Extrapolating Capacity Estimates to a Linear Stream Network

Kevin See

July 24, 2020

Contents

[1] Biomark, Inc.

Introduction

For several years, our development of a QRF capacity model has relied on extrapolation to master sample points to interpolate between CHaMP (Columbia Habitat Monitoring Program) sites and generate capacity estimates on larger spatial scales. Recently there has been interest in moving from a point-based prediction layer to a line-based prediction layer to aide with the interpretation and visualization of the capacity predictions. In this document, we present our method for doing so, and include some comparisons with the point-based estimates.

Methods

Master Sample Points

The master sample points were generated in the design phase of CHaMP. These 551,046 sites were selected from the NHD Plus 1:100,000 stream layer covering WA, OR and ID at an average density of one site per kilometer (Larsen et al. 2016). Each CHaMP site where direct QRF capacity estimates were made corresponds to one of these master sample points. CHaMP generated a number o of attributes for each master sample point, referred to here as globally available attributes (GAAs) because they are associated with every master sample point across all watersheds.

The original extrapolation model used the log of capacity estimates at each CHaMP site (fish / m) as the response, and various GAAs as covariates. The model was fit using the *svyglm* function in the *survey* (Lumley 2004) package with R software (R Core Team 2019), accounting for the various survey design weights within each CHaMP watershed. We then used that model to predict capacity at every master sample point that was not a CHaMP site. To roll up capacity estimates to larger spatial scales, the average predicted capacity of

Table 1: Correlation coefficient between capacity estimates at the population scale using each method.

Species	model_choice	r
Chinook	CHaMP	0.709
Chinook	DASH	0.851
Chinook	Redds	0.904
Steelhead	CHaMP	0.842
Steelhead	DASH	0.987
Steelhead	Redds	0.991

master sample points along a stream was multiplied by the length of that stream, and then combinations of streams could be added together to generate overall capacity estimates for a watershed.

Line Network

We adapted this method to using a stream layer created by Morgan Bond and Tyler Nodine at the Northwest Fisheries Science Center. This layer consisted of a line file divided into 200m reaches with various attributes attached to each reach. The line file is based on the National Hydrography Dataset High Resolution (NHDPlus HR) dataset, which has a higher resolution, 1:24,000, compared to the older layer that the master sample points were chosen from.

We determined which reach was closest to each CHaMP site, and then followed a similar process as described above to model the log of predicted capacity using the attributes attached to each 200m reach as covariates.

Capacity Comparisons

We computed the total capacity of each species in each population using both methods, for summer juveniles (using both CHaMP and DASH habitat metrics) and redds, and compared them. A handful of populations with extremely inflated estimates were left out. The correlations between the two estimates are shown in Table 1.

We plotted one estimate against the other in Figure 1, and showed the relative difference in Figure 2.

Maps

This shows the difference in how the results can be visualized.

Discussion

For both species, across all three QRF models, the two extrapolation models resulted in very similar estimates of total capacity at the population scale. Although the master sample point method has been used for several years, there is no reason to believe estimates from that method are inherently superior to using a line network, so even in the cases when the two models result in different estimates of capacity, it is difficult to say which is "better". On the other hand, there are several reasons to support using the line network method, apart from the actual results.

References

Larsen, D., C. Volk, D. Stevens Jr, A. Olsen, and C. Jordan. 2016. An overview of the Columbia Habitat Monitoring Program's (CHaMP) spatial-temporal design framework. South Fork Research.

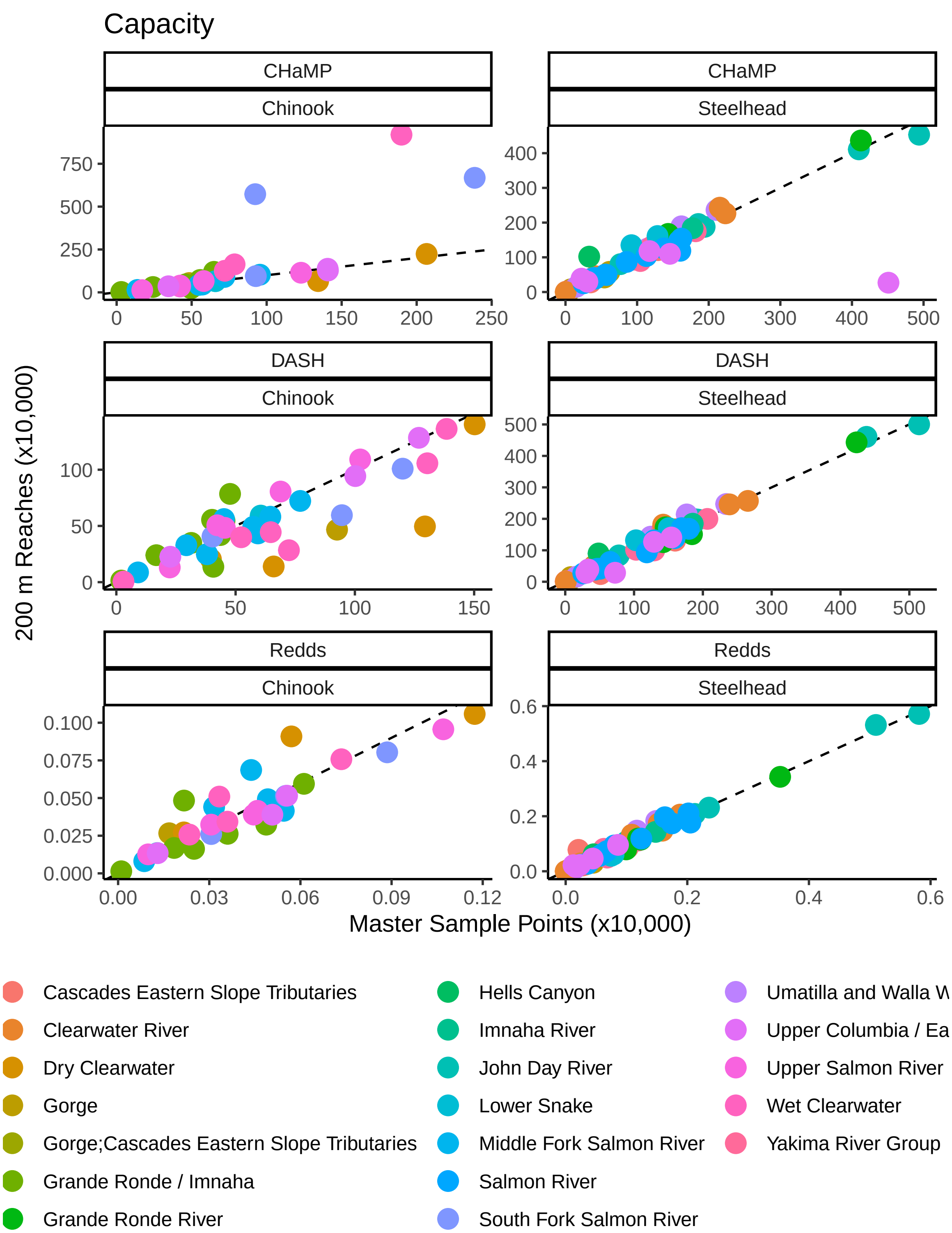

Figure 1: Capacity estimates for each population, calculated with the master sample points method on the x-axis and the line network on the y-axis.

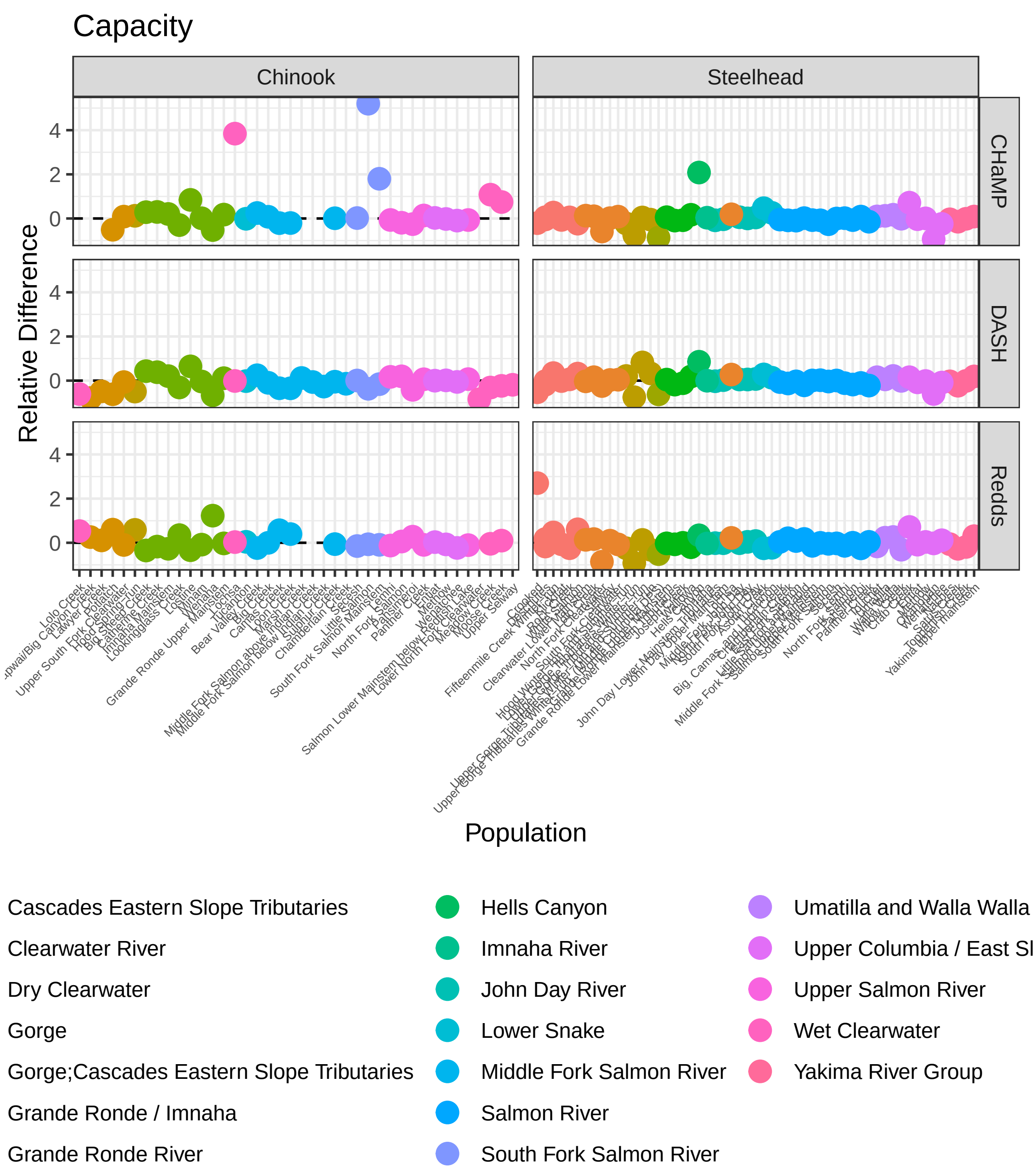

Figure 2: Relative difference between the capacity estimates for each population, using the master sample points method as the reference.

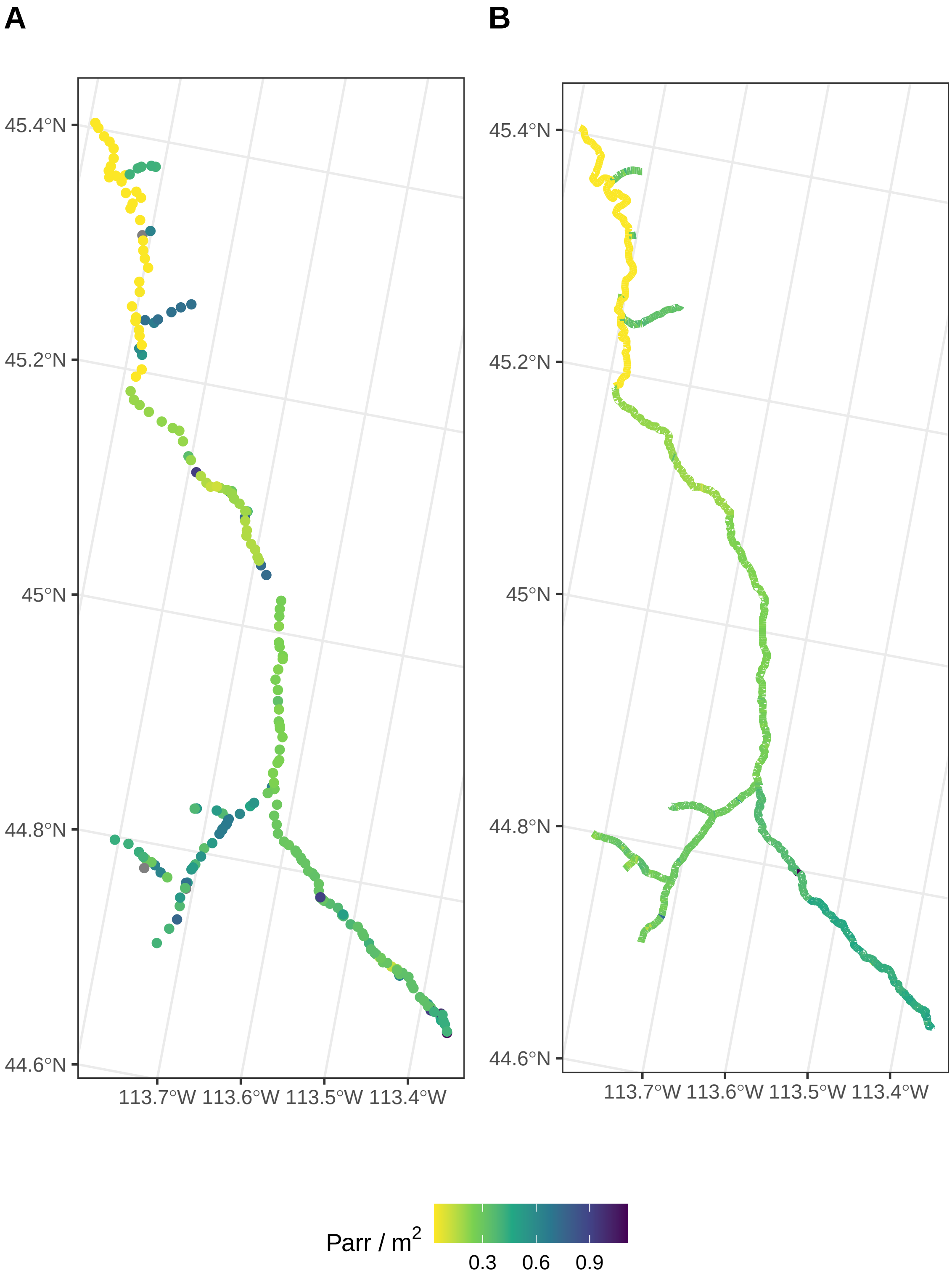

Figure 3: Plots of Chinook parr capacity in the Lemhi, using the master sample points method (A) and the 200 m reach method (B).

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

R Core Team. 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.