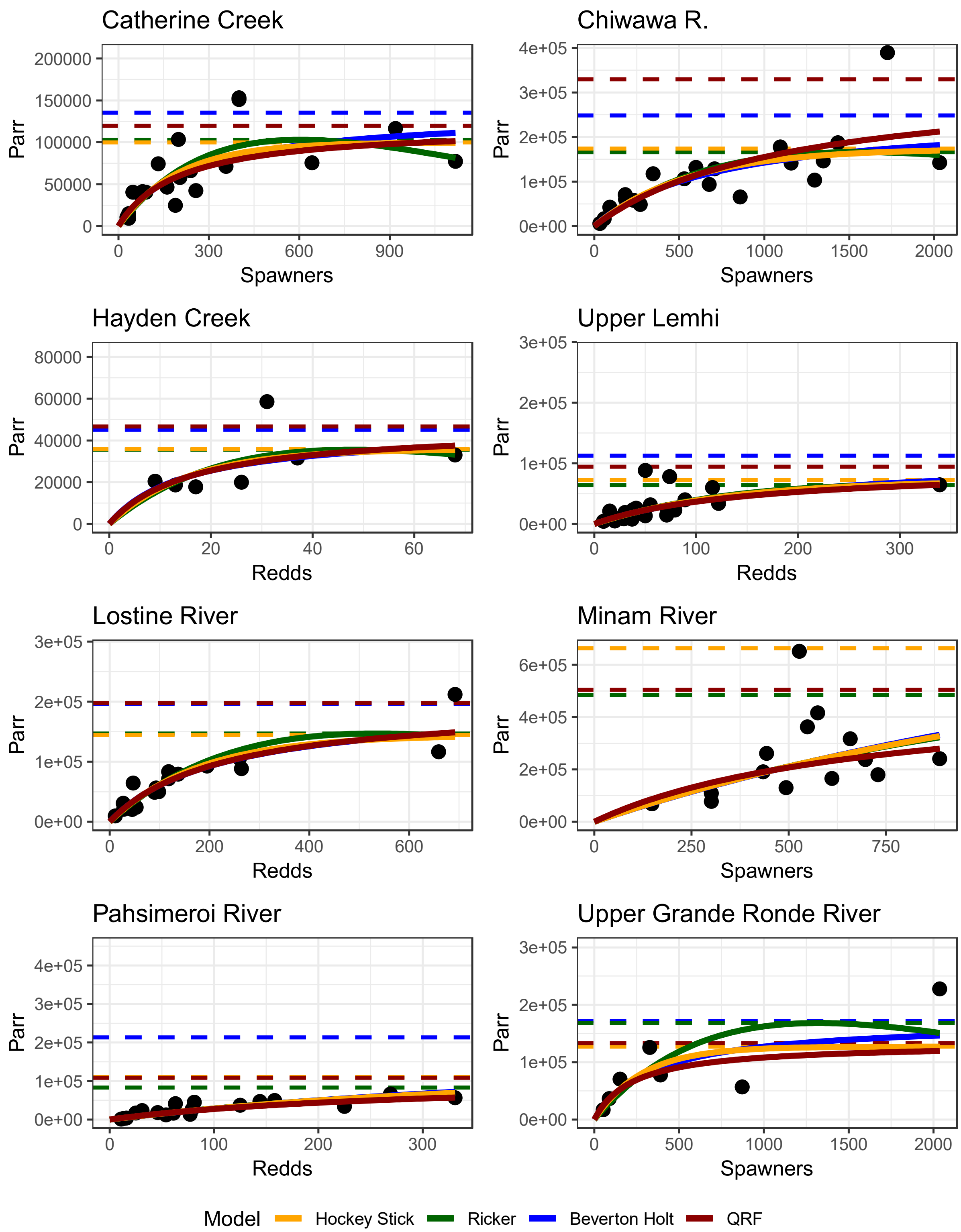


Figure 1. Spawner-recruit data from eight watersheds. Solid lines are the spawner-recruit curve, dashed lines are the estimated capacity, and shaded polygons depict the 95% confidence intervals of capacity. Blue corresponds to Beverton-Holt models, green to Ricker models, yellow to hockey stick models, and red to QRF estimates. The QRF solid curve is a Beverton-Holt model with the capacity parameter fixed to the QRF estimate of capacity.

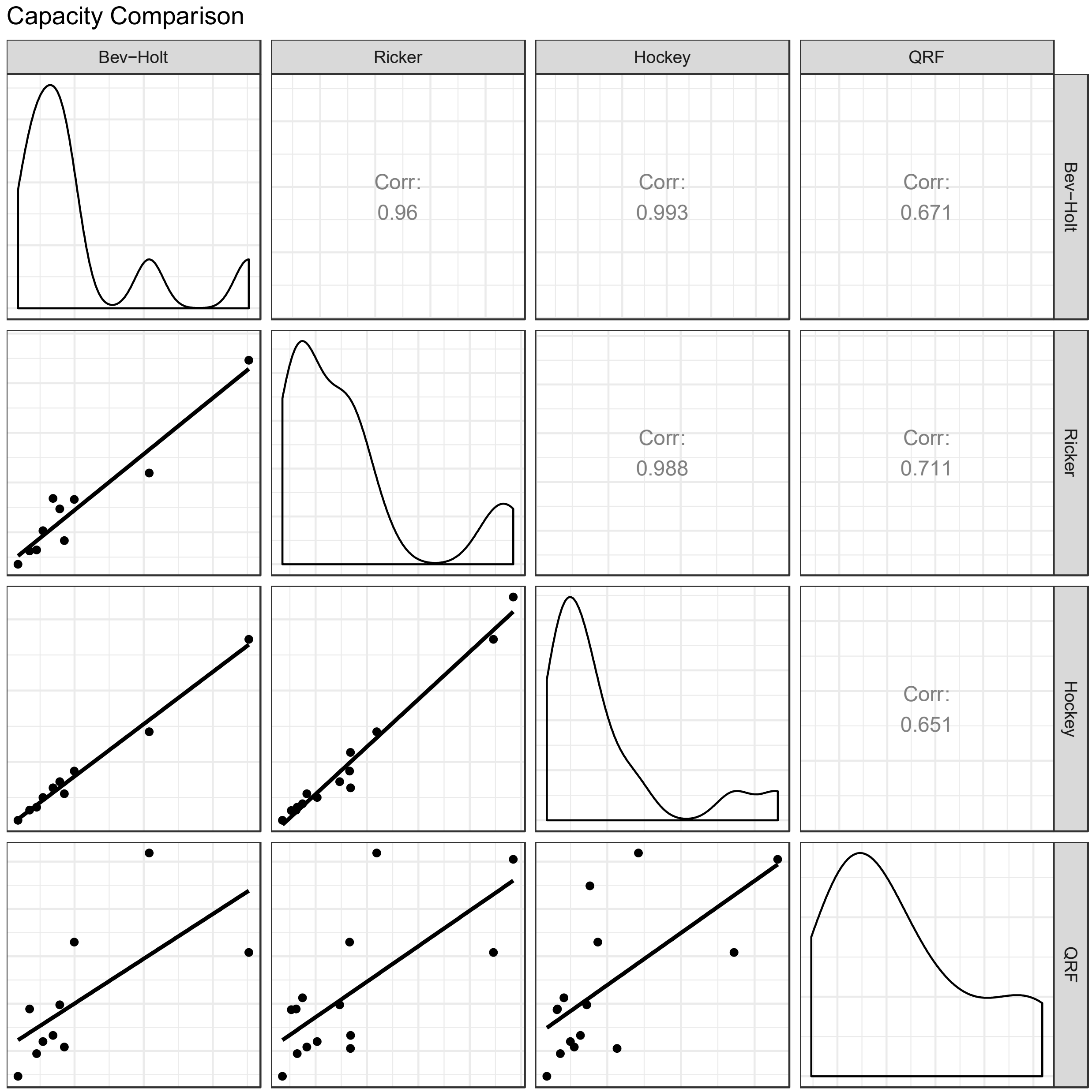


Figure 2. Scatterplots showing correlations among estimates of Chinook salmon parr carrying capacity from the quantile regression forest (QRF) model and three spawner-recruit models in selected watersheds.

DISCUSSION

We have described a novel approach to estimate the capacity of habitat, in wadable streams, to support Chinook salmon and steelhead juveniles (during summer and winter months) and redds in the interior Columbia River Basin. We have built QRF models for three different life-stages (summer parr, overwintering juveniles, and redds) for two different species (Chinook salmon and steelhead). The approach is entirely empirical, allowing fish-habitat relationships to emerge from the input data, even if they are non-linear in nature (as most ecological relationships are). For these species and life-stages, we have generated estimates of capacity where similar habitat data are available (i.e., at all CHaMP sites). In this report, we provide estimates of habitat capacity, and capacity maps, for the Lemhi and Upper Grande Ronde rivers that provide an example of outputs that can be produced using the QRF models. To date, we have validated the QRF estimates of Chinook summer parr with spawner-recruit curves from a variety of watersheds, and found them to match up very well, despite being based on entirely different data. Additional, QRF predictions of capacity can be built on habitat sampling conducted over a handful of years (or a single year with enough effort), whereas spawner-recruit curves, while often considered a gold standard for estimating capacity, require many years of data with plenty of contrast to be considered valid.

There are potential limitations to our approach that should be considered when interpreting results. First, we assume that at least some sites in our empirical dataset are at or near carrying capacity at the site level. Having at least some sites near capacity allows the random forest model to more accurately provide classification and regression trees, which in turn, allows better approximation of quantiles and capacity. However, this assumption may not be true in our case, especially since juvenile fish abundance/density and redd data used in the model have been collected during recent years of low escapement. Fortunately, if this assumption is not met, the QRF models will likely produce conservative (low) estimates of capacity (but the framework of the model is not wrong). To address this limitation, we hope to add paired fish-habitat data in the future from areas of increased escapement or that are likely near rearing capacity (e.g., Secesh River, ID or regions outside the Columbia River [Alaska]) to provide more accurate estimates of capacity. Adding fish-habitat data from additional areas has the added benefit of providing additional contrast in habitat data to the model, which can improve model predictions and extrapolation.

Our QRF models are populated using CHaMP habitat data and juvenile fish or redd abundance and density information collected within those CHaMP sites. Predictions of habitat capacity can then be made at locations where similar habitat data are available (i.e., all CHaMP sites) and then extrapolated to larger spatial scales using globally available attributes or similar. However, there are a few issues related to the extrapolation of QRF estimates of capacity to larger spatial scales that should be noted. Namely, it is difficult to make predictions of capacity and/or extrapolate estimates in areas where the habitat characteristics fall outside the habitat measurements used in the model (e.g., Clearwater River basin). Some master sample points had covariate values well outside the range of CHaMP sites, making it difficult and risky to extrapolate to those areas. The points are scattered infrequently within most watersheds in the interior Columbia River Basin, but some areas, notably the Clearwater River basin, are full of them. To address this, we hope to include paired fish-habitat data from areas such as the Clearwater to increase the scope of the models to those areas. Additionally, determining the downstream extent of wadable streams can be a challenge, and whether all the master sample points we include meet that definition is unclear. Fish-habitat relationships may change in deeper rivers, and these QRF models should currently only be applied to wadable areas of a watershed. In the future, we hope to explore the ability to apply QRF models to larger river systems where desired.

We recognize that occasionally estimates of winter capacity for a particular stream or watershed are higher for winter juveniles than summer parr (e.g., upper Grande Ronde steelhead), which was contrary to our expectations. Although this, in fact, may be true, there are other alternative potential expectations for these results. First, we assumed that the spatial extent for rearing during summer and winter months was the same. In reality, the winter extent for each species is likely not as broad as the summer rearing extent, so even if more fish could be supported at some sites during winter, there may be extents of the watershed not available to overwintering fish. However, without knowing the true winter distributional extent it is difficult to correct for varying summer/winter extents. To date, our winter fish sampling has focused on areas within the domain of Chinook salmon, so we do not have the observational data to restrict a species' winter range appropriately, and such data would be difficult to obtain.

Next Steps

The QRF models presented here are currently populated using habitat data collected by CHaMP (CHaMP 2017). However, with the reduction in effort of on-the-ground habitat data collection (CHaMP), habitat data and covariates used in the model may become outdated as habitat evolves year-to-year via natural and/or anthropogenic changes. As a result, the need for a broader, watershed scale, and cost-effective approach to sampling riverine habitat to populate fish-habitat models has become apparent. Remote sensing techniques, paired with minimal, streamlined on the ground sampling may allow for more rapid habitat data collection, at increased scale, and in a more cost-effective manner. Fish-habitat models, including QRF, would benefit by incorporating habitat data collected via remotely sensed platforms and at a greater scale. Emerging techniques, such as multi-spectral analysis, bathymetric LiDAR, and high-resolution RGB cameras are becoming more affordable and attainable for watershed scale habitat data collection. These type of data and sampling techniques may allow for cost-effective habitat sampling to occur over potentially large spatial scales. Further, if data availability via remotely sensed habitat information is adequate in detail and spatial scale, the need for extrapolation models may be removed completely. Use of continuous, remotely sensed habitat data, at the watershed scale, would provide accurate habitat data that can be used in QRF and similar fish-habitat data, while decreasing costs and potentially removing the need for extrapolation models where remotely sensed data are available.

QRF capacity modeling allows for the incorporation of multiple, adaptive habitat data sources and inputs, and so would benefit from improvements in habitat data collection, not only spatially and temporally, but also technologically. In addition, a need for a more comprehensive and adaptable toolkit for processing habitat data from various platforms (e.g., LiDAR, RGB imagery) would aide in rolling a new program out to multiple watersheds and collaborators. An open-source, adaptable toolkit could be developed to incorporate on-the-ground habitat data with remotely sensed data (e.g., aerial imagery, bathymetric LiDAR, satellite information, other future technologies). Habitat data collection is rapidly evolving and improving, and so would benefit from tools that can leverage new and developing technologies to incorporate habitat data into QRF capacity and similar models.

Habitat rehabilitation groups have requested further guidance on identification of limiting factors for Chinook salmon and steelhead and paths to address those limiting factors. Currently, fish and habitat data metrics used in our QRF models are collected at the reach (200 to 500 m) scale. However, fish and habitat can be very heterogeneous within that scale, and thus, identifying fine-scale (channel unit) fish-habitat relationships within the data can be difficult. Ideally, we would like to better understand fish-habitat relationships within individual channel units (e.g., pools, riffles, runs). Understanding relationships within individual channel units would allow us to identify what characteristics provide a high capacity pool, riffle, or similar, and further, would provide information on appropriate configuration of channel units to increase habitat capacity. We hope to build QRF models for estimating summer parr rearing capacity at the channel unit scale, similar to the winter presmolt capacity model we present here. A channel unit scale model would help to better translate fish-habitat relationships and allow for a more applicable assessment of restoration evaluation at a finer spatial scale. The channel unit scale is closer to the biological patches that fish occupy. Therefore, we hope to collect fish and habitat data at the channel unit scale in the future, and data can be 'lumped' to larger scales if desired. Channel unit scale information can be directly applied to restoration design and evaluation and assist engineers and geomorphologists.

Conclusion

Carrying capacity models based on QRF and habitat data, like those presented here, provide managers a framework to guide the identification, prioritization, and development of habitat rehabilitation actions to recover salmon populations. For perhaps the first time, the necessity of tributary habitat restoration actions can be demonstrated, and the magnitude of required change can be placed in context with the other ‘H’s’.

ADAPTIVE MANAGEMENT & LESSONS LEARNED

Our approach and models can be used to quantify juvenile rearing and adult spawning capacity in Chinook salmon and steelhead watersheds in the interior Columbia River Basin, and in turn, quantify the magnitude of tributary habitat rehabilitation necessary to support ESA delisting and recovery goals. Moreover, the models presented here could be applied to multiple stages within the life cycle (e.g., parr, presmolt, smolt, adult). Estimates of carrying capacity of habitat for multiple life stages will allow biologists and managers to identify where in the life cycle that habitat may be limiting. For example, QRF models and associated extrapolation models may demonstrate that habitat for a given population may be sufficient to support adult spawning required for ESA delisting, but that juvenile capacity may not be sufficient to, in turn, support those levels of adult spawning. In such a case, habitat restoration actions may be most cost-effectively and sustainably directed towards juvenile rearing habitat. Models to estimate habitat carrying capacity for multiple life stages will help to better direct habitat restoration actions. The QRF and extrapolation models provide useful tools towards the prioritization, implementation, and evaluation of habitat restoration actions to recover depleted populations.

More specifically, the QRF models presented here, or similar variations, could be used to:

- Identify low, medium, or high priority areas for habitat rehabilitation, by species and life stage. For example, capacity estimate and maps could be used to identify specific reaches or areas with low or high capacity to support Chinook salmon and steelhead life stages of interest. Such information could be combined with geomorphic information to identify low, medium, and high priority for habitat rehabilitation.
- Describe target habitat conditions to support increased capacity for life stages of Chinook salmon or steelhead. For example, if populated using fish and habitat data collected at the channel unit scale, QRF models could be used to describe those habitat characteristics for channel units that support increased capacity (i.e., what makes a high capacity pool?).
- Evaluate the potential of various rehabilitation designs. For example, these QRF models, or ones like them, could be used to evaluate a suite of potential rehabilitation actions and compare the predicted improvements to capacity to the current state to determine the most appropriate or cost-effective plan of action. Similarly, these models can be used to assess the potential capacity that could be added by reconnecting tributaries and opening additional habitat.
- Provide pre- and post-restoration monitoring to estimate the benefit of actions. Habitat data could be collected both before and after restoration activities so that capacity at a given site can be estimated before and after action, and in turn, estimate the capacity increase or benefit of that action.

Here, we presented QRF capacity predictions and maps for the Lemhi River and Upper Grande Ronde rivers. The QRF capacity predictions of capacity at all master sample points can be exported as shapefiles for easy distribution of spatially explicit data. We recommend discussions about how to best apply QRF results for particular projects or studies, so that we can best provide the most appropriate versions of predictions.

ACKNOWLEDGEMENTS

QRF development efforts have been funded by the Bonneville Power Administration through their ISEMP and CHaMP projects. Fish sampling work in the Lemhi was also funded through the Idaho Office of Species Conservation Pacific Coast Salmon Recovery Fund. Redd QRF model development was funded by BPA and through the Office of Species Conservation through a cooperative agreement with the Bureau of Reclamation. Special thanks to staff from the following agencies for providing spatial redd data: Idaho Department of Fish and Game, Nez Perce Tribe Department of Fisheries Resources, Oregon Department of Fish and Wildlife, U.S. Fish and Wildlife Service, and Washington Department of Fish and Wildlife. Special thanks to staff from the following agencies for providing data to calculate juvenile fish abundance and density estimates: Columbia River Inter-Tribal Fish Commission, Oregon Department of Fish and Wildlife, Washington Department of Fish and Wildlife and U.S. Fish and Wildlife Service. The models described in this study were improved by conversations with Eric Buhle.

REFERENCES

Albanese, D., M. Filosi, R. Visintainer, G. Jurman, and C. Furlanello. 2012. Minerva and minepy: a C engine for the MINE suite and its R, Python, and MATLAB wrappers. Bioinformatics. 29(3):407-408.

Breiman, L. 2001. Random forests. Machine Learning. 45(1):5-32.

Cade, B.S. and B.R. Noon. 2003. A gentle introduction to quantile regression for ecologists. Frontiers in Ecology and the Environment. 1:412-420.

Carle, F.L. and M.R. Strub. 1978. A new method for estimating population size from removal data. Biometrics. 34:621-630.

Chapman, D.G. 1951. Some properties of the hypergeometric distribution with applications to zoological sample censuses. University of California Press.

Hedger, R.D., E. De Eyto, M. Dillane., O.H. Diserud, K. Hindar, P. McGinnity, R. Poole, and G. Rogan. 2013. Improving abundance estimates from electrofishing removal sampling. Fisheries Research. 137:104-115.

ISEMP/CHaMP. 2015. Combined Annual Report for the Integrated Status and Effectiveness Monitoring Program and Columbia Habitat Monitoring Program: 2014. Prepared by ISEMP and CHaMP for the Bonneville Power Administration. Published by Bonneville Power Administration.

Knudby, A., A. Brenning, and E. LeDrew. 2010. New approaches to modeling fish-habitat relationships. Ecological Modeling. 221(3):503-511.

Liaw, A. and M. Wiener. 2002. Classification and regression by randomForest. R news 2:18-22.

Lumley, T. 2016. survey: analysis of complex survey samples. R package version 3.31-5.

Lumley, T. 2004. Analysis of complex survey samples. Journal of Statistical Software. 9(1):1-19.

McNyset, K.M., C.J. Volk, and C.E. Jordan. 2015. Developing an Effective Model for Predicting Spatially and Temporally Continuous Stream Temperatures from Remotely Sensed Land Surface Temperatures. Water. 7:6827-6846.

Meinshausen, N. 2006. Quantile regression forests. Journal of Machine Learning Research. 7:983-999.

Meinshausen, N. 2016. quantregForest: Quantile Regression Forests. R package version 1.3-5. https://CRAN.R-project.org/package=quantregForest

Ogle, D.H. 2016. FSA: Fisheries Stock Analysis: R package version 0.8.11.

Prasad, A., L. Iverson, and A. Liaw. 2006. Newer classification and regression tree techniques: Bagging and random forests for ecological prediction. Ecosystem. 9(2):181-199.

R Core Team. 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Reshef, D.N., Y.A. Reshef, H.K. Finucane, S.R. Grossman, G. McVean, P.J. Turnbaugh, E.S. Lander, M. Mitzenbacher, and P.C. Sabeti. 2011. Detecting Novel Associations in Large Data Sets. Science. 334:1518-1524.

Rivest, L.P. and S. Baillargeon. 2014. Rcapture: Loglinear models for capture-recapture experiments. R package version 1.4-2.

Robson, D. and H. Regier. 1964. Sample size in Peterson mark-recapture experiments. Transactions of the American Fisheries Society. 93:215-226.

Seber, G. 2002. The estimation of animal abundance and related parameters. Blackburn Press Caldwell, New Jersey.

Stekhoven, D.J. 2013. missForest: Nonparametric Missing Value Imputation using Random Forest. R package version 1.4.

Stekhoven, D.J. and P. Buehlmann. 2012. MissForest – non-parametric missing value imputation for mixed-type data. Bioinformatics. 28(1):112-118.

Stevens, D. and A. Olsen. 2004. Spatially balanced sampling of natural resources. Journal of the American Statistical Association. 99:262-278.

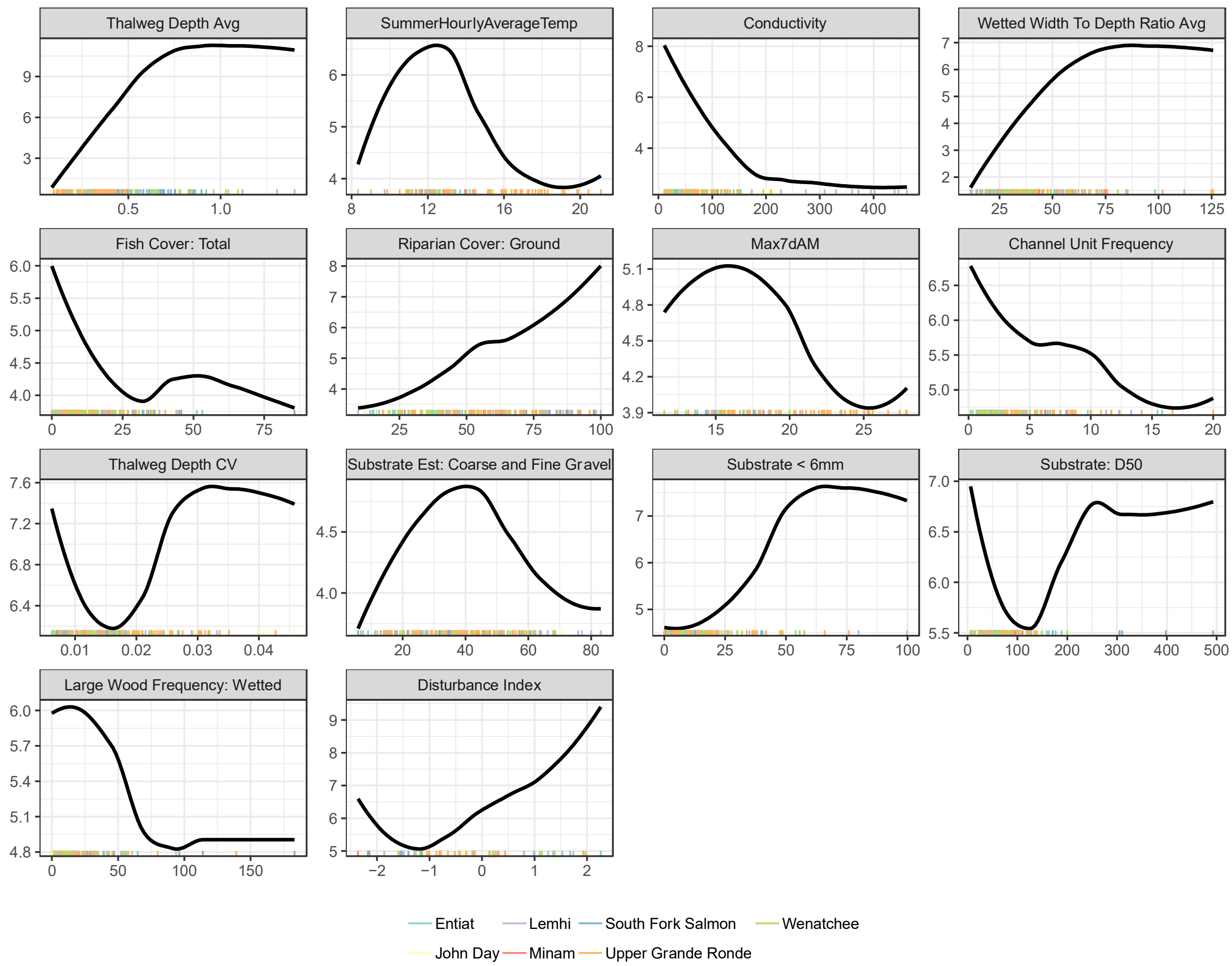
Terrell, J.W., B.S. Cade, J. Carpenter, and J.M. Thompson. 1996. Modeling Stream Fish Habitat Limitations from Wedge-Shaped Patterns of Variation in Standing Stock. Transactions of the American Fisheries Society. 125(1):104-117.

Whittier, T., A. Herlihy, C. Jordan, and C. Volk. 2011. Landscape classification of Pacific Northwest hydrologic units based on natural features and human disturbance to support salmonid research and management. NOAA, National Marine Fisheries Service. NOAA Contract #AB1133F10SE2464. Pp. 39.

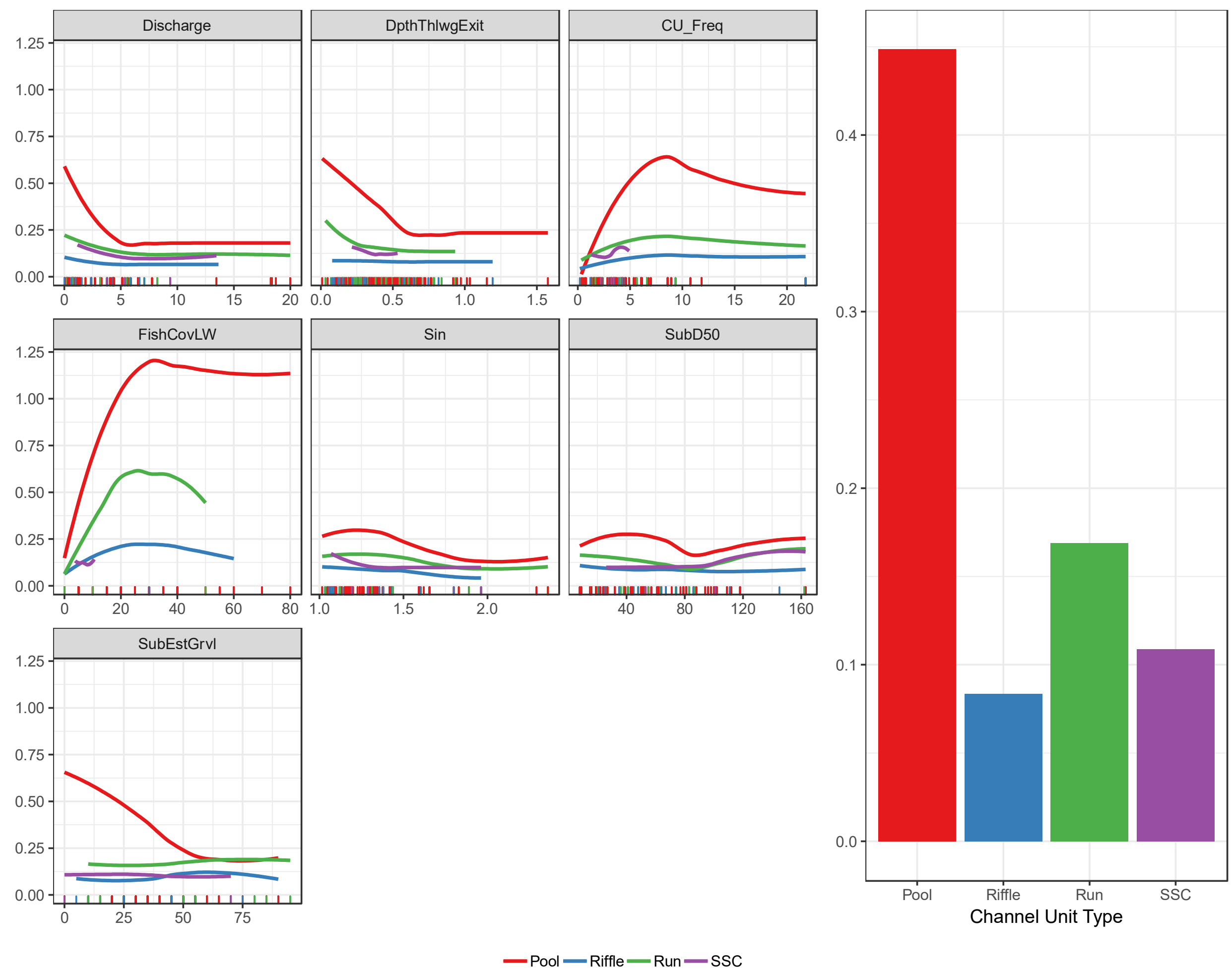
SUPPLEMENTAL TABLES AND FIGURES

Partial Dependence Plots

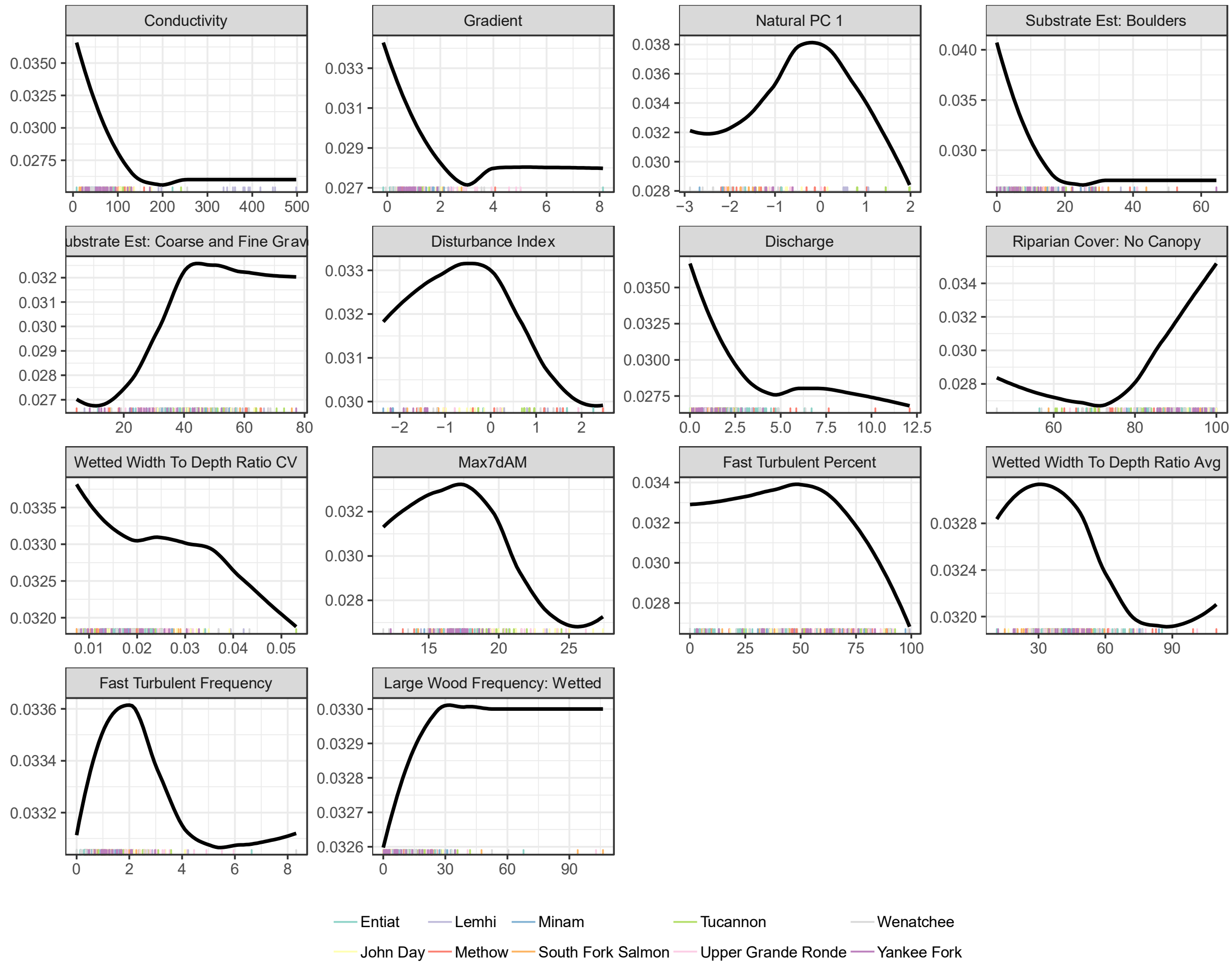
Chinook



Supplemental Figure 1. Partial dependence plots from the Chinook salmon parr (summer) capacity quantile regression forest (QRF) model, depicting how parr capacity shifts as each habitat metric changes, assuming all other habitat metrics remain at their mean values. Tick marks along the X-axis depict observed values, and the sub-basin they came from.

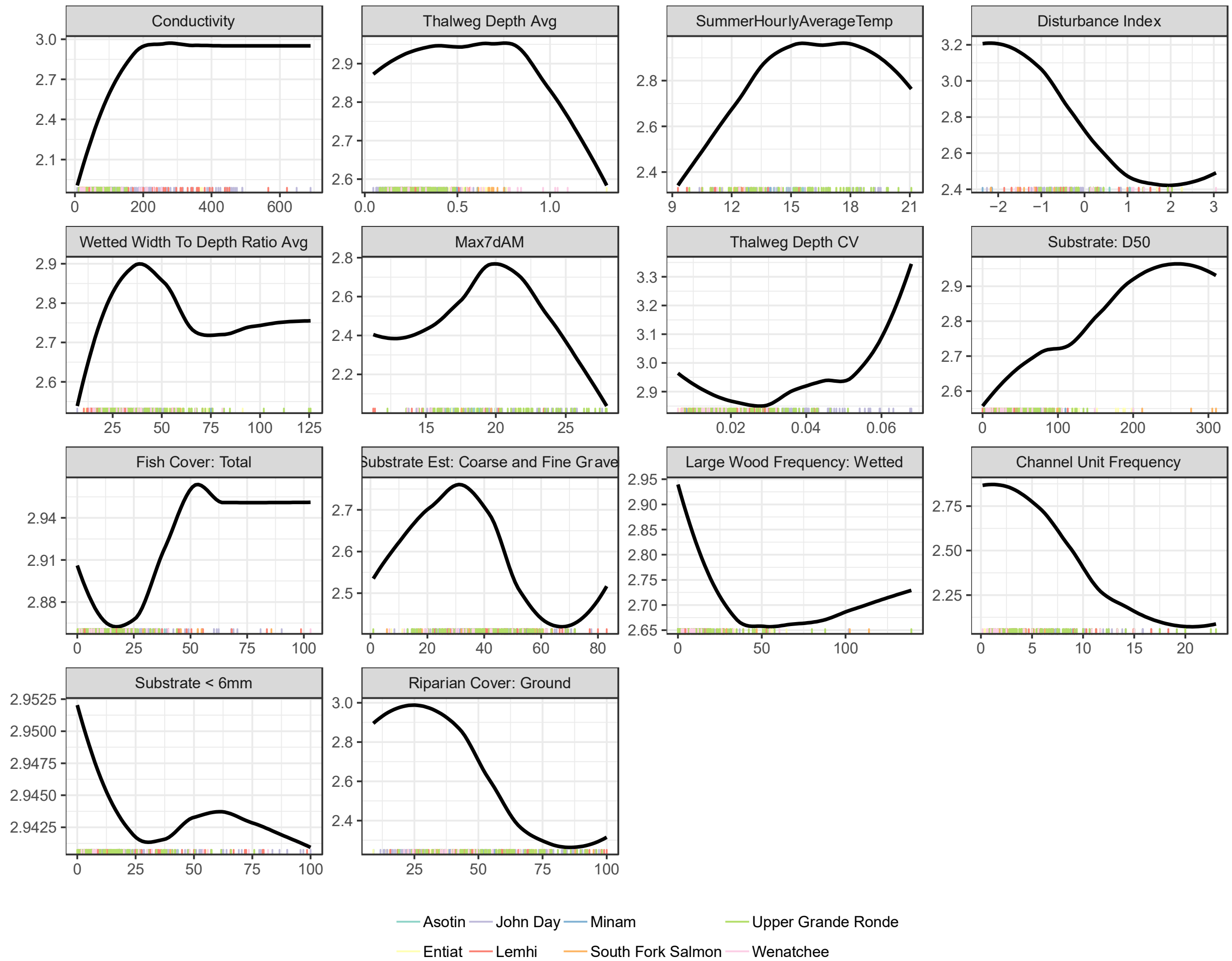


Supplemental Figure 2. Partial dependence plots from the Chinook salmon parr (winter) capacity quantile regression forest (QRF) model, depicting how parr capacity shifts as each habitat metric changes, assuming all other habitat metrics remain at their mean values. Tick marks along the X-axis depict observed values. Colors correspond to the type of channel unit (pool, riffle, run or small side channel).

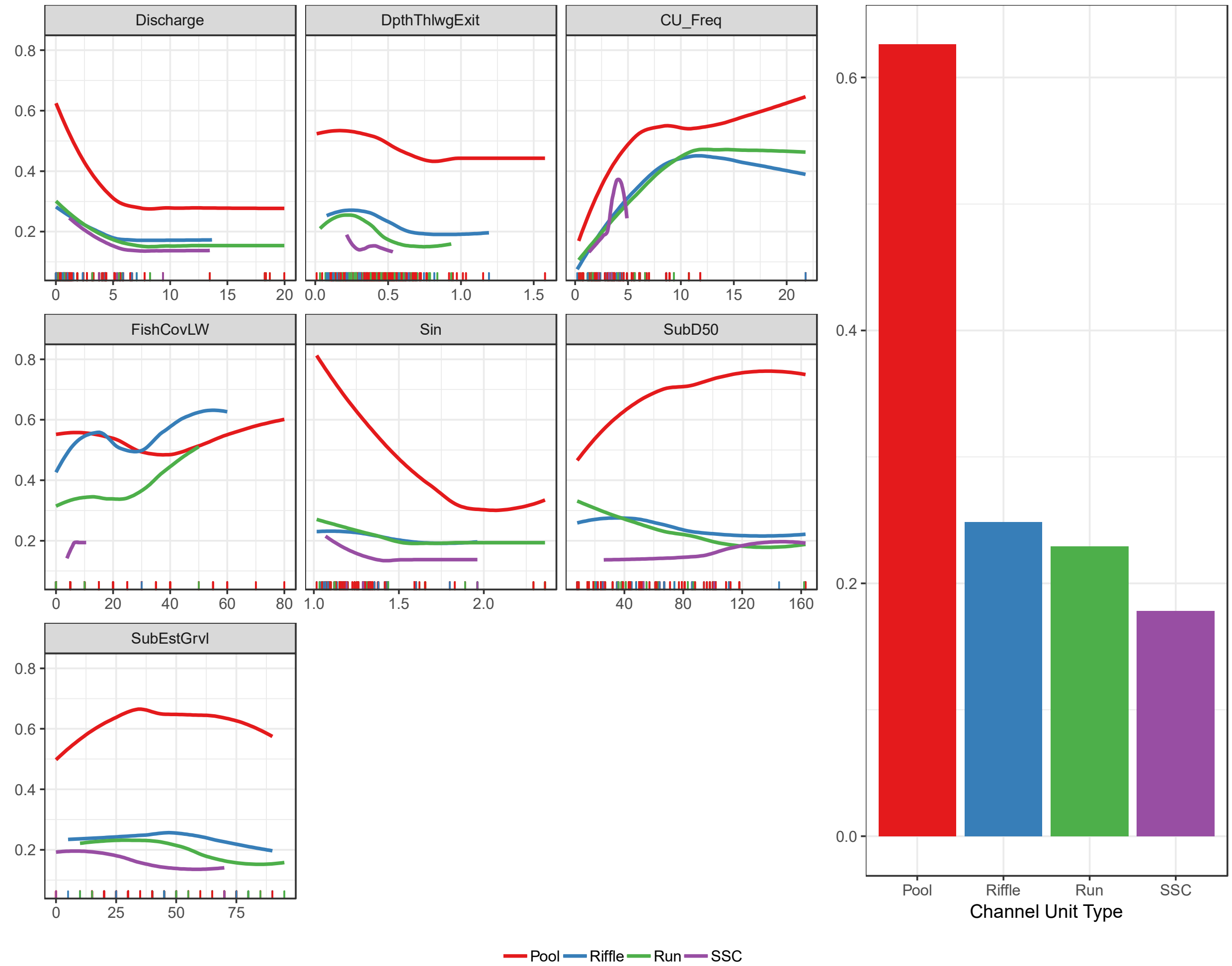


Supplemental Figure 3. Partial dependence plots from the Chinook salmon redd capacity quantile regression forest (QRF) model, depicting how redd capacity shifts as each habitat metric changes, assuming all other habitat metrics remain at their mean values. Tick marks along the X-axis depict observed values, and the sub-basin they came from.

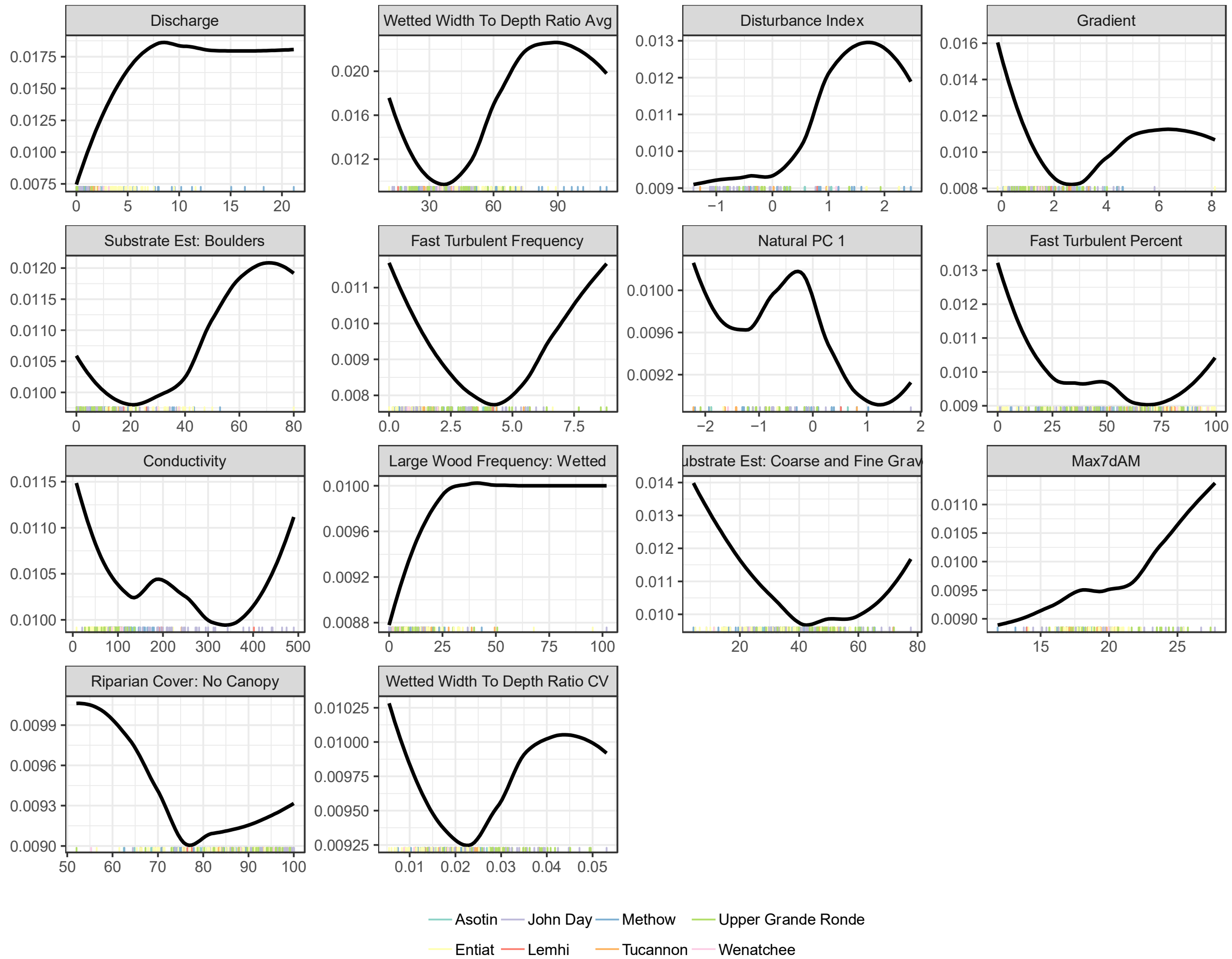
Steelhead



Supplemental Figure 4. Partial dependence plots from the steelhead juvenile (summer) capacity quantile regression forest (QRF) model, depicting how juvenile capacity shifts as each habitat metric changes, assuming all other habitat metrics remain at their mean values. Tick marks along the X-axis depict observed values, and the sub-basin they came from



Supplemental Figure 5. Partial dependence plots from the steelhead juvenile (winter) capacity quantile regression forest (QRF) model, depicting how juvenile capacity shifts as each habitat metric changes, assuming all other habitat metrics remain at their mean values. Tick marks along the X-axis depict observed values. Colors correspond to the type of channel unit (pool, riffle, run or small side channel).



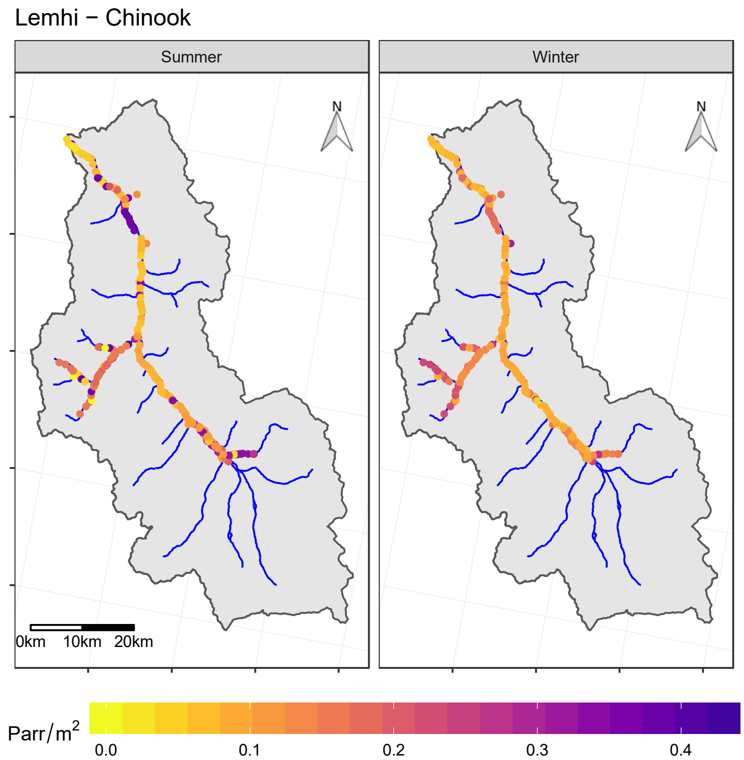
Supplemental Figure 6. Partial dependence plots from the steelhead redd capacity quantile regression forest (QRF) model, depicting how redd capacity shifts as each habitat metric changes, assuming all other habitat metrics remain at their mean values. Tick marks along the X-axis depict observed values, and the sub-basin they came from.

Chinook Salmon

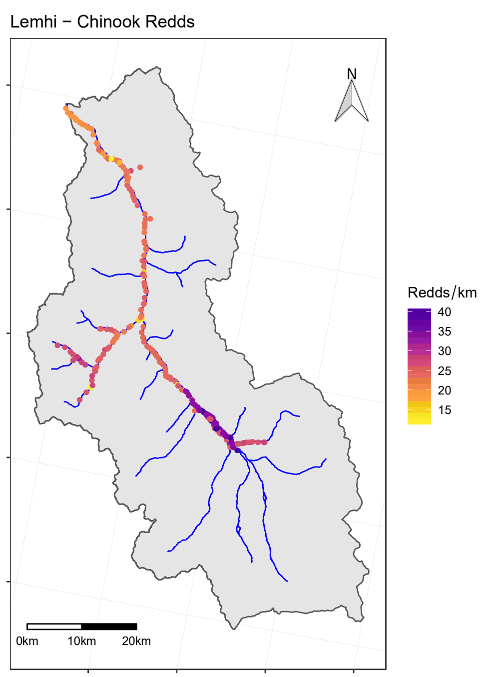
Lemhi River

Supplemental Table 1. Estimates of total Chinook salmon summer parr, overwintering parr and redd capacity for the Lemhi River, with standard error.

Stream	Stream length		Summer Parr		Winter Parr		Redd	
	km	mi	Capacity	SE	Capacity	SE	Capacity	SE
Basin Creek	4.3	2.7	3,235	2,410	3,485	744	108	7
Bear Valley Creek	9.4	5.9	4,949	3,855	8,889	1,524	256	15
Big Springs Creek	8	4.9	6,134	2,869	6,449	1,461	224	10
Hayden Creek	20.8	12.9	36,523	9,389	21,981	3,909	478	31
Lemhi River	92	57.2	200,856	36,395	104,687	16,523	2,367	138
Wright Creek	1.2	0.8	0	1,505	492	161	33	2
Total	**135.7**	**84.4**	**251,697**	**37,999**	**145,983**	**17,127**	**3,466**	**142.8**



Supplemental Figure 7. Map showing estimates of Chinook salmon summer and winter parr capacity (parr/m^2) for the Lemhi River.

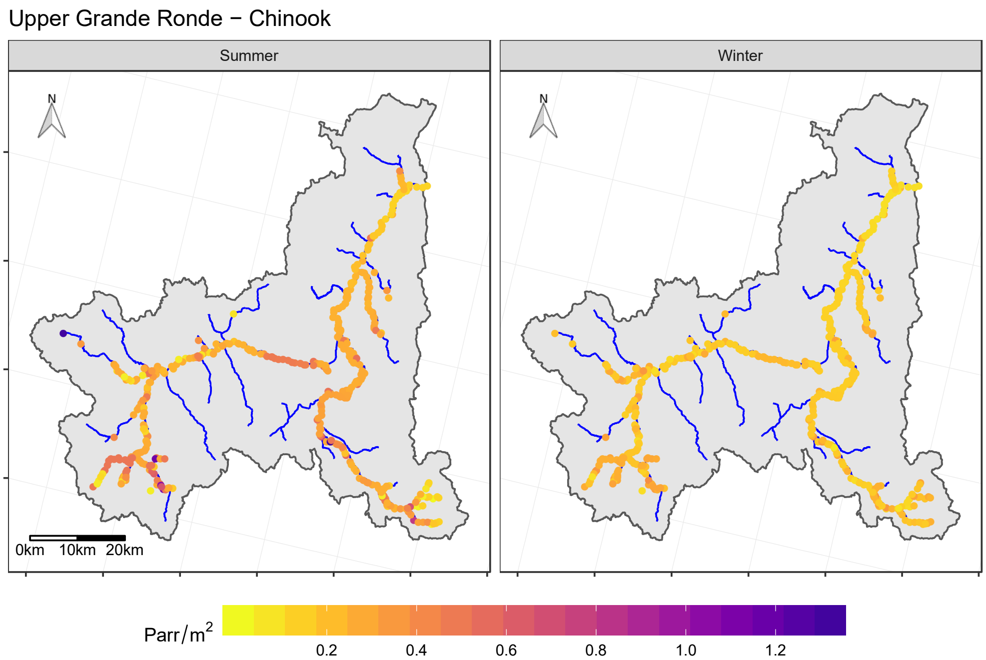


Supplemental Figure 8. Map showing estimates of Chinook salmon redd capacity (redds/km) for the Lemhi River.

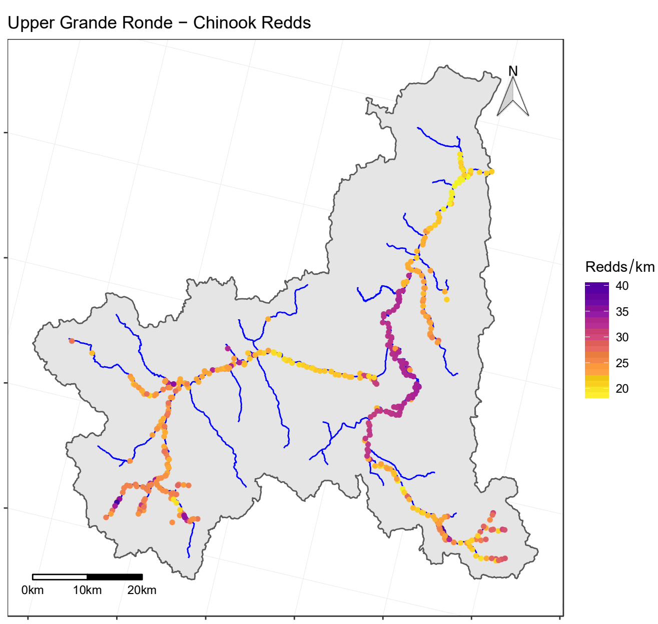
Upper Grande Ronde River

Supplemental Table 2. Estimates of total Chinook salmon summer parr, overwintering parr and redd capacity for the upper Grande Ronde River, with standard error.

Stream	Stream length		Summer Parr		Winter Parr		Redd	
	km	mi	Capacity	SE	Capacity	SE	Capacity	SE
Beaver Creek	23.5	14.6	59,950	8,600	43,138	4,161	576	27
Cabin Creek	28.7	17.8	100,174	19,486	61,726	10,913	601	53
Five Points Creek	33.3	20.7	127,691	22,630	65,017	10,502	730	61
Headwaters Grande Ronde River	73.7	45.8	170,204	23,907	104,973	9,512	1,912	81
Indian Creek	47.6	29.6	157,687	29,539	93,582	18,497	1,317	93
Lookingglass Creek	4.3	2.7	14,321	1,781	5,200	722	87	5
Lower Catherine Creek	57.3	35.6	243,867	50,625	140,511	31,783	1,822	136
Meadow Creek	17.9	11.1	36,743	4,711	29,524	1,841	448	16
Upper Catherine Creek	68.2	42.4	179,295	22,499	93,813	7,626	1,755	79
Total	**354.5**	**220.3**	**1,089,932**	**74,191**	**637,484**	**41,851**	**9,248**	**217.9**



Supplemental Figure 9. Map showing estimates of Chinook salmon summer and winter parr capacity (parr/m^2) for the upper Grande Ronde River.



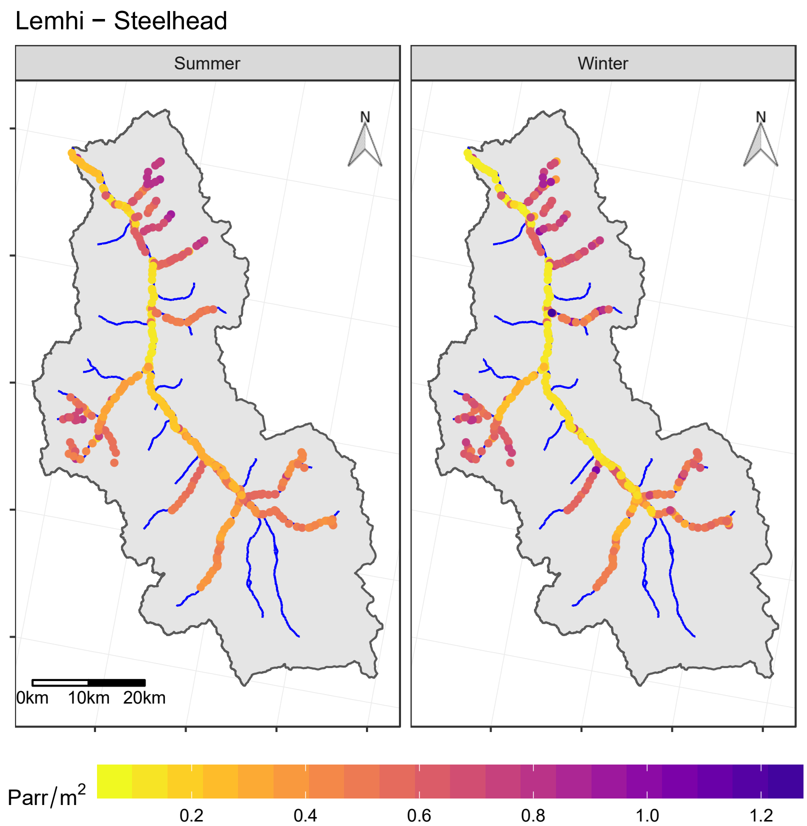
\Supplemental Figure 10. Map showing estimates of Chinook salmon redd capacity (redds/km) for the upper Grande Ronde River.

Steelhead

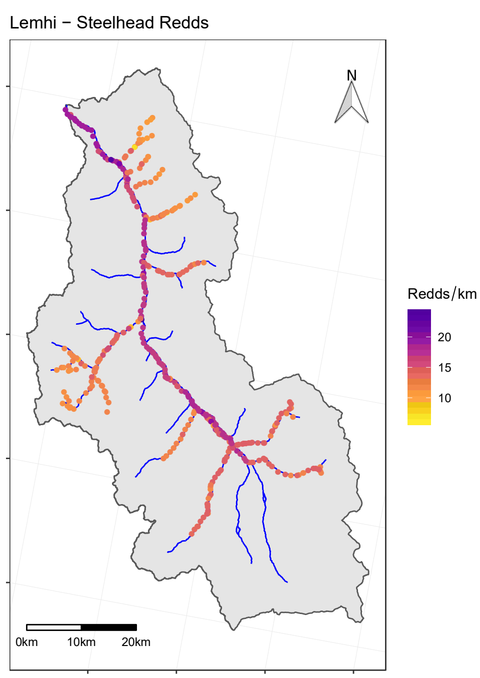
Lemhi River

Supplemental Table 3. Estimates of total steelhead summer juveniles, overwintering juveniles and redd capacity for the Lemhi River, with standard error.

Stream	Stream length		Summer Juv.		Winter Juv.		Redd	
	km	mi	Capacity	SE	Capacity	SE	Capacity	SE
Bear Valley Creek	7.8	4.8	14,173	1,094	18,312	1,282	94	6
Bohannon Creek	10.7	6.7	23,256	1,696	21,539	1,618	126	10
East Fork Hayden Creek	7.8	4.8	12,418	1,483	17,502	2,082	93	9
Hayden Creek	27	16.8	64,564	4,487	59,223	5,868	342	23
Kadletz Creek	2.6	1.6	4,507	494	4,462	597	30	4
Kenney Creek	11.8	7.4	22,839	1,697	26,828	1,766	139	9
Lemhi River	92	57.2	256,996	14,175	200,894	17,747	1,542	70
West Fork Hayden Creek	1	0.6	1,649	197	2,442	248	11	1
Wright Creek	4.4	2.7	7,780	875	8,418	1,404	52	6
Total:	**165.1**	**102.6**	**408,182**	**15,207**	**359,620**	**19,066**	**2,429**	**76.03**



Supplemental Figure 11. Map showing estimates of steelhead summer and winter juvenile capacity (fish/m^2) for the Lemhi River.

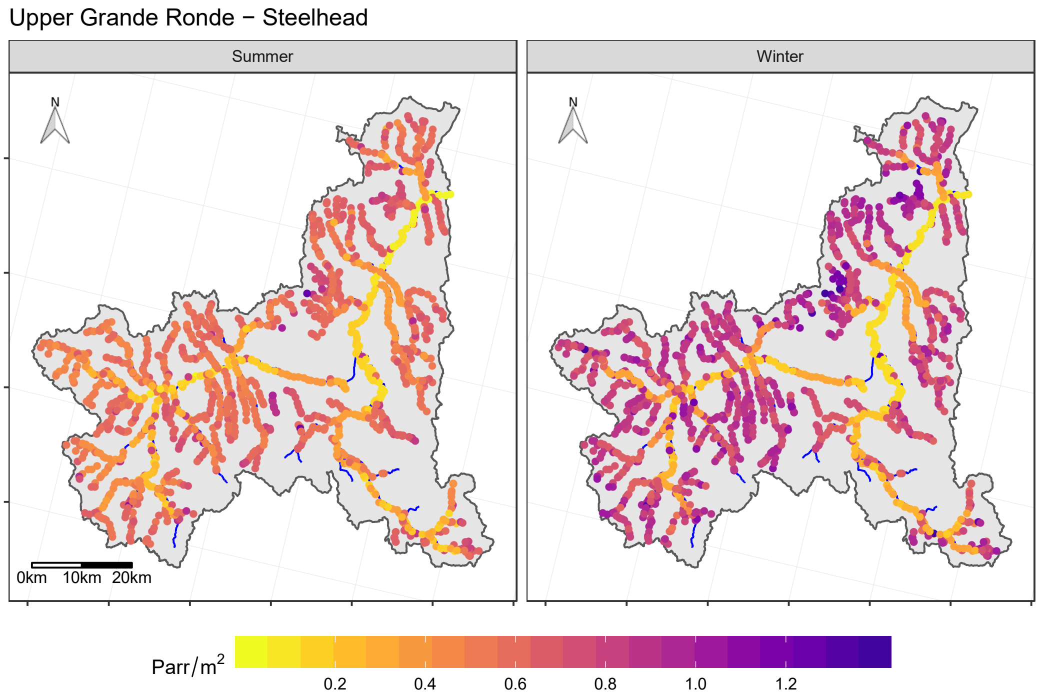


Supplemental Figure 12. Map showing estimates of steelhead redd capacity (redds/km) for the Lemhi River.

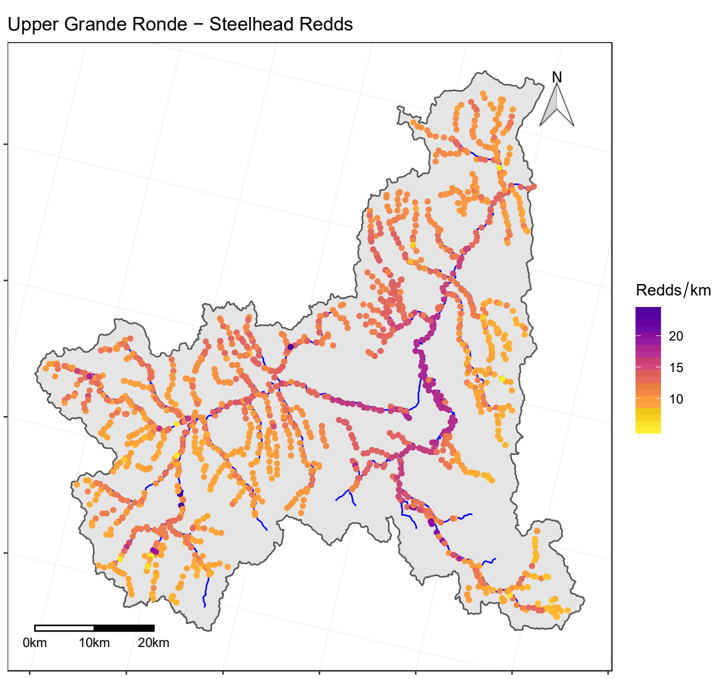
Upper Grande Ronde River

Supplemental Table 4. Estimates of total steelhead summer juveniles, overwintering juveniles and redd capacity for the upper Grande Ronde River, with standard error.

Stream	Stream length		Summer Juv.		Winter Juv.		Redd	
	km	mi	Capacity	SE	Capacity	SE	Capacity	SE
Beaver Creek	196.1	121.9	331,693	26,034	463,811	33,761	2,120	148
Cabin Creek	133.2	82.8	278,044	19,942	288,020	24,254	1,513	117
Five Points Creek	103.3	64.2	210,408	16,316	237,511	21,352	1,328	91
Headwaters Grande Ronde River	177	110	293,871	21,012	417,633	28,391	1,875	123
Indian Creek	137.9	85.7	257,716	21,282	286,790	29,073	1,550	126
Ladd Creek	46.4	28.8	110,753	9,984	96,342	15,737	600	44
Lookingglass Creek	82.9	51.5	152,163	11,775	173,973	15,966	879	71
Lower Catherine Creek	84.9	52.8	179,006	16,389	171,354	27,062	1,179	90
Meadow Creek	165.3	102.7	280,210	18,851	404,122	25,460	1,804	106
Upper Catherine Creek	107.4	66.7	192,561	12,125	237,985	17,259	1,211	72
Willow Creek	104.3	64.8	213,379	16,480	236,194	25,101	1,356	93
Total	**1,339**	**831.9**	**2,499,804**	**59,304**	**3,013,735**	**81,525**	**15,415**	**339.2**



Supplemental Figure 13. Map showing estimates of steelhead summer and winter juvenile capacity (fish/m^2) for the upper Grande Ronde River.



Supplemental Figure 14. Map showing estimates of steelhead redd capacity (redds/m) for the upper Grande Ronde River.