Estimates of Wenatchee Steelhead Spawners

Spawn Years 1987-2021

Kevin See[1,*]

January 03, 2023

Contents

[1] Washington Department of Fish & Wildlife

[*] Correspondence: Kevin See <Kevin.See@dfw.wa.gov>

1 Goal

The current method of estimating spawners in the Wenatchee subbasin involves using a PIT-tag based escapement model (DABOM) to estimate tributary spawners (Waterhouse et al. 2020) and adjust the observed redd counts in the mainstem Wenatchee from two observers with a redd observer error model (Murdoch et al. 2018). These adjusted redd counts are combined with redd counts in tributaries below the PIT tag arrays. The PIT tags observed moving into the mainstem (or the tributaries) are used to calculate a fish / redd estimate (males/females + 1 (Murdoch et al. 2009)) and the proportion of hatchery fish on the spawning grounds (pHOS), both of which are used to translate estimates of redds into estimates of hatchery and natural origin spawners. This method has been utilized from spawn year 2014 until the present.

From 2011-2013, the exact same methods were used, except redd counts were adjusted for observer error estimated using the one-observer net error model from Murdoch et al. (2018), because redd surveys in the Wenatchee during that time used a one-observer methodology.

From 2004-2010, estimates of spawners come mainly from redd surveys, which are adjusted using the one-observer net error model from Murdoch et al. (2018). Estimates of fish / redd and pHOS come from fish sampled at Dryden dam. There were three tributaries (Mission, Chumstick and Chiwaukum) that were not part of the redd sampling frame. However, when PIT tag arrays were placed in those tributaries after 2011, some steelhead spawning was observed. Therefore, for 2004-2010, we expanded the estimate of hatchery and natural origin spawners by the mean proportion of overall Wenatchee spawners in those tributaries from 2011 on.

This results in a complete time series from 2004-2021 of estimates of hatchery and natural origin spawners, with associated standard errors. We believe these estimates to be unbiased, based on Murdoch et al. (2018) and Waterhouse et al. (2020).

There is another time series of estimates, from 1987 - 2021, using older methods based on dam counts at the mainstem dams on the Upper Columbia. The goal of this work is to establish a relationship between the two time-series, and use that relationship to "adjust" the older time-series, from 1987-2003, to better match the more recent time-series.
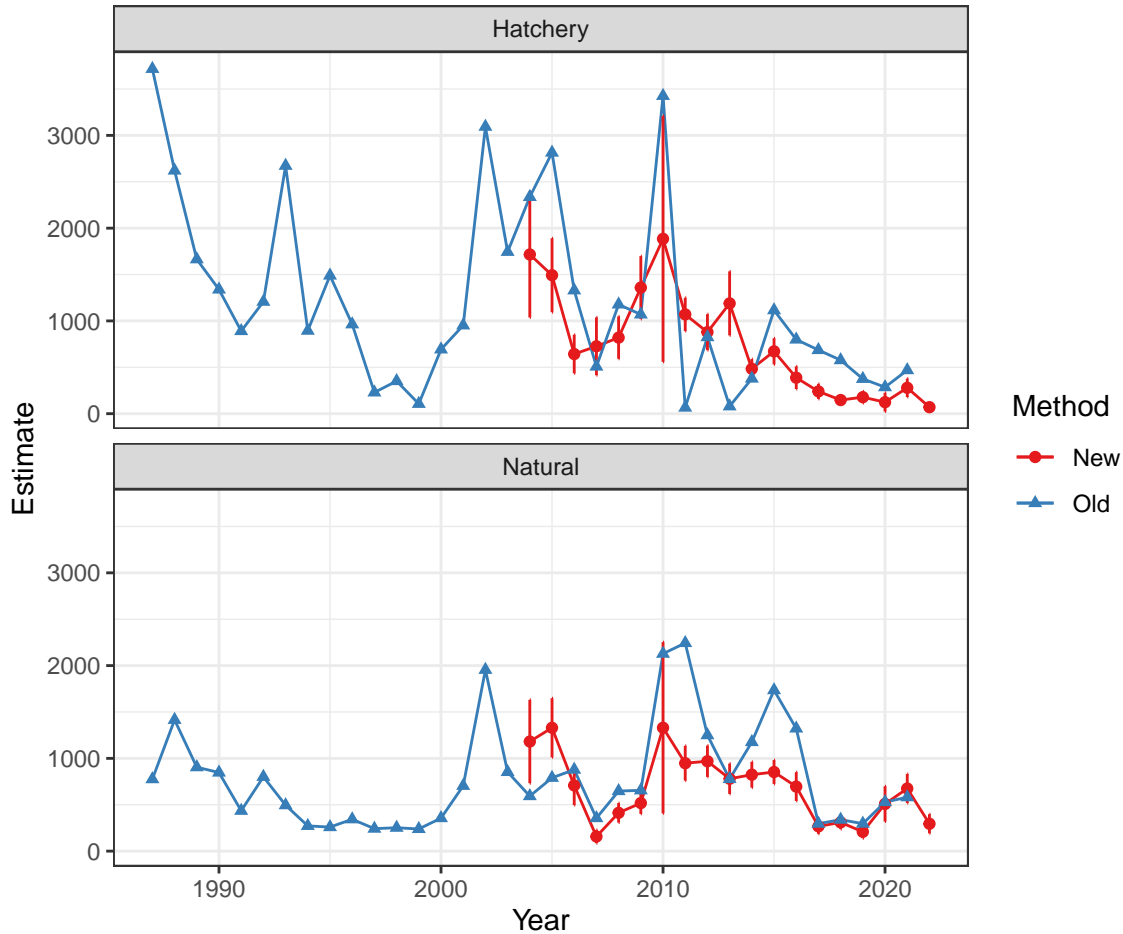

Figure 1: Time-series of hatchery and natural origin spawners in the Wenatchee, colored by what method was used. Error bars represent 95% confidence intervals where available.

2 Methods and Results

2.1 Linear Model

Treat each year as independent, and fit a linear model that includes interactions with origin for both the intercept and slope, with the new estimates as the independent variable and old estimates as the dependent variable. We also tested a log-log linear regression, which involved taking the natural logarithm of each time-series before fitting a linear regression.
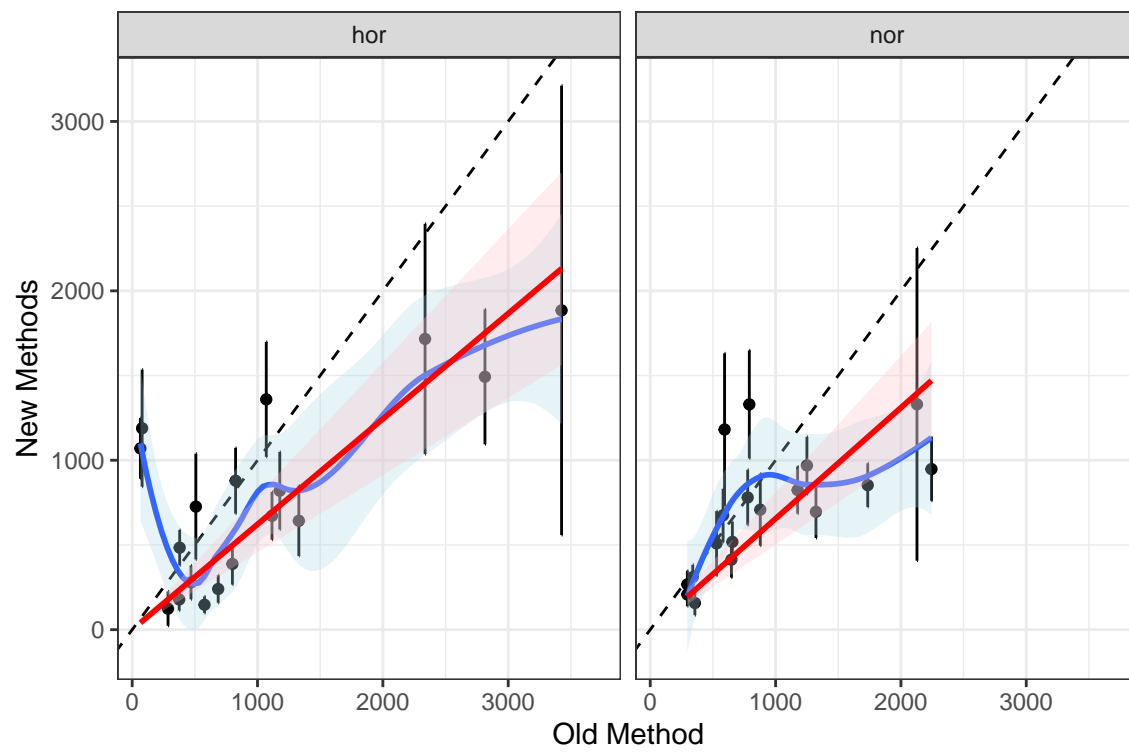
Figure 2: Scatterplots of hatchery and natural origin spawners in the Wenatachee, as estimated by the old method (x-axis) and new methods (y-axis). The blue line is a loess fit, and the red line shows a linear fit forced through the origin.

Figure 3: Log-log scatterplots of hatchery and natural origin spawners in the Wenatachee, as estimated by the old method (x-axis) and new methods (y-axis). The blue line is a loess fit, and the red line shows a linear fit forced through the origin.

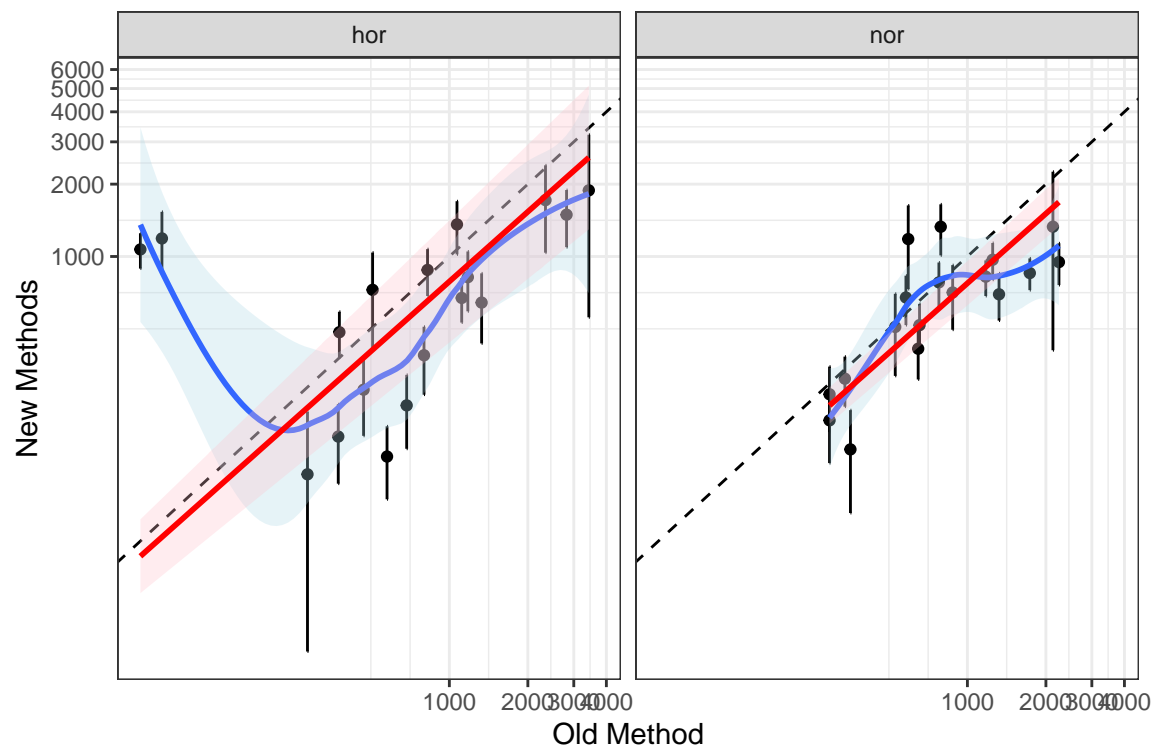Figure 4: Black lines show linear regression estimates with the 95% confidence intervals depicted as grey ribbons. Blue triangles depict estimates from the old time-series, while red points and 95% confidence intervals are from the new time-series.
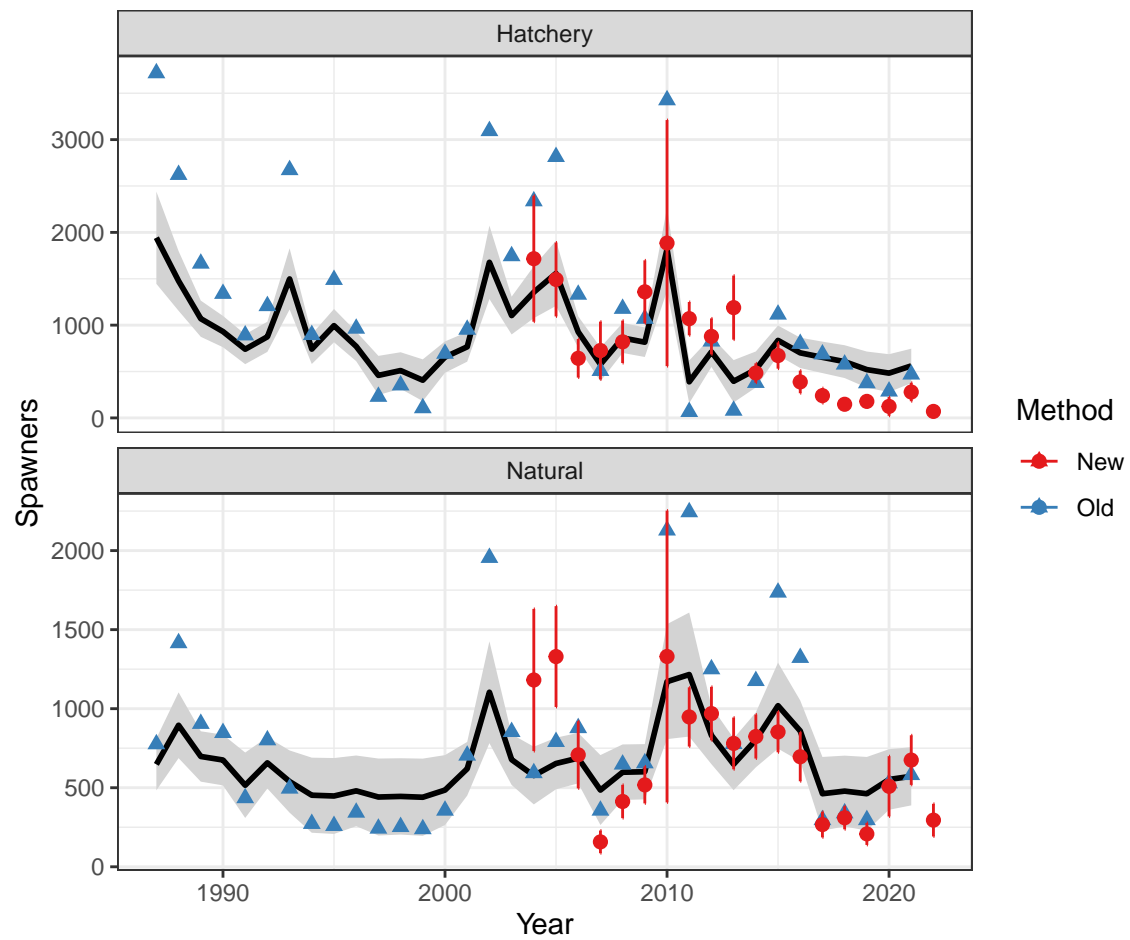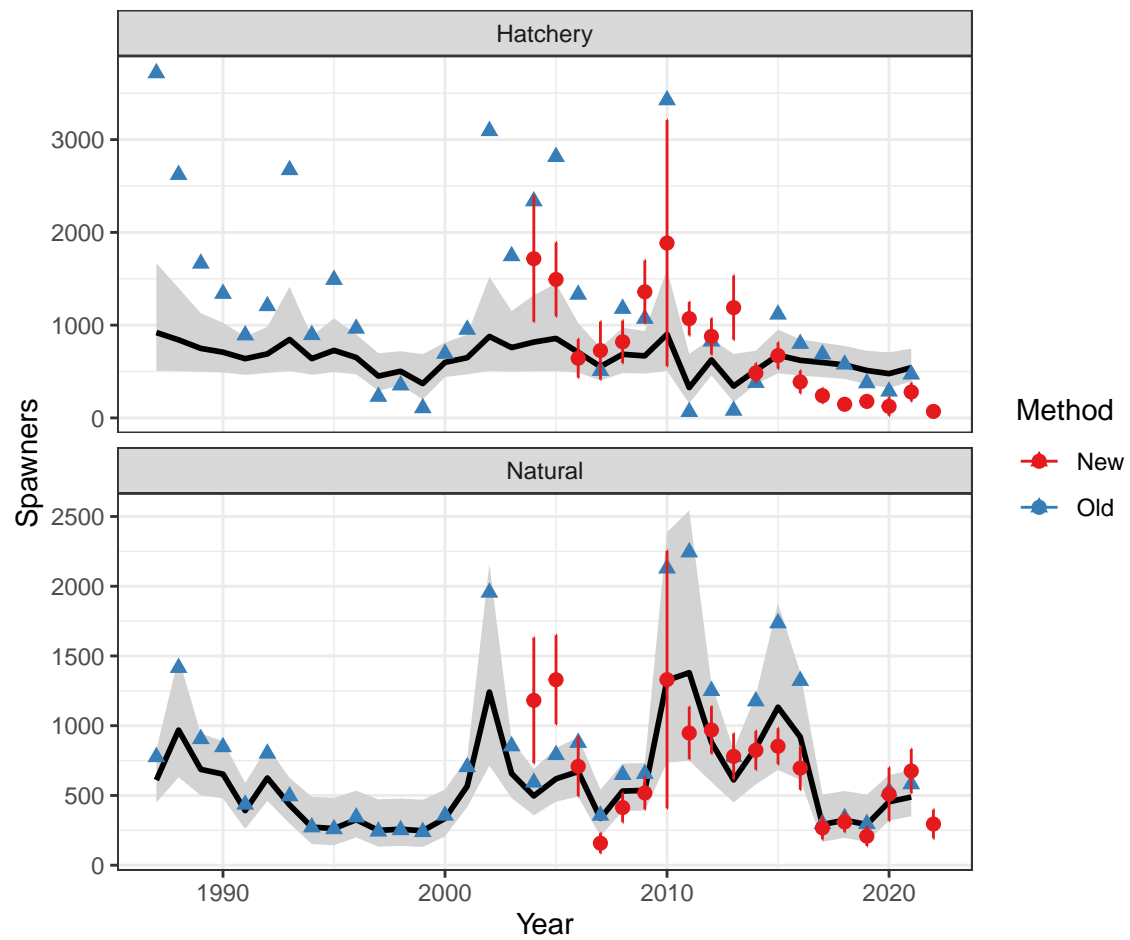
Figure 5: Black lines show log-log linear regression estimates with the 95% confidence intervals depicted as grey ribbons. Blue triangles depict estimates from the old time-series, while red points and 95% confidence intervals are from the new time-series.

2.1.1 Linear Modeling Results

Neither a linear nor a log-log linear model fit the data very well (Figures 2 and 3). A linear fit to these scatter plots would imply a consistent bias (either additive or multiplicative). The lack of such an obvious fit implies the relationship between the two time-series is more complicated. Both appeared to underestimate abundance during years when the older method predicted high numbers steelhead spawners (Figures 4 and 5).

2.2 MARSS

Fit a multivariate auto-regressive state-space (MARSS) model (Holmes et al. 2012, 2021) to the two time-series, ensuring that the only offset of the true states is for the old time-series and that the observation error of the new time-series is informed by mean standard error from the new time-series.

A MARSS model is of the form:

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{u} + \mathbf{C}_t\mathbf{c}_t + \mathbf{w}_t, \text{where } \mathbf{w}_t \sim MVN(0, \mathbf{Q})$$
$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}_t\mathbf{d}_t + \mathbf{v}_t, \text{where } \mathbf{v}_t \sim MVN(0, \mathbf{R})$$

where $\mathbf{x}_t$ represents the true state at time t, which change as a correlated random walk through time. The $\mathbf{u}$ term represents average drift or trend through time. Meanwhile, $\mathbf{y}_t$ represent the observations of those true states, $\mathbf{x}_t$. Which state each element of $\mathbf{y}_t$ is an observation of is determined by the $\mathbf{Z}$ matrix, while $\mathbf{a}$ represents a fixed offset between different elements of $\mathbf{y}$. $\mathbf{C}_t$ and $\mathbf{D}_t$ are possible parameters that show how inputs $\mathbf{c}_t$ and $\mathbf{d}_t$ influence the states ($\mathbf{x}_t$) or observations ($\mathbf{y}_t$); in other words they are covariates. Finally $\mathbf{Q}$ is the process error variance, while $\mathbf{R}$ is the observation error covariance matrix. This framework works best in log-space, so we log-transformed $\mathbf{y}_t$. Further details of MARSS models can be found in the MARSS user guide.

- We set $\mathbf{y}_{1,t}$ and $\mathbf{y}_{3,t}$ to be the estimates of hatchery and wild spawners using the most updated methods, while $\mathbf{y}_{2,t}$ and $\mathbf{y}_{4,t}$ are the vector of estimates of hatchery and wild spawners using the older method.

- We fixed the first and third element of $\mathbf{a}$ to be 0, to ensure there was no offset between the updated estimates and the MARSS model states (The second and fourth element of $\mathbf{a}$ was estimated, as the average multiplicative offset between the older time-series and the true states).

- We set $\mathbf{B}$ to be the identity matrix

- We tested setting $\mathbf{u}$ to 0, the equivalent of a random walk model, and allowing it be estimated, the equivalent of a random walk with drift or trend model.

- The other element we wanted to feed *a priori* into the MARSS framework was the observation error variance, based on the estimated standard errors in the updated estimates. Because the model is set in log-space, we transformed the estimated standard errors by calculating the coefficient of variation, adding 1, logging that value and then calculating the square root. We then took the mean of the log-space standard errors before squaring it. These two values for hatchery and wild observation error were set as the first and third term along the diagonal of the $\mathbf{R}$ matrix, while the off-diagonals were set to 0 and the observation variance of the older methods was left for the MARSS model to estimate.

- Because hatchery and natural origin returns may be correlated to other dam counts, we compiled time-series of counts from several other Columbia River dams: Bonneville, Ice Harbor, McNary, Prosser and Rock Island dams. These were treated as separate states in the MARSS framework, each with a single observation. For all dams, counts were summed from June 1 the year prior to May 31 of that spawn year. These counts are plotted in Figure 6.

Table 1: AICc values for all models.

Model Num.	Description	n Params	LogLik	AICc	delta AICc
3	Q unconstrained, no covariates	44	-24.9	154.2	0.0
6	Q unconstrained, smolt covariate	45	-24.9	156.9	2.7
9	Q unconstrained, no covariates, U unequal	51	-17.0	158.5	4.3
12	Q unconstrained, smolt covariate, U unequal	52	-16.9	161.3	7.1
2	Q mostly independent, no covariates	24	-189.1	430.8	276.6
5	Q mostly independent, smolt covariate	25	-189.1	433.1	278.9
8	Q mostly independent, no covariates, U unequal	31	-183.6	436.9	282.7
11	Q mostly independent, smolt covariate, U unequal	32	-182.5	437.3	283.1
1	Q diag and unequal, no covariates	23	-194.6	439.5	285.3
4	Q diag and unequal, smolt covariate	24	-194.6	441.8	287.7
7	Q diag and unequal, no covariates, U unequal	30	-191.9	451.1	296.9
10	Q diag and unequal, smolt covariate, U unequal	31	-191.9	453.5	299.3

- We also compiled one more possible input, hatchery releases of smolts. We hypothesized that the hatchery release numbers from previous years might inform the predicted returns of adults. We used the weighted average of salt age 1 and salt age 2 releases, weighted 70% towards salt age 1 and 30% towards salt age 2 based on average age composition data. Salt age 1 fish returned 2 years after their release, while salt age 2 fish returned after 3 years. This time-series extended back to 1987 and was normalized to have a mean of zero and standard deviation of one. This was treated as a possible covariate for the estimated state of hatchery spawners. This time series is shown in Figure 7.

- We tested several configurations of this model:

    1. Treated all states (Wenatchee hatchery and wild spawners, and other dam counts) as independent, by setting the off-diagonal terms of **Q** to 0. (`Q = "diagonal and unequal`)
    2. Similar to (1), but allowed for the process errors of Wenatchee hatchery and wild spawners to co-vary by estimating a single off-diagonal element of **Q**.
    3. Allowed the process errors to co-vary across all states, and estimated their covariance as the off-diagonal term of **Q**. (`Q = "unconstrained`)
    4. Same as (1), but included a covariate of hatchery smolt releases to inform hatchery returns.
    5. Same as (2), but included a covariate of hatchery smolt releases to inform hatchery returns.
    6. Same as (3), but included a covariate of hatchery smolt releases to inform hatchery returns. 7-12. Same as above, but included a possible trend (`U = "unequal"`).

Models 1, 2, 4 and 5 essentially ignore the dam counts when it comes to fitting and predicting for the Wenatchee states. Models 1 and 4 treat hatchery and wild spawners as independent time-series which is the equivalent of fitting separate models for wild and hatchery spawners.

- All models were compared with AICc.

- All models were fit using the `MARSS` package in R.

2.2.1 MARSS Results

The results (Table 1) show the third model to be best supported by the data. This model allows for correlated process errors between hatchery and natural spawners and various dam counts. The second best model by AICc was model 6, which was identical to model 3 but also included a covariate of previous smolt releases to help predict hatchery spawners.
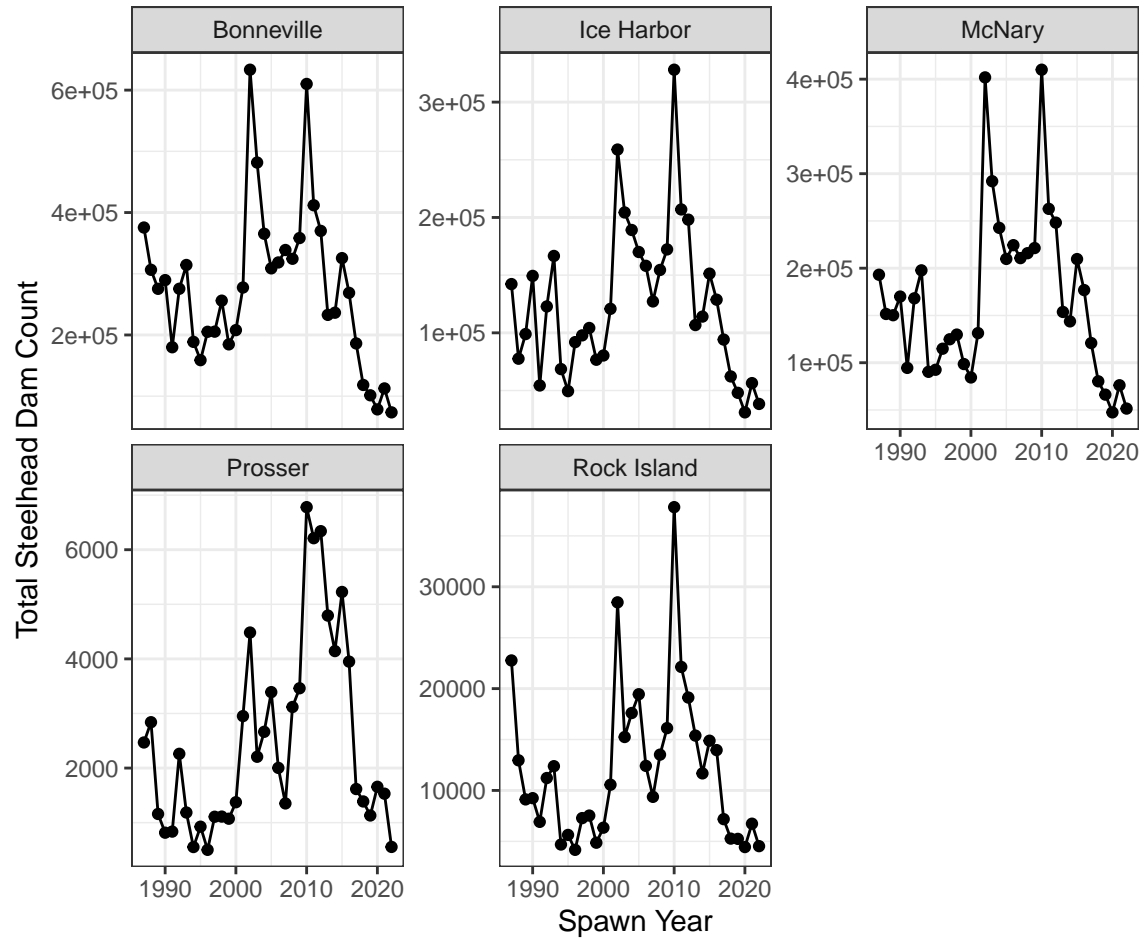
Figure 6: Time-series of counts from various Columbia River dams, from June 1 the year prior to May 31 of that spawn year.

Table 2: Estimates of Q matrix from model 3, showing variance and co-variance estimates.

	Wen. Hatch	Wen. Wild	BON	IHR	MCN	PRO	RIS
Wen. Hatch	0.233	0.188	0.142	0.161	0.163	0.160	0.183
Wen. Wild	0.188	0.217	0.097	0.105	0.111	0.209	0.174
BON	0.142	0.097	0.098	0.115	0.114	0.084	0.110
IHR	0.161	0.105	0.115	0.138	0.136	0.085	0.125
MCN	0.163	0.111	0.114	0.136	0.136	0.089	0.128
PRO	0.160	0.209	0.084	0.085	0.089	0.231	0.162
RIS	0.183	0.174	0.110	0.125	0.128	0.162	0.159

Table 3: Estimates of selected parameters from the best model.

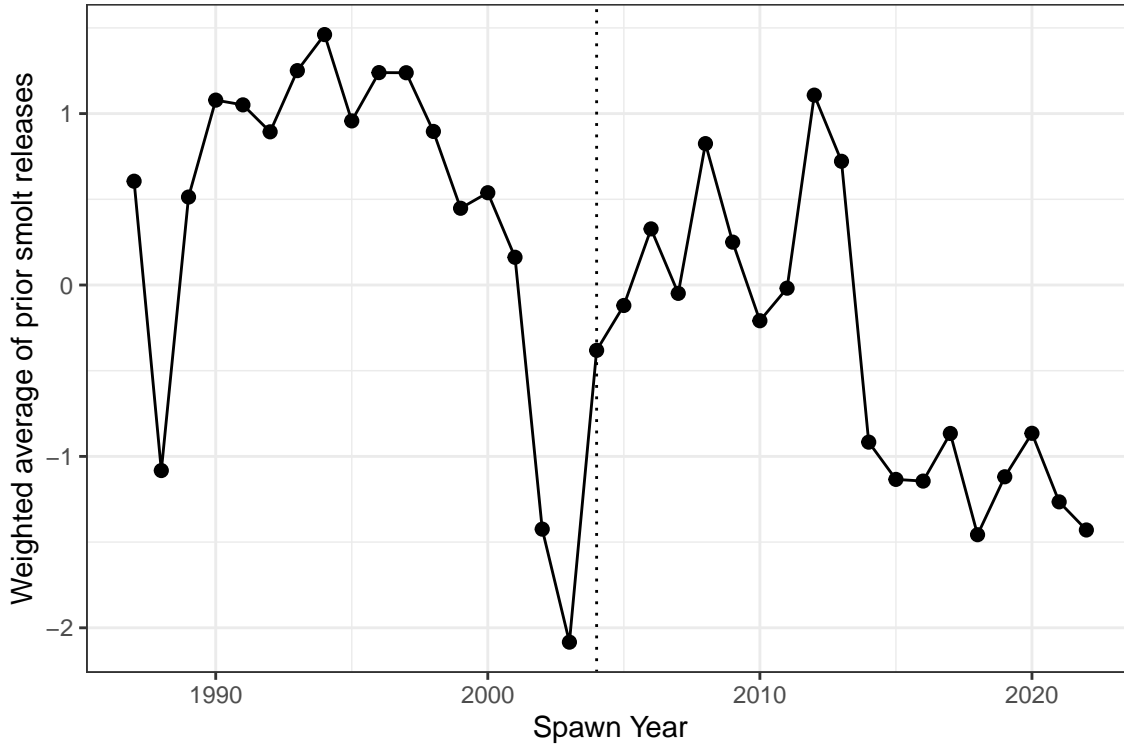| term	estimate	std.error	conf.low	conf.up
A.a__old__hor	0.135	0.195	-0.247	0.516
A.a__old__nor	0.247	0.106	0.039	0.456
R.r__old__hor	0.673	0.166	0.348	0.998
R.r__old__nor	0.086	0.025	0.037	0.136

Figure 7: Time-series of normalized weighted average of smolt releases prior to the spawn year (x-axis). Dotted line shows when the new time-series begins.

Table 2 shows the estimates of the process error covariance matrix, **Q**. Table 3 shows other parameter estimates from the selected model.

Figure 8 compares the predictions of hatchery spawners from a model that does not use smolt releases as a covariate and one that does, although both have unconstrained **Q** matrices. (models 2 and 4). Predictions are greater for the model with a smolt release covariate, but only in the earlier years.

3 Conclusions

The MARSS framework appear to fit the data better than the linear regression for several reasons, so we chose to use that. First, there does not appear to be a consistent additive or multiplicative bias between the two time-series. Second, a MARSS model is explicitly a time-series model, which is appropriate for this comparison. Finally, the MARSS framework allowed us to test a variety of model structures, including bringing in other time-series and covariates. AICc supported a model that included several time-series of various dam counts, with correlated process errors (true year-to-year variability), including a positive correlation between hatchery and natural origin spawners. This positive correlation could reflect the impacts of shared ocean conditions. There was slightly less support for the same model that also included a covariate of weighted average of previous smolt releases to for the hatchery returns (but not natural origin returns). Because the coefficient of that covariate was negative, with confidence intervals that overlapped 0, and because including smolt releases had very little effect on spawner estimates (Figure 8), we decided against using that model and chose the one with the lowest AICc score.

Table 4 shows the updated time-series, including a brief description of the method used to generate the estimate each year.
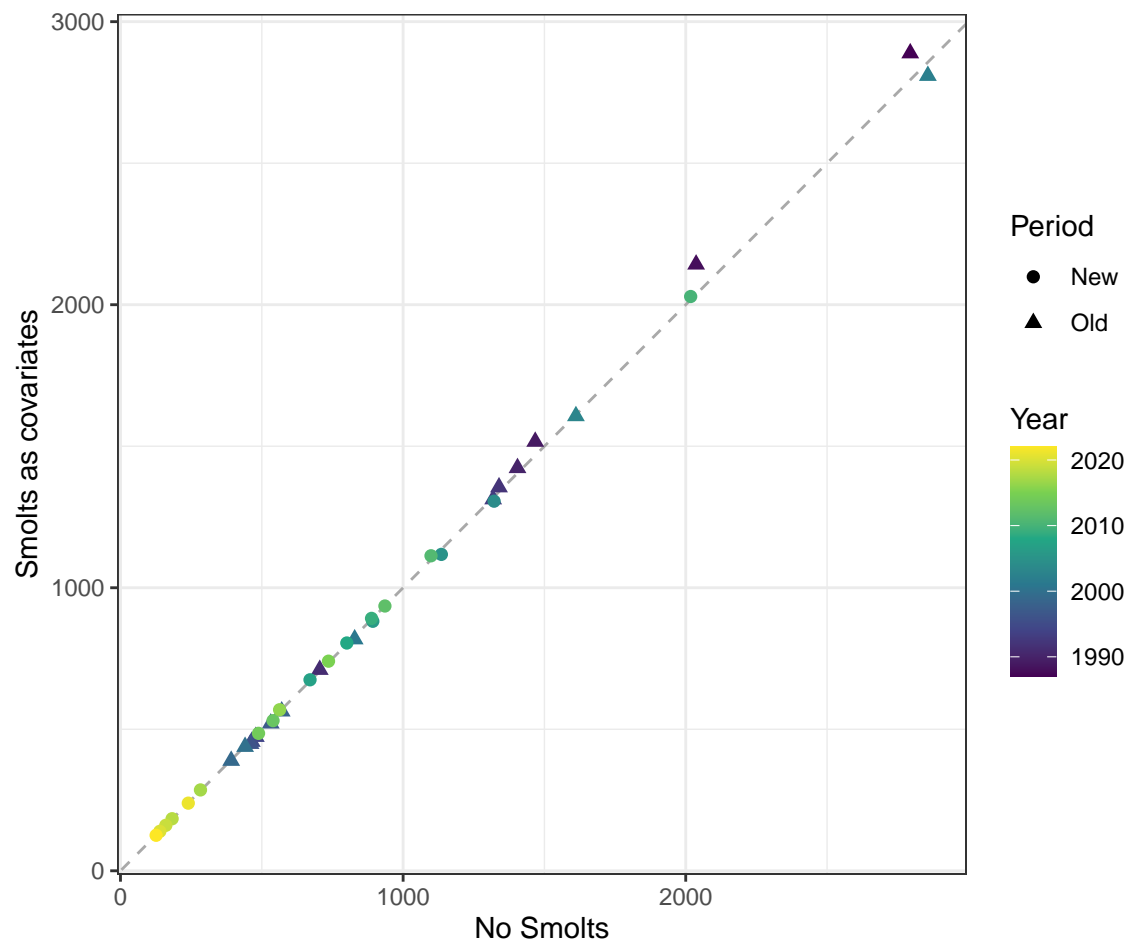
Figure 8: Comparison of predicted states of hatchery spawners for a model with no smolt release covariate (x-axis) and one that includes that covariate (y-axis). The period refers to whether the new time-series estimates exist.
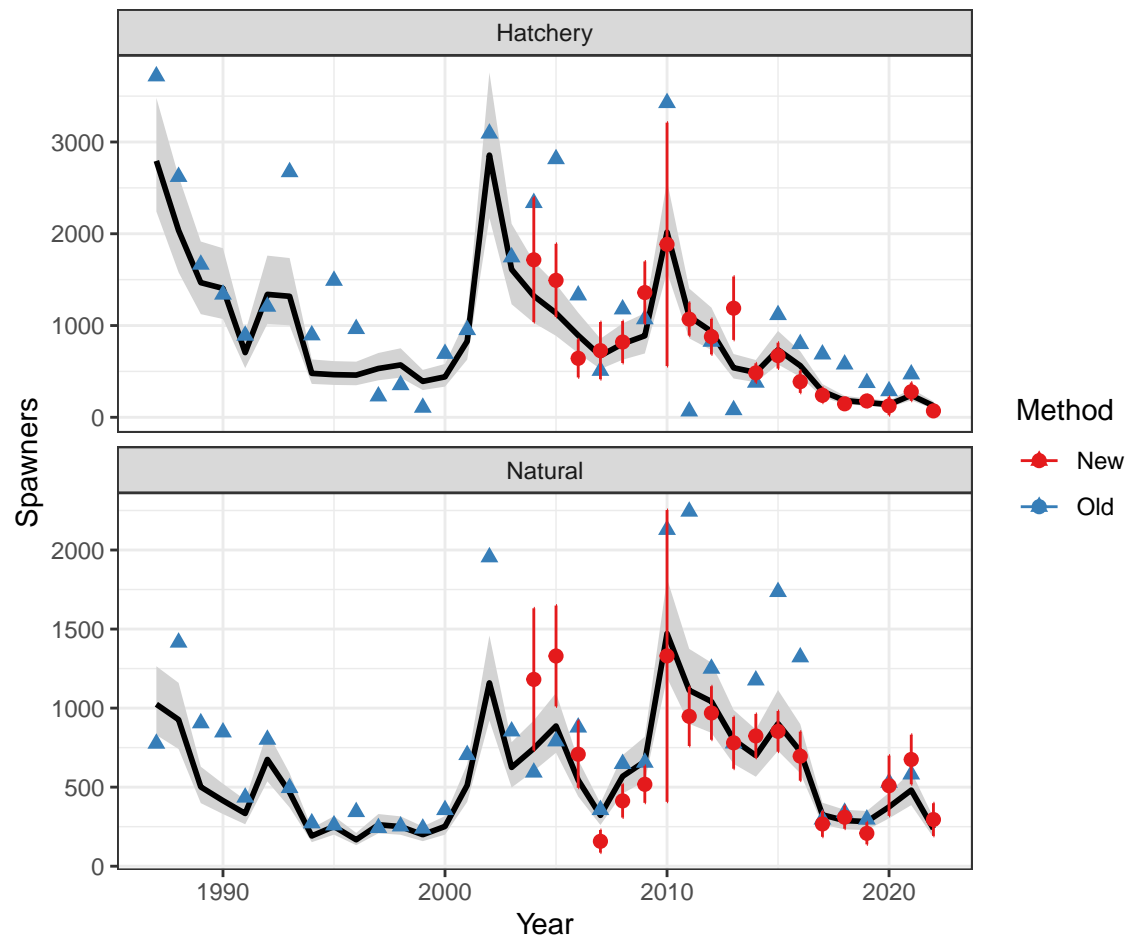
Figure 9: Estimates of spawners through the years, faceted by origin. Predicted spawners is the black line with 95% confidence interval in gray. Blue triangles depict estimates from the old time-series, while red points and 95% confidence intervals are from the new time-series.

Figure 10: Updated estimates of spawners through the years, colored by origin, showing point estimates and 95% confidence intervals. Dashed vertical line differentiates older and newer time-series.
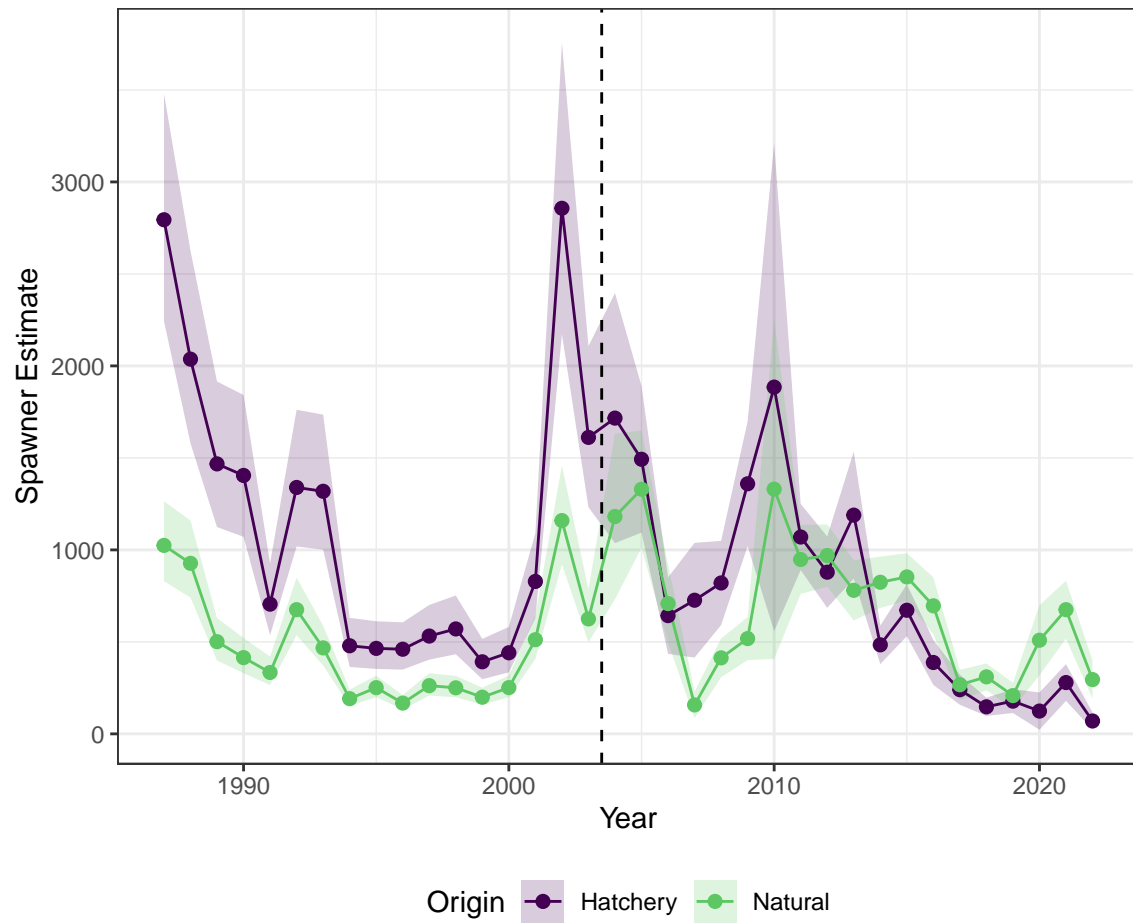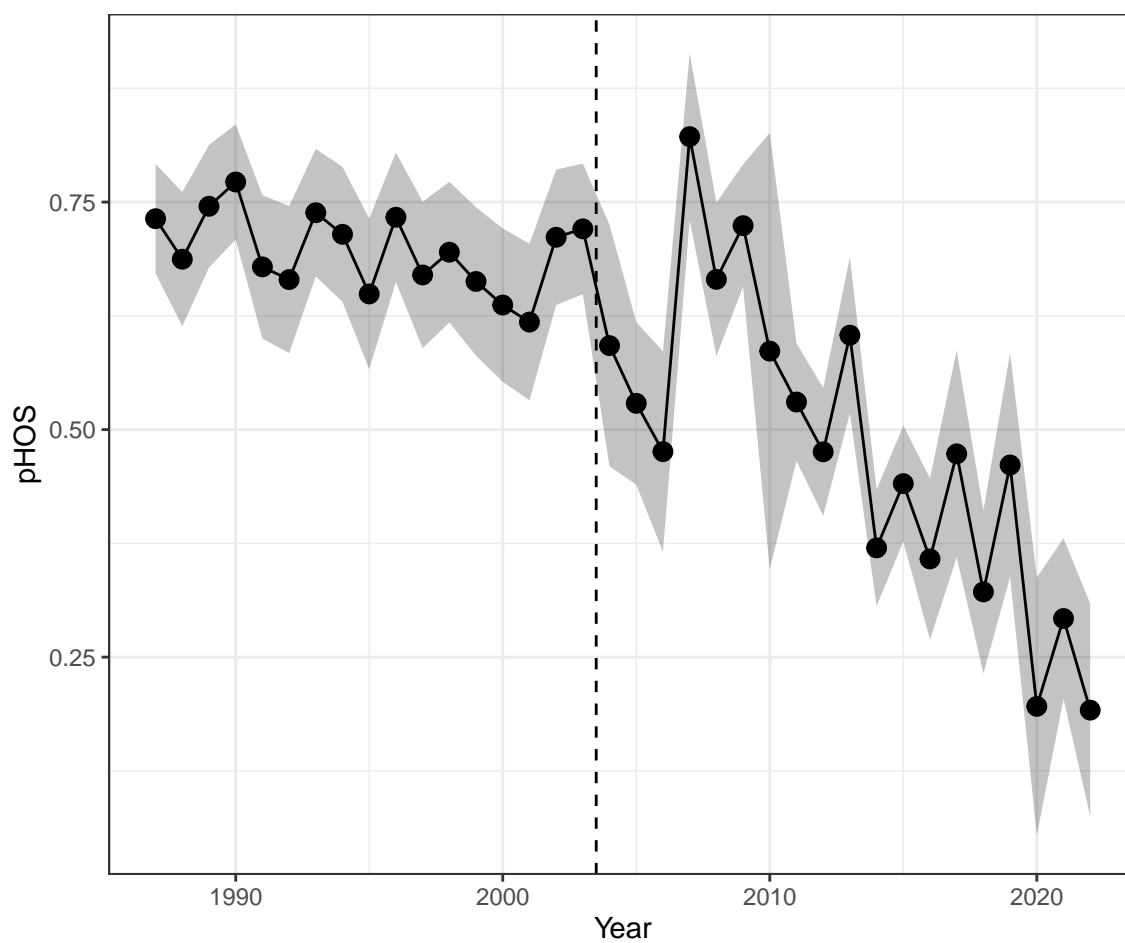
Figure 11: Estimates of pHOS based on the updated time-series, showing 95% confidence intervals. Dashed vertical line differentiates older and newer time-series.

Table 4: Updated time-series of steelhead spawners in the Wenatchee, by origin.

Year	Origin	Method	Estimate	SE	LCI	UCI
1987	Hatchery	MARSS	2,794	316.2	2,243	3,481
1987	Natural	MARSS	1,024	111.0	830	1,264
1988	Hatchery	MARSS	2,037	267.5	1,580	2,626
1988	Natural	MARSS	927	106.9	741	1,160
1989	Hatchery	MARSS	1,467	202.2	1,124	1,915
1989	Natural	MARSS	501	58.8	399	629
1990	Hatchery	MARSS	1,405	197.1	1,071	1,842
1990	Natural	MARSS	415	48.9	330	521
1991	Hatchery	MARSS	705	99.7	537	927
1991	Natural	MARSS	334	39.5	265	420
1992	Hatchery	MARSS	1,339	190.1	1,018	1,762
1992	Natural	MARSS	675	79.9	537	850
1993	Hatchery	MARSS	1,319	187.6	1,002	1,735
1993	Natural	MARSS	467	55.4	371	588
1994	Hatchery	MARSS	479	68.2	364	631
1994	Natural	MARSS	191	22.7	152	241
1995	Hatchery	MARSS	465	66.3	353	612
1995	Natural	MARSS	251	29.8	200	316
1996	Hatchery	MARSS	460	65.7	349	606
1996	Natural	MARSS	167	19.8	133	211
1997	Hatchery	MARSS	532	76.0	403	701
1997	Natural	MARSS	262	31.1	208	330
1998	Hatchery	MARSS	570	81.6	433	752
1998	Natural	MARSS	250	29.7	199	315
1999	Hatchery	MARSS	392	56.1	297	517
1999	Natural	MARSS	199	23.7	158	251
2000	Hatchery	MARSS	441	63.1	334	581
2000	Natural	MARSS	251	29.9	200	317
2001	Hatchery	MARSS	829	118.3	629	1,092
2001	Natural	MARSS	512	60.8	407	644
2002	Hatchery	MARSS	2,857	405.2	2,172	3,757
2002	Natural	MARSS	1,159	137.3	921	1,459
2003	Hatchery	MARSS	1,611	224.0	1,232	2,108
2003	Natural	MARSS	625	73.4	498	785
2004	Hatchery	1 Obs. Model & Expansion	1,716	346.6	1,037	2,396
2004	Natural	1 Obs. Model & Expansion	1,181	229.5	732	1,631
2005	Hatchery	1 Obs. Model & Expansion	1,493	203.3	1,094	1,891
2005	Natural	1 Obs. Model & Expansion	1,330	162.7	1,011	1,649
2006	Hatchery	1 Obs. Model & Expansion	643	106.2	435	851
2006	Natural	1 Obs. Model & Expansion	708	108.4	496	921
2007	Hatchery	1 Obs. Model & Expansion	726	159.1	415	1,038
2007	Natural	1 Obs. Model & Expansion	157	36.7	86	229
2008	Hatchery	1 Obs. Model & Expansion	820	116.8	591	1,049
2008	Natural	1 Obs. Model & Expansion	413	53.8	308	518
2009	Hatchery	1 Obs. Model & Expansion	1,359	173.1	1,020	1,699
2009	Natural	1 Obs. Model & Expansion	518	59.8	401	635
2010	Hatchery	1 Obs. Model & Expansion	1,885	676.7	558	3,211
2010	Natural	1 Obs. Model & Expansion	1,330	470.9	407	2,253
2011	Hatchery	1 Obs. Model & DABOM	1,070	91.7	890	1,250

Table 4: Updated time-series of steelhead spawners in the Wenatchee, by origin. *(continued)*

Year	Origin	Method	Estimate	SE	LCI	UCI
2011	Natural	1 Obs. Model & DABOM	948	95.6	760	1,135
2012	Hatchery	1 Obs. Model & DABOM	879	99.0	685	1,073
2012	Natural	1 Obs. Model & DABOM	969	86.3	800	1,138
2013	Hatchery	1 Obs. Model & DABOM	1,189	176.2	844	1,534
2013	Natural	1 Obs. Model & DABOM	780	83.5	617	944
2014	Hatchery	2 Obs. Model & DABOM	484	53.4	379	589
2014	Natural	2 Obs. Model & DABOM	824	71.1	685	963
2015	Hatchery	2 Obs. Model & DABOM	672	72.0	531	813
2015	Natural	2 Obs. Model & DABOM	853	65.7	724	982
2016	Hatchery	2 Obs. Model & DABOM	388	62.3	266	510
2016	Natural	2 Obs. Model & DABOM	696	79.2	541	851
2017	Hatchery	2 Obs. Model & DABOM	240	41.6	158	322
2017	Natural	2 Obs. Model & DABOM	267	41.4	186	348
2018	Hatchery	2 Obs. Model & DABOM	147	25.2	98	196
2018	Natural	2 Obs. Model & DABOM	310	37.2	237	383
2019	Hatchery	2 Obs. Model & DABOM	178	33.0	113	243
2019	Natural	2 Obs. Model & DABOM	208	35.7	138	278
2020	Hatchery	2 Obs. Model & DABOM	124	51.7	23	225
2020	Natural	2 Obs. Model & DABOM	509	97.2	318	700
2021	Hatchery	2 Obs. Model & DABOM	279	50.9	179	379
2021	Natural	2 Obs. Model & DABOM	675	80.1	518	832
2022	Hatchery	2 Obs. Model & DABOM	70	23.8	23	117
2022	Natural	2 Obs. Model & DABOM	295	52.7	192	398

References

Holmes, E. E., E. J. Ward, M. D. Scheuerell, and K. Wills. 2021. MARSS: Multivariate autoregressive state-space modeling.

Holmes, E. E., E. J. Ward, and K. Wills. 2012. MARSS: Multivariate autoregressive state-space models for analyzing time-series data. The R Journal 4(1):30.

Murdoch, A. R., C. J. Herring, C. H. Frady, K. See, and C. E. Jordan. 2018. Estimating observer error and steelhead redd abundance using a modified Gaussian area-under-the-curve framework. Canadian Journal of Fisheries and Aquatic Sciences 75(12):2149–2158.

Murdoch, A. R., T. N. Pearsons, and T. W. Maitland. 2009. The Number of Redds Constructed per Female Spring Chinook Salmon in the Wenatchee River Basin. North American Journal of Fisheries Management 29(2):441–446.

Waterhouse, L., J. White, K. See, A. Murdoch, and B. X. Semmens. 2020. A Bayesian nested patch occupancy model to estimate steelhead movement and abundance. Ecological Applications 30(8).