Estimates of Wenatchee Steelhead Spawners

Spawn Years 1987-2022

Kevin See1,✉

May 16, 2023

Table of Contents

1 Washington Department of Fish & Wildlife

✉ Correspondence: [Kevin See <[Kevin.See@dfw.wa.gov](mailto:Kevin.See@dfw.wa.gov)>](mailto: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 and a Gaussian area-under-the-curve method, as described in 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 observer error was estimated with 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 or from broodstock collection. 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 2004-2022 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.

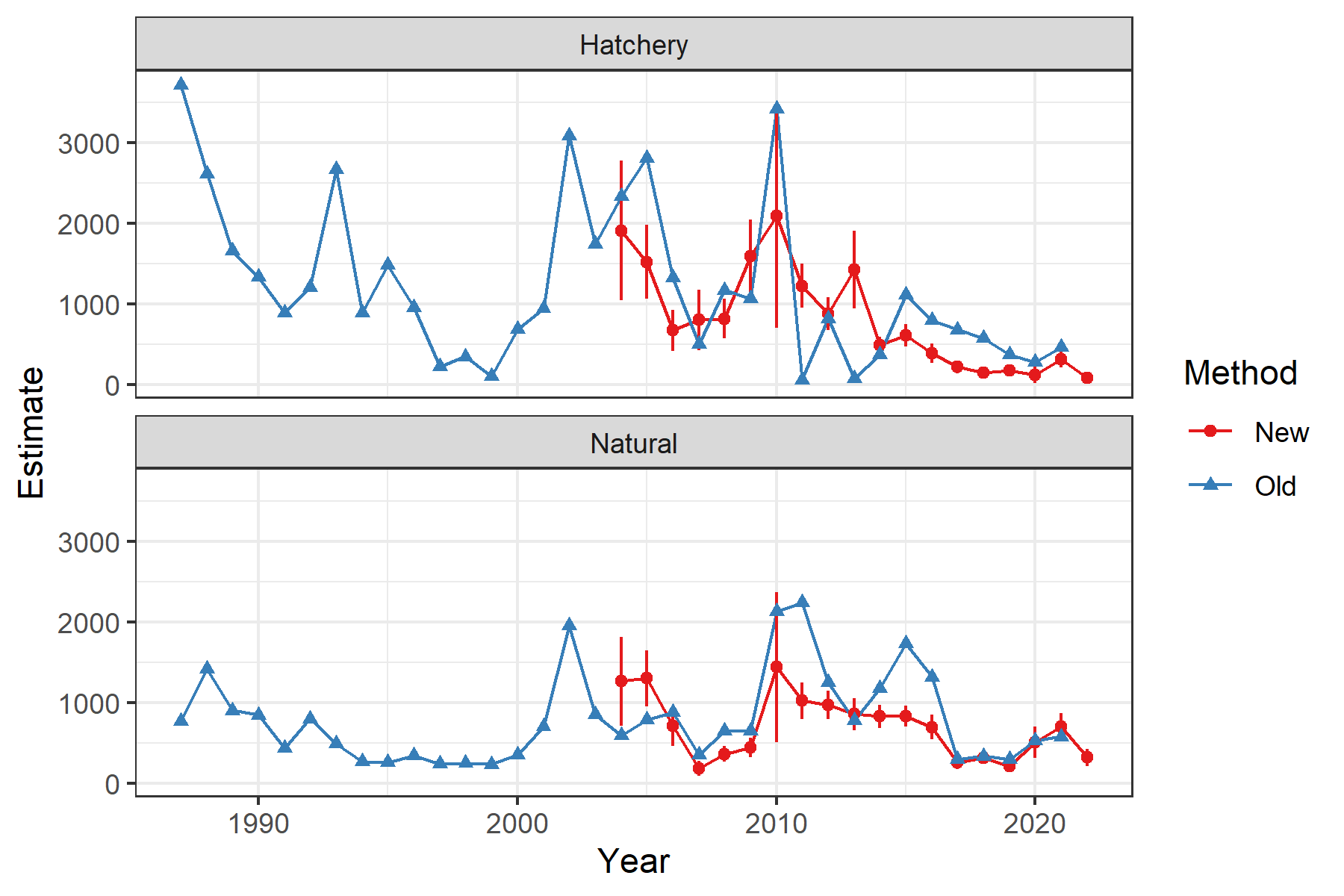


Figure 1.1: Time-series of hatchery and natural origin spawners in the Wenatchee, colored by what method was used. Error bars represent 95% confidence intervals where available.

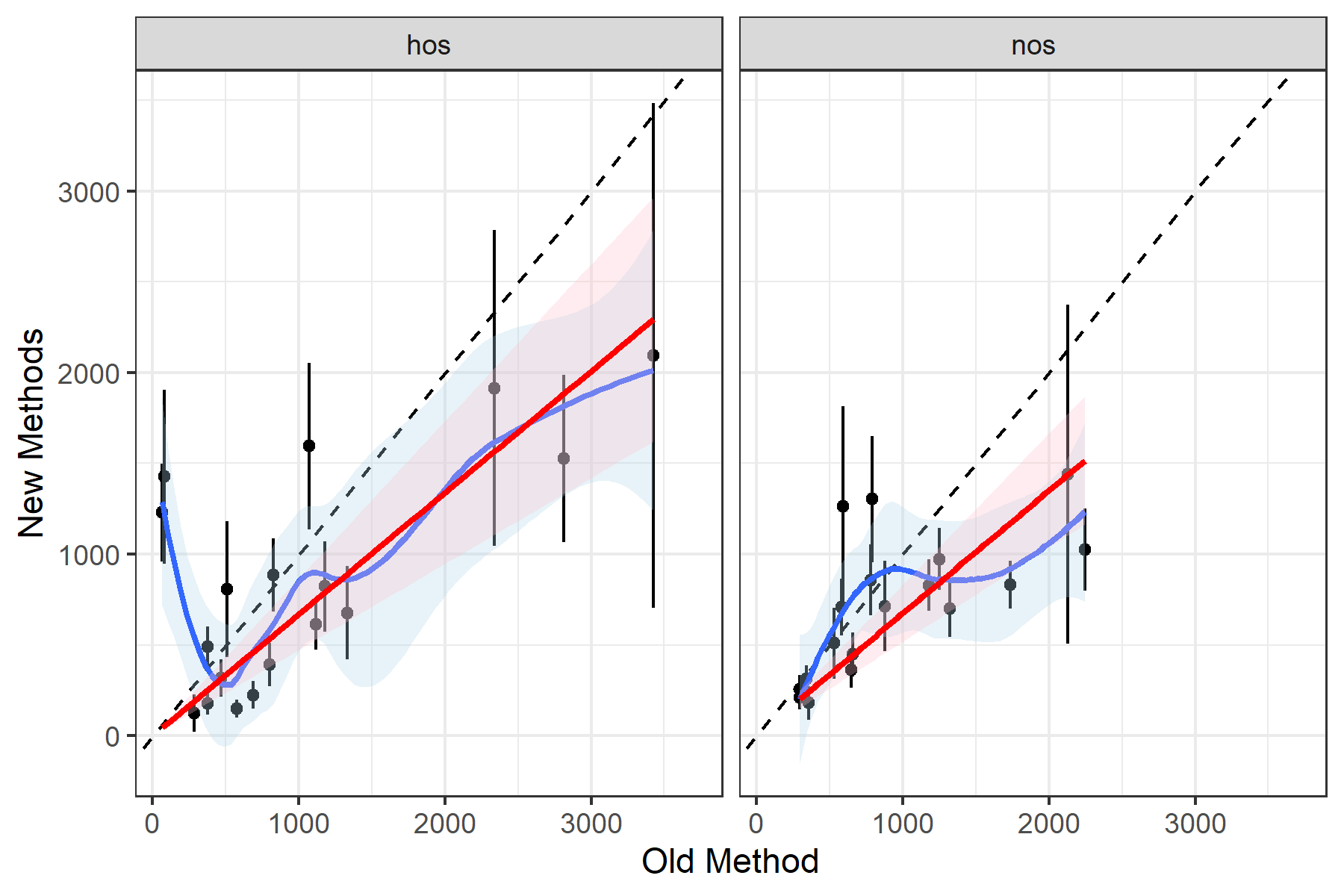


Figure 1.2: Scatterplots of hatchery and natural origin spawners in the Wenatachee, as estimated by the old method (x-axis) and new methods (y-axis). The blue line is a loess fit, and the red line shows a linear fit forced through the origin.

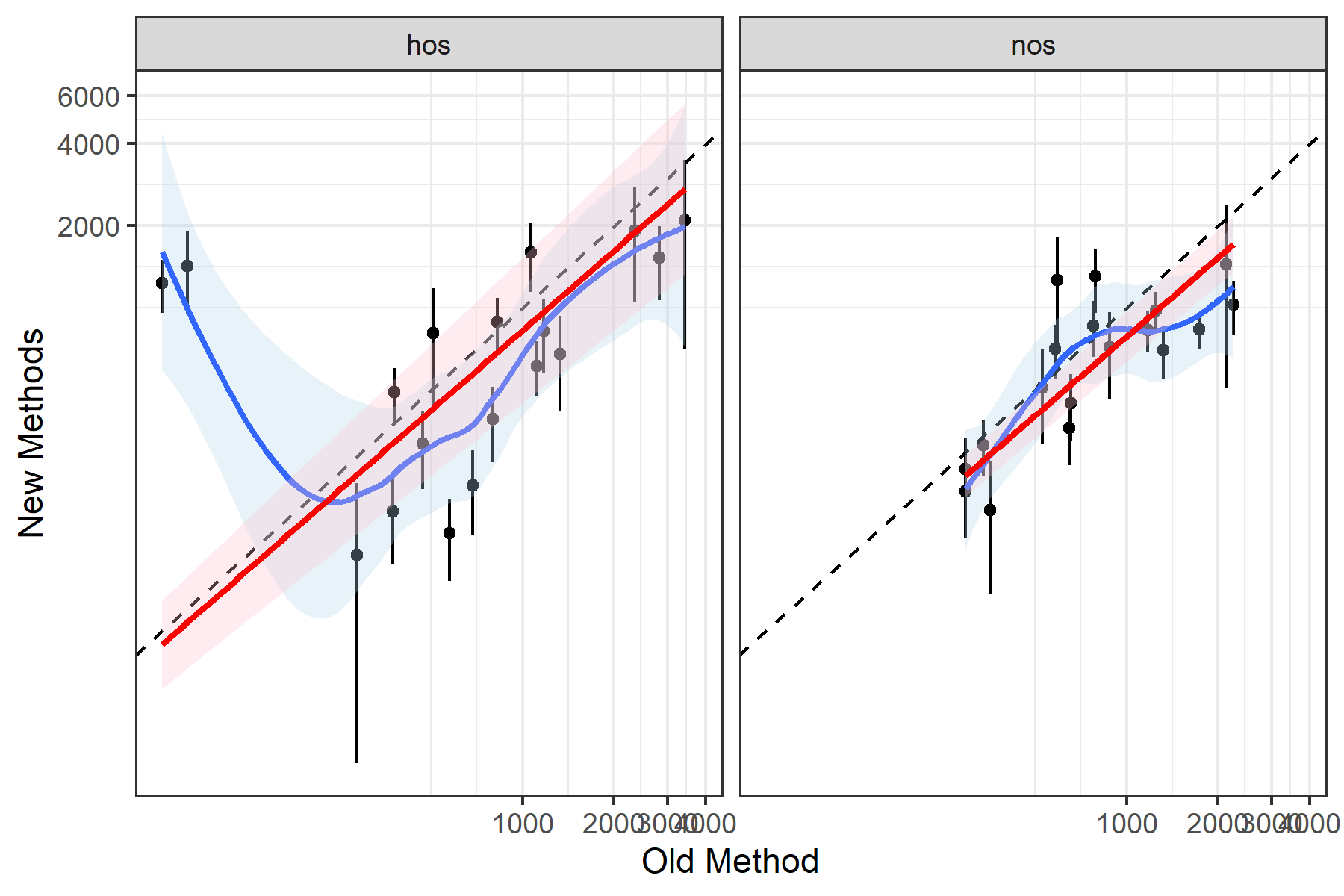


Figure 1.3: Log-log scatterplots of hatchery and natural origin spawners in the Wenatachee, as estimated by the old method (x-axis) and new methods (y-axis). The blue line is a loess fit, and the red line shows a linear fit forced through the origin.

# 2 Methods and Results

## 2.1 Linear Model

Our first approach was to 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.

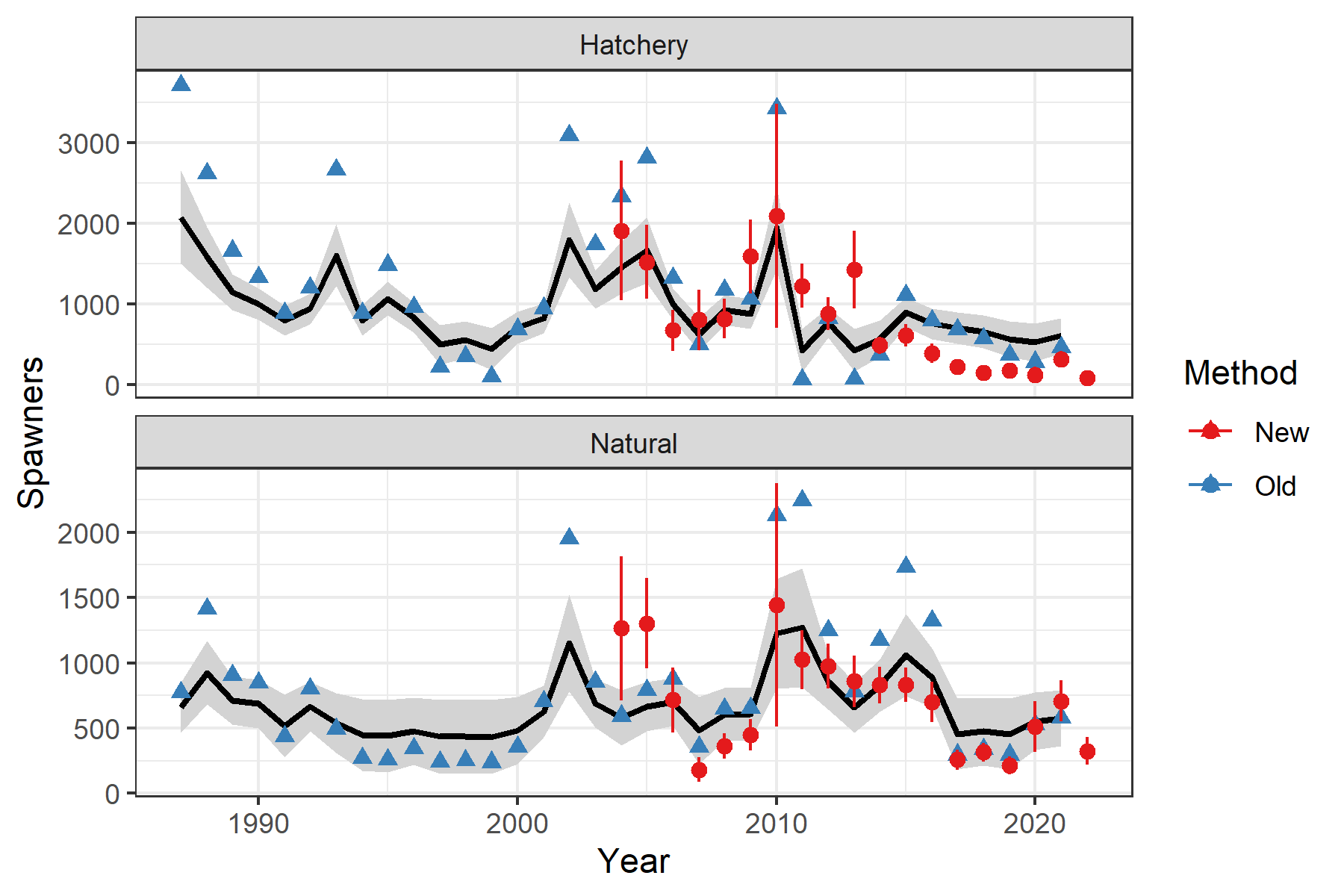


Figure 2.1: Black lines show linear regression estimates with the 95% confidence intervals depicted as grey ribbons. Blue triangles depict estimates from the old time-series, while red points and 95% confidence intervals are from the new time-series.

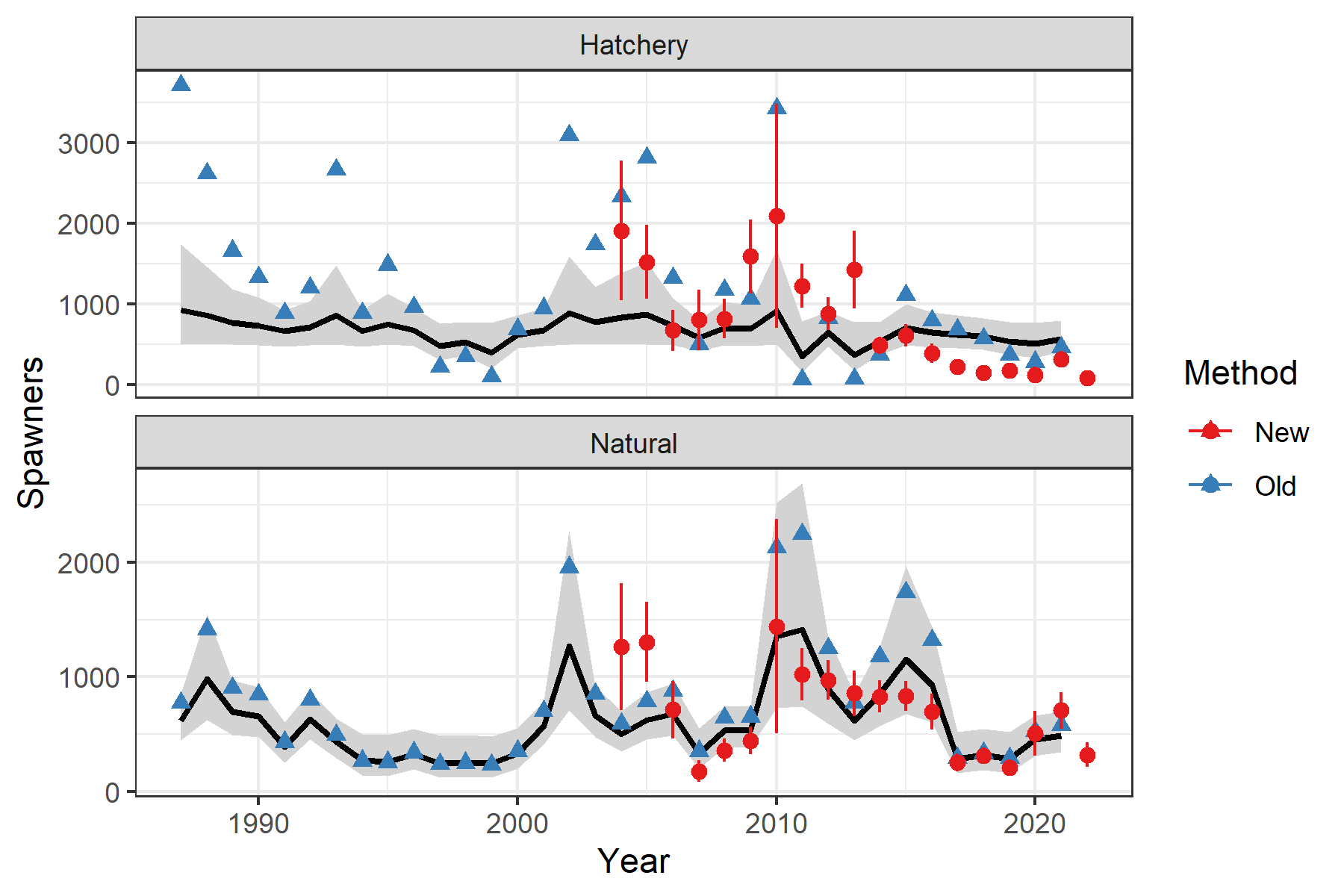


Figure 2.2: Black lines show log-log linear regression estimates with the 95% confidence intervals depicted as grey ribbons. Blue triangles depict estimates from the old time-series, while red points and 95% confidence intervals are from the new time-series.

### 2.1.1 Linear Modeling Results

Neither a linear nor a log-log linear model fit the data very well (Figures 1.2 and 1.3). 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 2.1 and 2.2).

## 2.2 MARSS

Our next approach was to 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:

where represents the true state at time , which change as a correlated random walk through time. The term represents average drift or trend through time. Meanwhile, represent the observations of those true states, . Which state each element of is an observation of is determined by the matrix, while represents a fixed offset between different elements of . and are possible parameters that show how inputs and influence the states () or observations (); in other words they are covariates. Finally is the process error variance, while is the observation error covariance matrix. This framework works best in log-space, so we log-transformed . Further details of MARSS models can be found in the [MARSS user guide](https://cran.r-project.org/web/packages/MARSS/vignettes/UserGuide.pdf).

* We set and to be the estimates of hatchery and wild spawners using the most updated methods, while and are the vector of estimates of hatchery and wild spawners using the older method.
* We fixed the first and third element of to be 0, to ensure there was no offset between the updated estimates and the MARSS model states (The second and fourth element of was estimated, as the average multiplicative offset between the older time-series and the true states).
* We set to be the identity matrix.
* We tested setting 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 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 2.3. The hypothesis here is that other dam counts may inform the $\textbf(Q)$ matrix, allowing for better inference of the states we are interested in.
* 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 2.4.
* 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 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 .
  3. Allowed the process errors to co-vary across all states, and estimated their covariance as the off-diagonal term of . (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. 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.

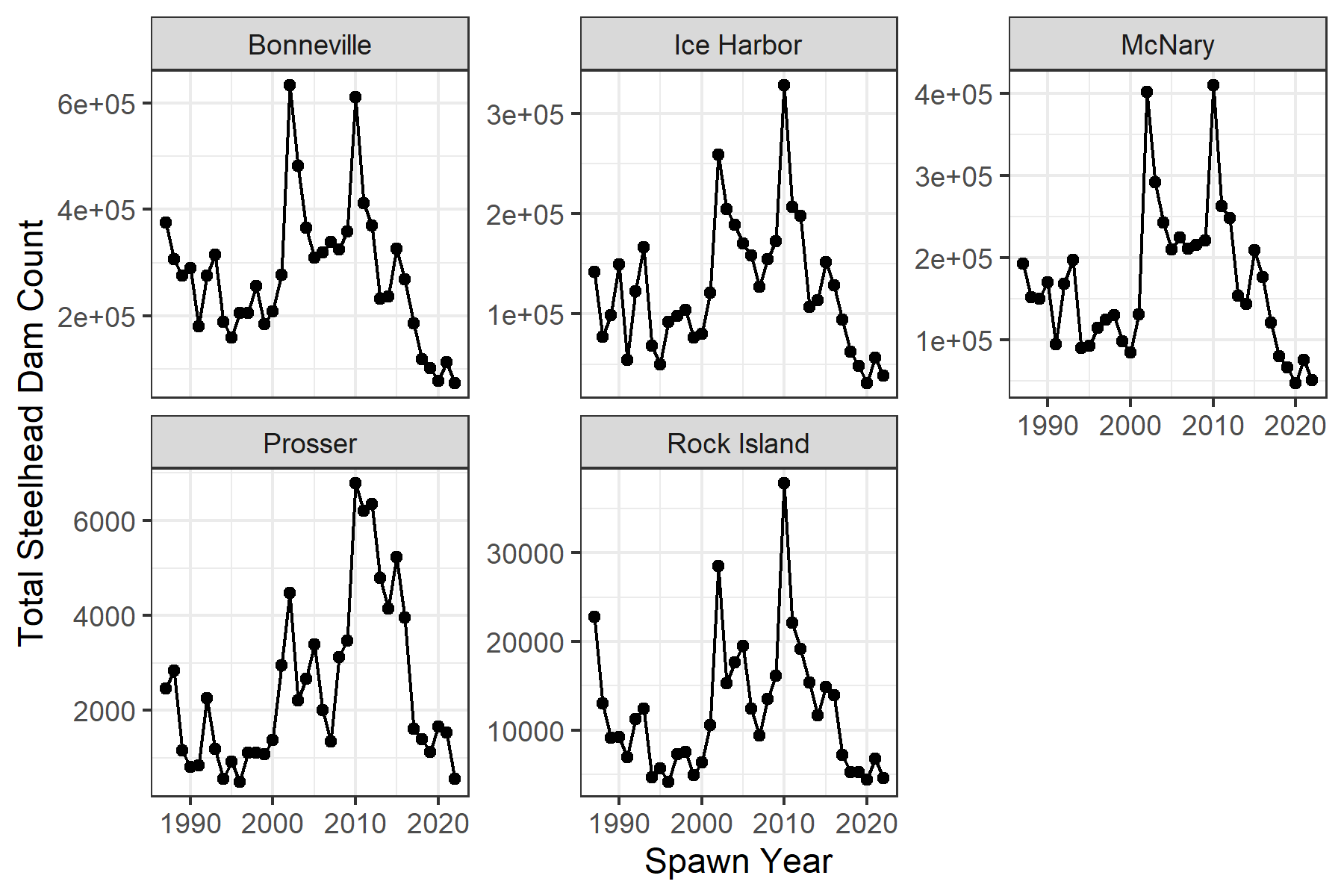


Figure 2.3: Time-series of counts from various Columbia River dams, from June 1 the year prior to May 31 of that spawn year.

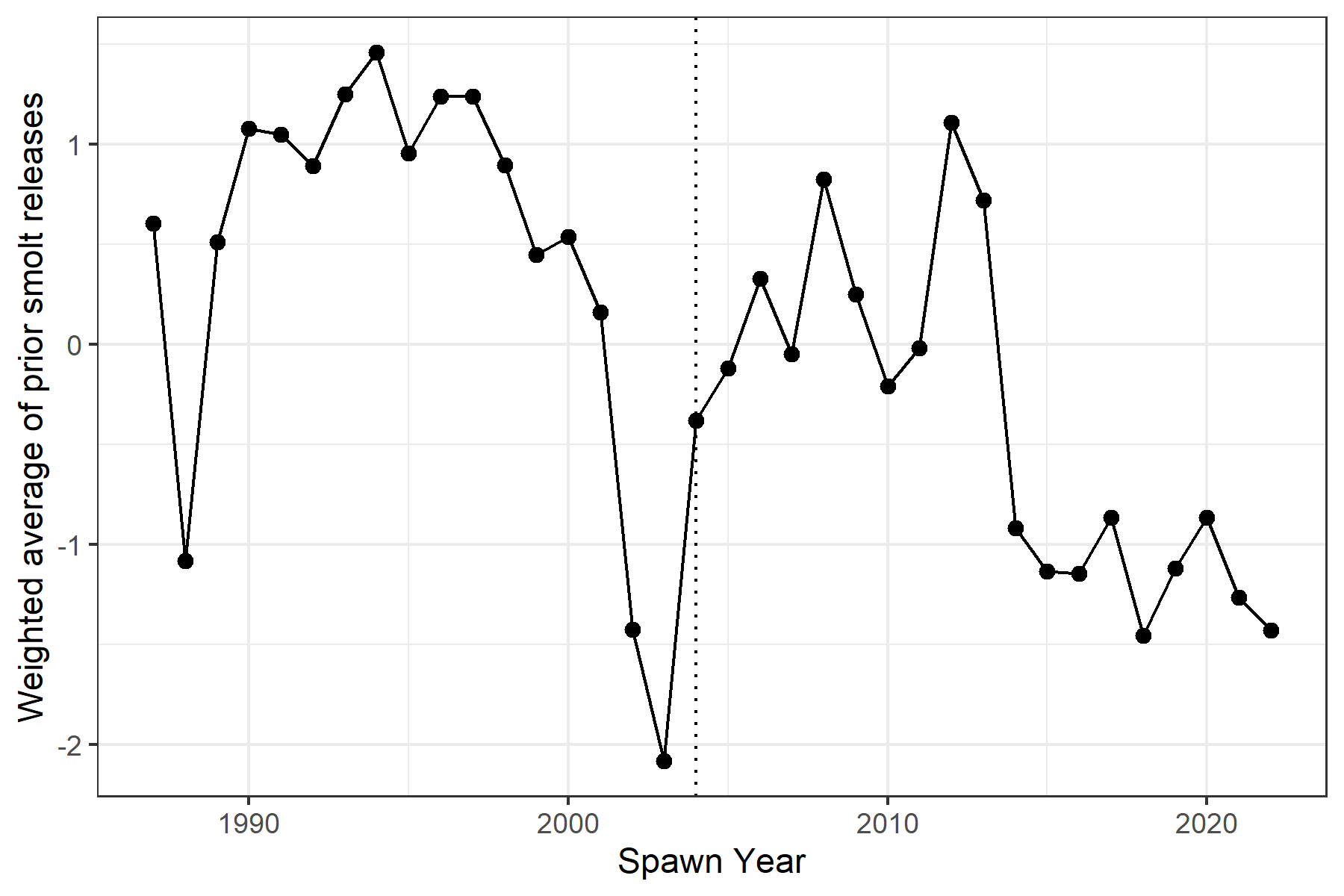


Figure 2.4: Time-series of normalized weighted average of smolt releases prior to the spawn year (x-axis). Dotted line shows when the new time-series begins.

### 2.2.1 MARSS Results

Table 2.1: AICc values for all models.

| Model Num. | Description | n Params | LogLik | AICc | delta AICc | Model Weight |
| --- | --- | --- | --- | --- | --- | --- |
| 3 | Q unconstrained, no covariates | 44 | -26.8 | 157.9 | 0.0 | 0.722 |
| 6 | Q unconstrained, smolt covariate | 45 | -26.7 | 160.6 | 2.7 | 0.184 |
| 9 | Q unconstrained, no covariates, U unequal | 51 | -18.9 | 162.3 | 4.5 | 0.076 |
| 12 | Q unconstrained, smolt covariate, U unequal | 52 | -18.9 | 165.2 | 7.4 | 0.018 |
| 2 | Q mostly independent, no covariates | 24 | -190.4 | 433.3 | 275.5 | 0.000 |
| 5 | Q mostly independent, smolt covariate | 25 | -190.3 | 435.6 | 277.8 | 0.000 |
| 8 | Q mostly independent, no covariates, U unequal | 31 | -185.1 | 439.9 | 282.0 | 0.000 |
| 11 | Q mostly independent, smolt covariate, U unequal | 32 | -184.1 | 440.5 | 282.6 | 0.000 |
| 1 | Q diag and unequal, no covariates | 23 | -195.6 | 441.4 | 283.6 | 0.000 |
| 4 | Q diag and unequal, smolt covariate | 24 | -195.6 | 443.8 | 285.9 | 0.000 |
| 7 | Q diag and unequal, no covariates, U unequal | 30 | -193.0 | 453.2 | 295.4 | 0.000 |
| 10 | Q diag and unequal, smolt covariate, U unequal | 31 | -193.0 | 455.7 | 297.8 | 0.000 |

The results (Table 2.1) show model number 3 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. The next two models (by AICc) were identical to the previous two but included a possible trend in each time-series. Although the trends for all states were estimated to be slightly negative in both models, the 95% confidence intervals overlapped zero in every case. These models also had very low weight (Table 2.1).

Table 2.2: Estimates of Q matrix from model 3, showing variance and co-variance estimates.

|  | Wen. Hatch | Wen. Wild | BON | IHR | MCN | PRO | RIS |
| --- | --- | --- | --- | --- | --- | --- | --- |
| Wen. Hatch | 0.236 | 0.188 | 0.142 | 0.162 | 0.165 | 0.158 | 0.184 |
| Wen. Wild | 0.188 | 0.213 | 0.097 | 0.106 | 0.112 | 0.204 | 0.173 |
| BON | 0.142 | 0.097 | 0.098 | 0.115 | 0.114 | 0.085 | 0.110 |
| IHR | 0.162 | 0.106 | 0.115 | 0.138 | 0.136 | 0.086 | 0.125 |
| MCN | 0.165 | 0.112 | 0.114 | 0.136 | 0.136 | 0.090 | 0.128 |
| PRO | 0.158 | 0.204 | 0.085 | 0.086 | 0.090 | 0.226 | 0.161 |
| RIS | 0.184 | 0.173 | 0.110 | 0.125 | 0.128 | 0.161 | 0.159 |

Table 2.3: Estimates of selected parameters from the best model. Terms containing “new” were inputs to the model, derived from the new time-series estimates.

| Term | Estimate | Std Error | Conf Low | Conf Up |
| --- | --- | --- | --- | --- |
| A.a\_old\_hos | 0.078 | 0.196 | -0.306 | 0.461 |
| A.a\_old\_nos | 0.236 | 0.108 | 0.025 | 0.447 |
| R.r\_old\_hos | 0.673 | 0.166 | 0.348 | 0.999 |
| R.r\_old\_nos | 0.087 | 0.026 | 0.037 | 0.137 |
| R.r\_new\_hos | 0.169 | 0.073 | 0.096 | 0.369 |
| R.r\_new\_nos | 0.137 | 0.067 | 0.078 | 0.301 |

Table 2.2 shows the estimates of the process error covariance matrix, , from the best supported model. Table 2.3 shows other parameter estimates from the selected model.

Figure 2.5 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 matrices. (models 3 and 6). Predictions are slightly greater for the model with a smolt release covariate, but only in the earlier years.

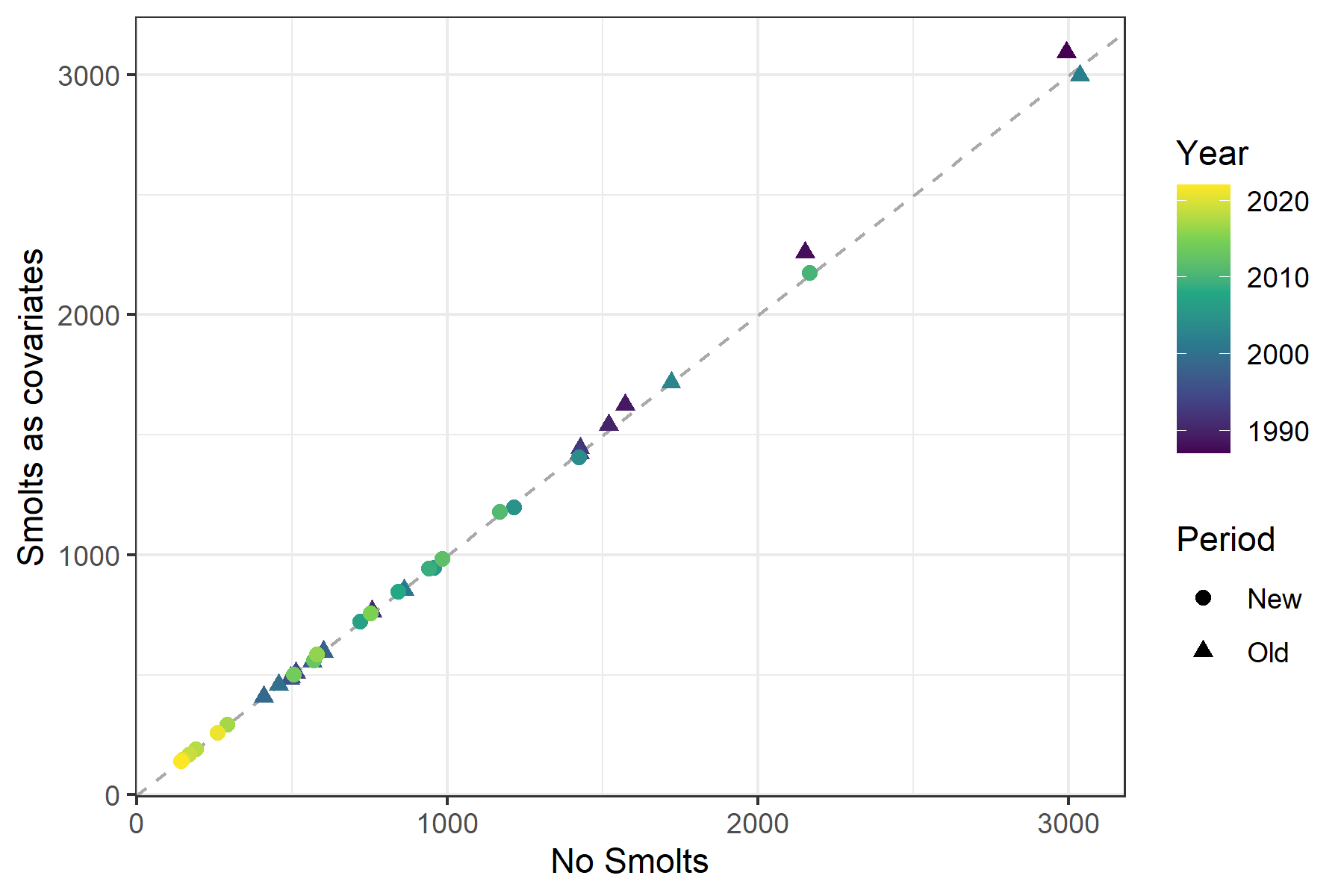


Figure 2.5: Comparison of predicted states of hatchery spawners for a model with no smolt release covariate (x-axis) and one that includes that covariate (y-axis). The period refers to whether the new time-series estimates exist.

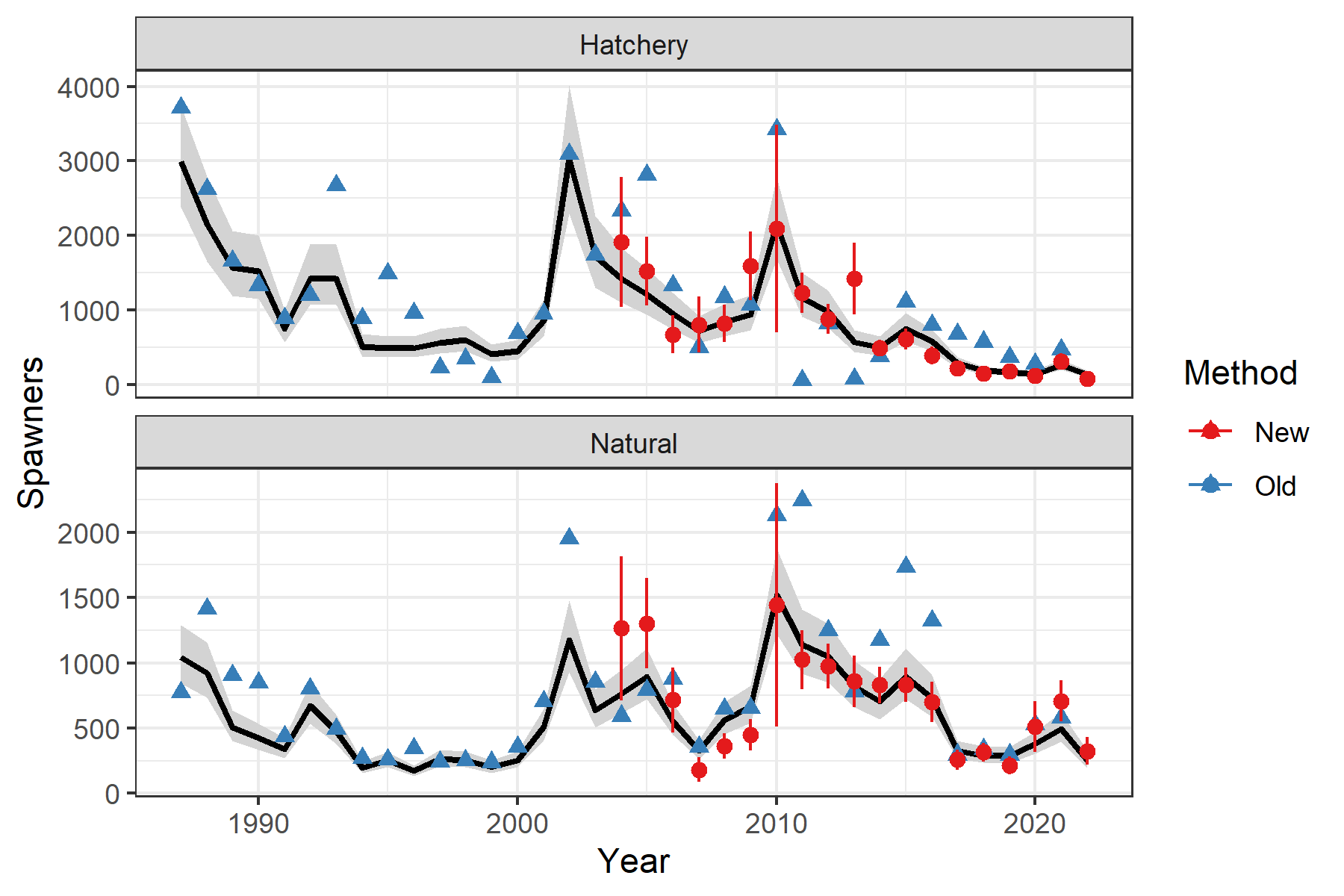


Figure 2.6: Estimates of spawners through the years, faceted by origin. Predicted spawners is the black line with 95% confidence interval in gray. Blue triangles depict estimates from the old time-series, while red points and 95% confidence intervals are from the new time-series.

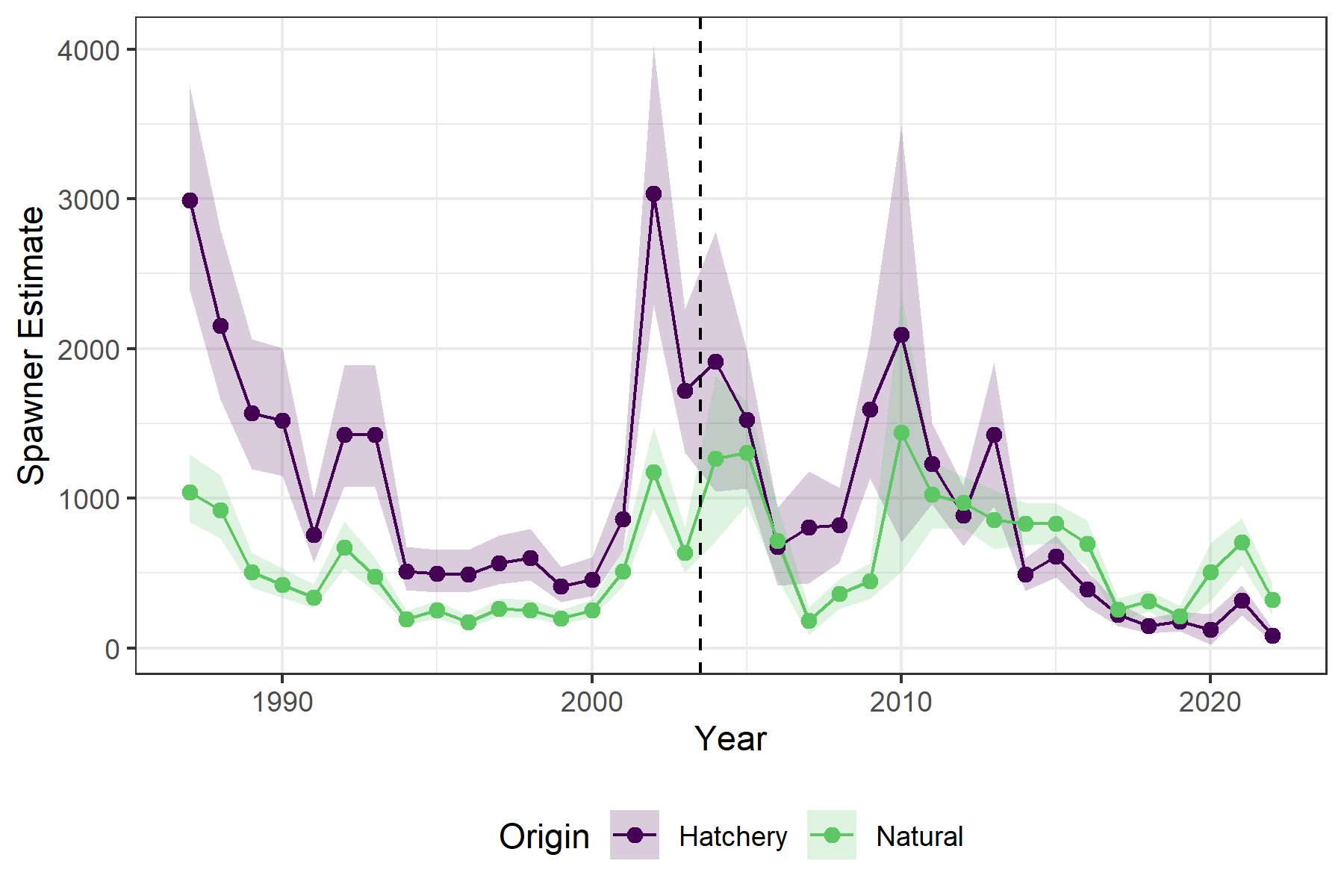


Figure 2.7: Updated estimates of spawners through the years, colored by origin, showing point estimates and 95% confidence intervals. Dashed vertical line differentiates older and newer time-series.

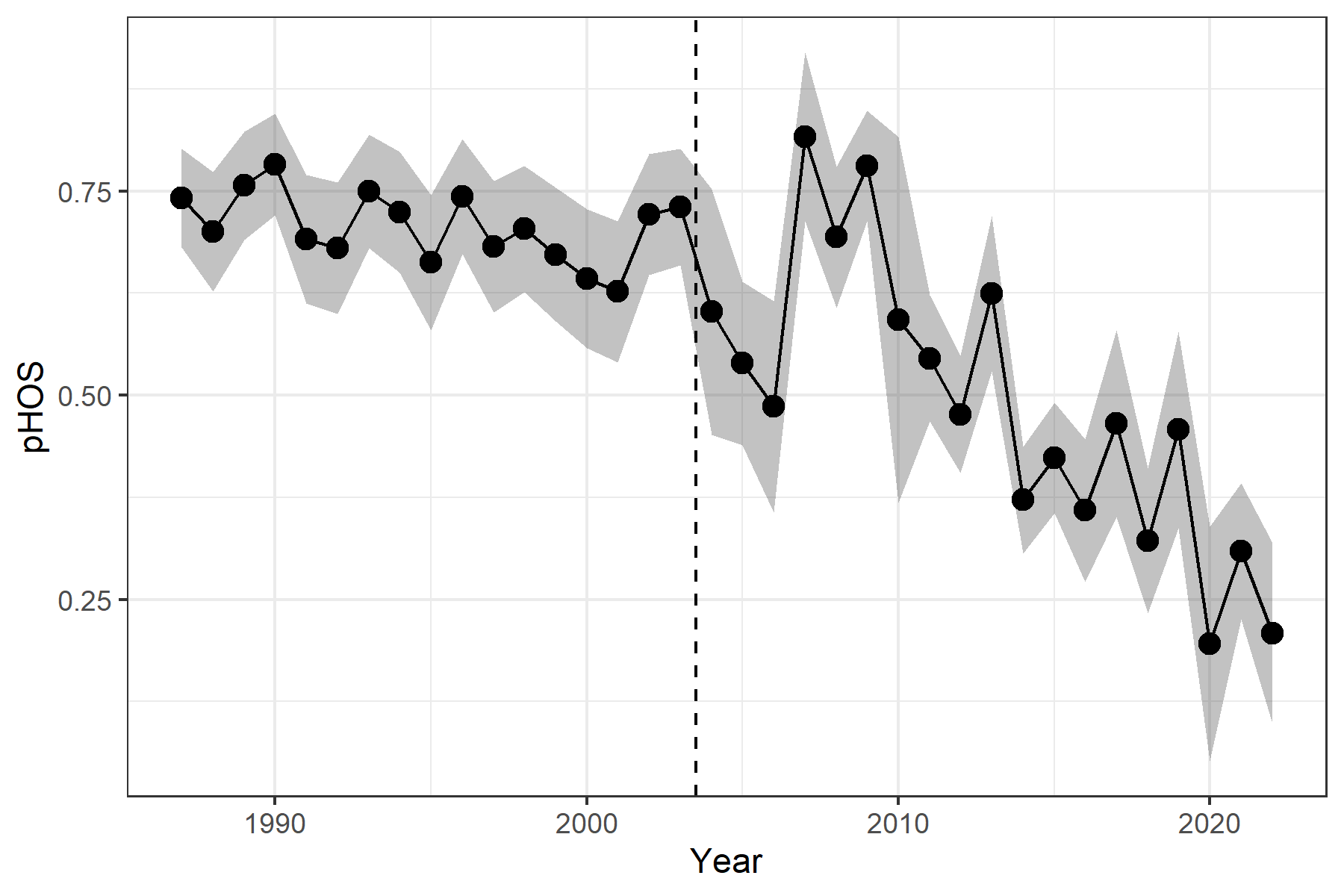


Figure 2.8: Estimates of pHOS based on the updated time-series, showing 95% confidence intervals. Dashed vertical line differentiates older and newer time-series.

# 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 2.5), we decided against using that model and chose the one with the lowest AICc score.

Table 5.1 in Appendix 5.1 shows the updated time-series, including a brief description of the method used to generate the estimate each year.

# 4 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).

# Appendices

# 5 Appendix A

## 5.1 Updated Time Series

Table 5.1: Updated time-series of steelhead spawners in the Wenatchee, by origin.

| Year | Origin | Method | Estimate | SE | LCI | UCI |
| --- | --- | --- | --- | --- | --- | --- |
| 1987 | Hatchery | MARSS | 2,992 | 347.1 | 2,389 | 3,747 |
| 1987 | Natural | MARSS | 1,042 | 115.3 | 840 | 1,292 |
| 1988 | Hatchery | MARSS | 2,151 | 289.2 | 1,658 | 2,790 |
| 1988 | Natural | MARSS | 920 | 107.9 | 733 | 1,156 |
| 1989 | Hatchery | MARSS | 1,571 | 221.5 | 1,197 | 2,063 |
| 1989 | Natural | MARSS | 505 | 60.1 | 401 | 636 |
| 1990 | Hatchery | MARSS | 1,520 | 218.1 | 1,152 | 2,005 |
| 1990 | Natural | MARSS | 422 | 50.5 | 335 | 533 |
| 1991 | Hatchery | MARSS | 757 | 109.5 | 573 | 1,001 |
| 1991 | Natural | MARSS | 339 | 40.6 | 269 | 427 |
| 1992 | Hatchery | MARSS | 1,427 | 207.0 | 1,078 | 1,888 |
| 1992 | Natural | MARSS | 670 | 80.3 | 531 | 845 |
| 1993 | Hatchery | MARSS | 1,426 | 207.3 | 1,077 | 1,888 |
| 1993 | Natural | MARSS | 476 | 57.1 | 377 | 600 |
| 1994 | Hatchery | MARSS | 512 | 74.6 | 387 | 679 |
| 1994 | Natural | MARSS | 195 | 23.4 | 154 | 246 |
| 1995 | Hatchery | MARSS | 496 | 72.4 | 375 | 658 |
| 1995 | Natural | MARSS | 252 | 30.3 | 200 | 319 |
| 1996 | Hatchery | MARSS | 493 | 72.0 | 372 | 654 |
| 1996 | Natural | MARSS | 170 | 20.5 | 135 | 215 |
| 1997 | Hatchery | MARSS | 565 | 82.6 | 426 | 749 |
| 1997 | Natural | MARSS | 263 | 31.7 | 209 | 333 |
| 1998 | Hatchery | MARSS | 601 | 88.0 | 454 | 798 |
| 1998 | Natural | MARSS | 253 | 30.5 | 201 | 320 |
| 1999 | Hatchery | MARSS | 410 | 59.9 | 309 | 543 |
| 1999 | Natural | MARSS | 200 | 24.0 | 158 | 252 |
| 2000 | Hatchery | MARSS | 457 | 66.9 | 345 | 607 |
| 2000 | Natural | MARSS | 254 | 30.6 | 201 | 321 |
| 2001 | Hatchery | MARSS | 860 | 125.6 | 649 | 1,140 |
| 2001 | Natural | MARSS | 512 | 61.6 | 406 | 647 |
| 2002 | Hatchery | MARSS | 3,036 | 440.3 | 2,295 | 4,017 |
| 2002 | Natural | MARSS | 1,173 | 140.8 | 929 | 1,480 |
| 2003 | Hatchery | MARSS | 1,720 | 244.5 | 1,307 | 2,263 |
| 2003 | Natural | MARSS | 635 | 75.6 | 504 | 800 |
| 2004 | Hatchery | 1 Obs. Model & Expansion | 1,915 | 443.1 | 1,047 | 2,783 |
| 2004 | Natural | 1 Obs. Model & Expansion | 1,264 | 280.7 | 714 | 1,814 |
| 2005 | Hatchery | 1 Obs. Model & Expansion | 1,526 | 235.5 | 1,064 | 1,988 |
| 2005 | Natural | 1 Obs. Model & Expansion | 1,303 | 177.0 | 956 | 1,650 |
| 2006 | Hatchery | 1 Obs. Model & Expansion | 676 | 131.6 | 418 | 934 |
| 2006 | Natural | 1 Obs. Model & Expansion | 714 | 127.6 | 464 | 964 |
| 2007 | Hatchery | 1 Obs. Model & Expansion | 807 | 190.8 | 433 | 1,181 |
| 2007 | Natural | 1 Obs. Model & Expansion | 181 | 47.1 | 89 | 273 |
| 2008 | Hatchery | 1 Obs. Model & Expansion | 822 | 126.7 | 574 | 1,070 |
| 2008 | Natural | 1 Obs. Model & Expansion | 363 | 50.2 | 265 | 461 |
| 2009 | Hatchery | 1 Obs. Model & Expansion | 1,595 | 233.3 | 1,138 | 2,052 |
| 2009 | Natural | 1 Obs. Model & Expansion | 447 | 61.9 | 326 | 568 |
| 2010 | Hatchery | 1 Obs. Model & Expansion | 2,095 | 708.9 | 706 | 3,484 |
| 2010 | Natural | 1 Obs. Model & Expansion | 1,441 | 476.1 | 508 | 2,374 |
| 2011 | Hatchery | 1 Obs. Model & DABOM | 1,229 | 137.8 | 959 | 1,499 |
| 2011 | Natural | 1 Obs. Model & DABOM | 1,025 | 115.7 | 798 | 1,252 |
| 2012 | Hatchery | 1 Obs. Model & DABOM | 885 | 102.7 | 684 | 1,086 |
| 2012 | Natural | 1 Obs. Model & DABOM | 972 | 87.5 | 800 | 1,144 |
| 2013 | Hatchery | 1 Obs. Model & DABOM | 1,427 | 244.5 | 948 | 1,906 |
| 2013 | Natural | 1 Obs. Model & DABOM | 858 | 100.6 | 661 | 1,055 |
| 2014 | Hatchery | 2 Obs. Model & DABOM | 491 | 55.9 | 381 | 601 |
| 2014 | Natural | 2 Obs. Model & DABOM | 829 | 71.8 | 688 | 970 |
| 2015 | Hatchery | 2 Obs. Model & DABOM | 612 | 71.0 | 473 | 751 |
| 2015 | Natural | 2 Obs. Model & DABOM | 833 | 67.1 | 702 | 964 |
| 2016 | Hatchery | 2 Obs. Model & DABOM | 391 | 61.2 | 271 | 511 |
| 2016 | Natural | 2 Obs. Model & DABOM | 698 | 78.7 | 544 | 852 |
| 2017 | Hatchery | 2 Obs. Model & DABOM | 223 | 39.0 | 147 | 299 |
| 2017 | Natural | 2 Obs. Model & DABOM | 256 | 39.7 | 178 | 334 |
| 2018 | Hatchery | 2 Obs. Model & DABOM | 149 | 25.4 | 99 | 199 |
| 2018 | Natural | 2 Obs. Model & DABOM | 314 | 37.3 | 241 | 387 |
| 2019 | Hatchery | 2 Obs. Model & DABOM | 179 | 32.7 | 115 | 243 |
| 2019 | Natural | 2 Obs. Model & DABOM | 212 | 35.1 | 143 | 281 |
| 2020 | Hatchery | RT Survival | 124 | 52.4 | 21 | 227 |
| 2020 | Natural | RT Survival | 509 | 99.4 | 314 | 704 |
| 2021 | Hatchery | 2 Obs. Model & DABOM | 317 | 51.9 | 215 | 419 |
| 2021 | Natural | 2 Obs. Model & DABOM | 708 | 80.6 | 550 | 866 |
| 2022 | Hatchery | 2 Obs. Model & DABOM | 85 | 25.3 | 35 | 135 |
| 2022 | Natural | 2 Obs. Model & DABOM | 323 | 54.6 | 216 | 430 |