MONITORING AND EVALUATION PLAN FOR PUD HATCHERY PROGRAMS

2019 UPDATE

December 18, 2019

Prepared by:

Tracy Hillman
Tom Kahler
Greg Mackey
Andrew Murdoch
Keely Murdoch
Todd Pearsons
Mike Tonseth
Catherine Willard

Prepared for:

HCPs and PRCC Hatchery Committees

Citation: Hillman, T., T. Kahler, G. Mackey, Andrew Murdoch, K. Murdoch, T. Pearsons, M. Tonseth, and C. Willard. 2017. Monitoring and evaluation plan for PUD hatchery programs: 2017 update. Report to the HCP and PRCC Hatchery Committees, Wenatchee and Ephrata, WA.

TABLE OF CONTENTS

LIST OF APPENDICES

SECTION 1: INTRODUCTION

This document is an update of the monitoring and evaluation (M&E) plan of the salmon and steelhead hatchery programs funded by Douglas, Chelan, and Grant County Public Utility Districts (PUDs). Programmatic changes, evaluation of data collection methods, and M&E results from the past several years, along with shifting management paradigms affect M&E needs, all of which have occurred under advancing fish culture and monitoring techniques. As required by the programs, this document is a result of a five-year review intended to expand on and coalesce previous M&E documents (BAMP 1998; Cates et al. 2005; Murdoch and Peven 2005; Hays et al. 2006; Pearsons and Langshaw 2009a, 2009b; Hillman et al. 2013) with inclusion of new information.

Fishery management agencies developed the following general goal statements for hatchery programs, which were adopted by the HCPs Hatchery Committees and PRCC Hatchery Sub-Committee (hereafter, Hatchery Committees):

1. Support the recovery of ESA-listed species by increasing the abundance of the natural adult population, while ensuring appropriate spatial distribution, genetic stock integrity, and adult spawner productivity.
2. Increase the abundance of the natural adult population of unlisted plan species, while ensuring appropriate spatial distribution, genetic stock integrity, and adult spawner productivity. In addition, provide harvest opportunities in years when spawning escapement is sufficient to support harvest.
3. Provide salmon for harvest and increase harvest opportunities, while segregating returning adults from natural tributary spawning populations.

Following the development of Hatchery and Genetic Management Plans (HGMPs), artificial supplementation programs are now characterized into three categories. The first type, integrated conservation programs, are intended to support or restore natural populations. These programs focus on increasing the natural production of targeted fish populations. A fundamental assumption of this strategy is that hatchery fish returning to the spawning grounds are reproductively similar to naturally produced fish. The second type, safety-net programs, are extensions of conservation programs, but are intended to function as reserve capacity for conservation programs in years of low returns. The safety-net provides a demographic and genetic reserve for the natural population. That is, in years of abundant returns they function like segregated programs, and in low return years they can be managed as conservation programs. Lastly harvest augmentation programs are intended to increase harvest opportunities while limiting interactions with wild-origin counterparts.

Monitoring is needed to determine if the hatchery programs are meeting the intended management objectives of conservation, safety-net, or harvest augmentation programs. Objectives for hatchery programs are generally grouped into three categories of performance indicators:

1. In-Hatchery: Is the program meeting the hatchery production objectives?
2. In-Nature: How do fish from the program perform after release?
    a. Conservation Program:

        - How does the program affect target population abundance and productivity?
        - How does the program affect target population long-term fitness?
    b. Safety-Net Program:
        - How does the program affect target population long-term fitness?
    c. Harvest Augmentation Program:
        - Does the program provide harvest opportunities?
3. Risk Assessment: Does the program pose risks to other populations?

Objectives in this plan have been organized in a hierarchy where productivity indicators are the primary metrics used to assess if conservation and safety-net program goals have been met; harvest rates and effects on non-targeted populations are used for harvest programs. In cases where productivity indicators are not available or results are equivocal, monitoring indicators may be used to help evaluate the performance of the program. Evaluations of monitoring indicators may not provide sufficiently powerful conclusions on which to base management actions, although they may provide insight as to why a productivity indicator did or did not meet the program goal. Therefore, the relationship between hatchery programs and indicators can be viewed in a chain-of-causation: management actions within the hatchery programs affect the status of monitoring indicators, which in turn influence productivity indicators (Figure 1).

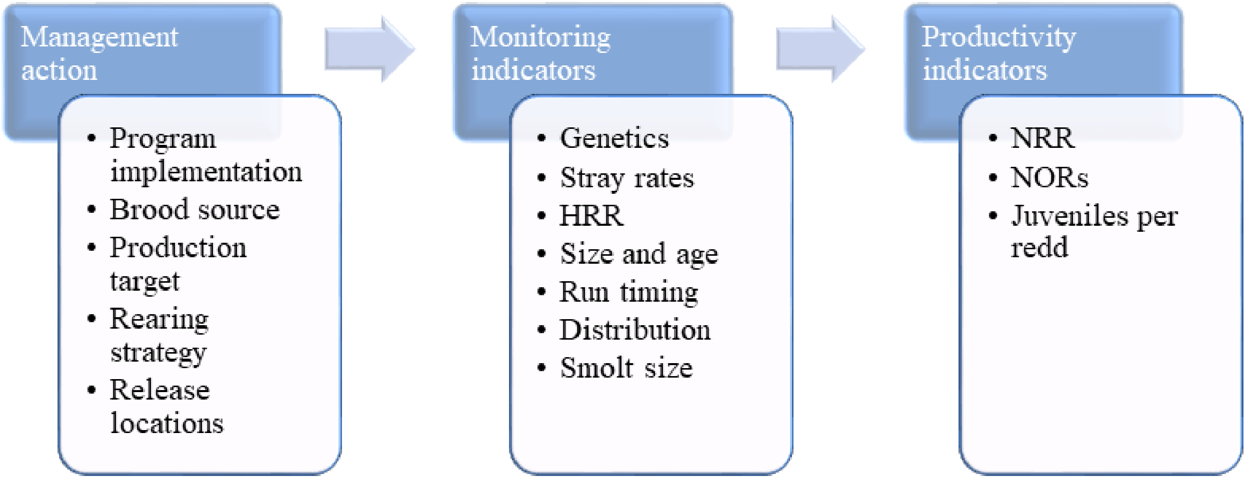


Figure 1. Relationship of indicators to the assessment of supplementation programs. Management actions affect monitoring indicators, which influence productivity indicators. Monitoring indicators may be used to hypothesize the magnitude of influence on productivity.

The primary goal of a conservation program is to contribute to the rebuilding and recovery of naturally reproducing populations within their native habitat. In this plan, natural replacement rates (NRR), recruitment of naturally-produced fish (NOR), and juvenile productivity (juveniles per redd) are important indicators for assessing the success of supplementation. These indicators are difficult to measure precisely and are quite variable in space and time. Therefore, monitoring indicators can be evaluated to help assess if productivity was related to the hatchery programs or other factors (Table 1).

Table 1. Program objectives, indicators, and goals for conservation hatchery programs including productivity and monitoring indicators (also applies to safety-net programs when used to support a conservation program).

				Program goals		
	Objective	**Indicator**	**Target**	**Rebuild natural populations**	**Maintain genetic diversity**	**Opportunity for harvest**
Productivity indicators	Determine if the program has increased the number of naturally spawning adults	Abundance of natural spawners	Increase	✓		✓
		Adult productivity (NRR)	No decrease	✓		
	Determine if the proportion of hatchery fish affects freshwater productivity	Residuals vs. pHOS	No relationship	✓		
		Juveniles per redd vs. pHOS	No relationship	✓		
Monitoring indicators	Determine if run timing and distribution meets objectives	Migration timing	No difference	✓	✓	
		Spawn timing[1]	No difference	✓	✓	
		Redd distribution[2]	No difference	✓	✓	
	Determine if program has affected genetic diversity and population structure	Allele frequency (hatchery vs. wild)	No difference		✓	
		Genetic distance between populations	No difference		✓	
		Effective population size	Increase		✓	
		Age and size at maturity	No difference		✓	
	Determine if hatchery survival meets expectations	HRR	HRR > NRR	✓		
		HRR	HRR ≥ Goal[3]	✓		
	Determine if recipient stray rate of hatchery fish is acceptable	Out of basin	≤ 5%	✓	✓	
		Within basin	≤ 10%	✓	✓	
	Determine if hatchery fish were released at program targets	Size and number	= Target[4]	✓		
	Provide harvest opportunities when appropriate	Harvest	Escapement goals			✓

[1] Hatchery and natural-origin fish should spawn at the same time across the range of elevations within the spawning distribution of each stock.

[2] Hatchery and natural-origin fish should spawn in the same locations. Exceptions are the Carlton and Dryden Summer Chinook programs (see Appendix 4).

[3] HRR targets are identified in Appendix 2.

[4] Number and size targets are identified in Table 3 and Appendix 5.

A flow of information following sequential, logical steps will be employed to evaluate supplementation programs, consistent with the indicators described in Table 1. For example, a hatchery program, at a minimum, must be able to produce more adults per spawner than would occur in the natural environment. Should the program fail this test, hatchery operations should be evaluated to determine if improvements can correct the problem. If a program successfully replaces the required number of adults, it is then evaluated against a reference population or condition, if available, to determine if it has increased the overall number of naturally spawning fish (including both hatchery- and natural-origin adults), increased the number of natural-origin spawners, and to test if productivity of the natural population has changed. When these goals are met, the program is considered successful. When these goals are not met, monitoring indicators may infer why the program is not achieving its goals

If suitable reference populations are not available, other comparisons can be used to help evaluate treatment responses. Evaluation of programs may pursue the following approaches:

- Comparison to reference population(s) that do not contain pre-treatment data.
- Before treatment and after treatment comparisons.
- Comparison to standard(s).
- Comparison to other suitable reference conditions.

Methodologies for selecting reference streams, analyzing data from treatment and reference stream comparisons, and other comparisons are presented in Hillman et al. (2012) (see Appendix 6).

The primary goals of a safety-net program are to provide demographic and genetic reserves for a population that is supplemented by a conservation program (Table 2). Harvest and adult management may be used to control escapement of spawners when appropriate. Monitoring focuses on estimating the number of fish that escape to spawn naturally and stray rates and in-hatchery performance evaluation.

Table 2. Program objectives, indicators, and goals for segregated harvest augmentation hatchery programs including monitoring indicators.

				Program goals		
	Objective	**Indicator**	**Target**	**Rebuild natural populations**	**Maintain genetic diversity**	**Opportunity for harvest**
Monitoring indicators	Determine if hatchery survival meets expectations	HRR	HRR > NRR			✓
		HRR	HRR ≥ Goal[1]			✓
	Determine if stray rate of hatchery fish is acceptable	Out of basin	≤ 5%		✓	
		Within basin	≤ 10%		✓	
	Determine if hatchery fish were released at program targets	Size and number	= Target[2]			✓
	Provide harvest opportunities when appropriate	Harvest	When greater than escapement goals			✓

[1] HRR targets are identified in Appendix 2.

[2] Number and size targets are identified in Table 3 and Appendix 5.

The primary goal of a harvest augmentation program is to increase harvest opportunities, while segregating adults from natural spawning populations. In this plan, harvest opportunity, survival rates, and stray rates are important indicators for assessing the success of harvest augmentation. These indicators are more readily quantified compared to productivity indicators (Table 2). A flow of information will be employed to evaluate harvest augmentation programs. Since harvest augmentation programs are typically segregated, monitoring indicators will be used to determine the success of a program.

Both monitoring and productivity indicators will be used to evaluate the success of hatchery programs. In the event that the statistical power of tests that involve productivity indicators is insufficient to inform sound management decisions, some of the monitoring indicators may be used to guide management. Figure 2 show the categories of indicators associated with each component of monitoring.

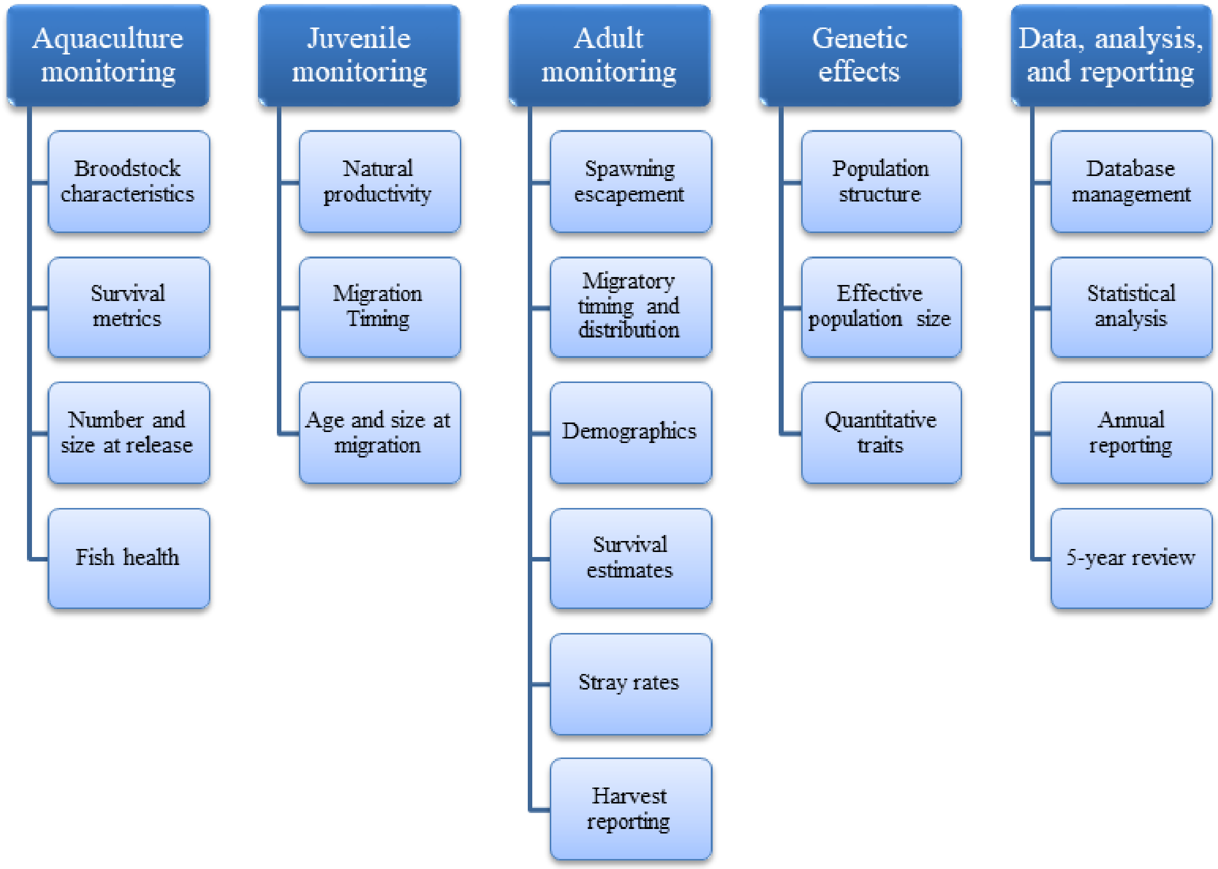


Figure 2. Overview of Monitoring and Evaluation Plan Categories and Components (not including regional objectives).

The overarching goals of conservation, safety-net, and harvest augmentation programs, as described above, are shown in detail in Figure 3. The flow chart (Figure 3) shows the relationship of overarching program goals, the strategies used to meet the goals, the monitoring and evaluation objectives used to evaluate the strategies and determine if goals are being met, and the adaptive management cycle associated with the programs (see Tables 1 and 2 for the indicators under each objective). The logic depicted in this flow chart shall be used to assess M&E results and apply those results to management decisions. Table 3 presents the current hatchery programs releasing fish in the Upper Columbia Basin.

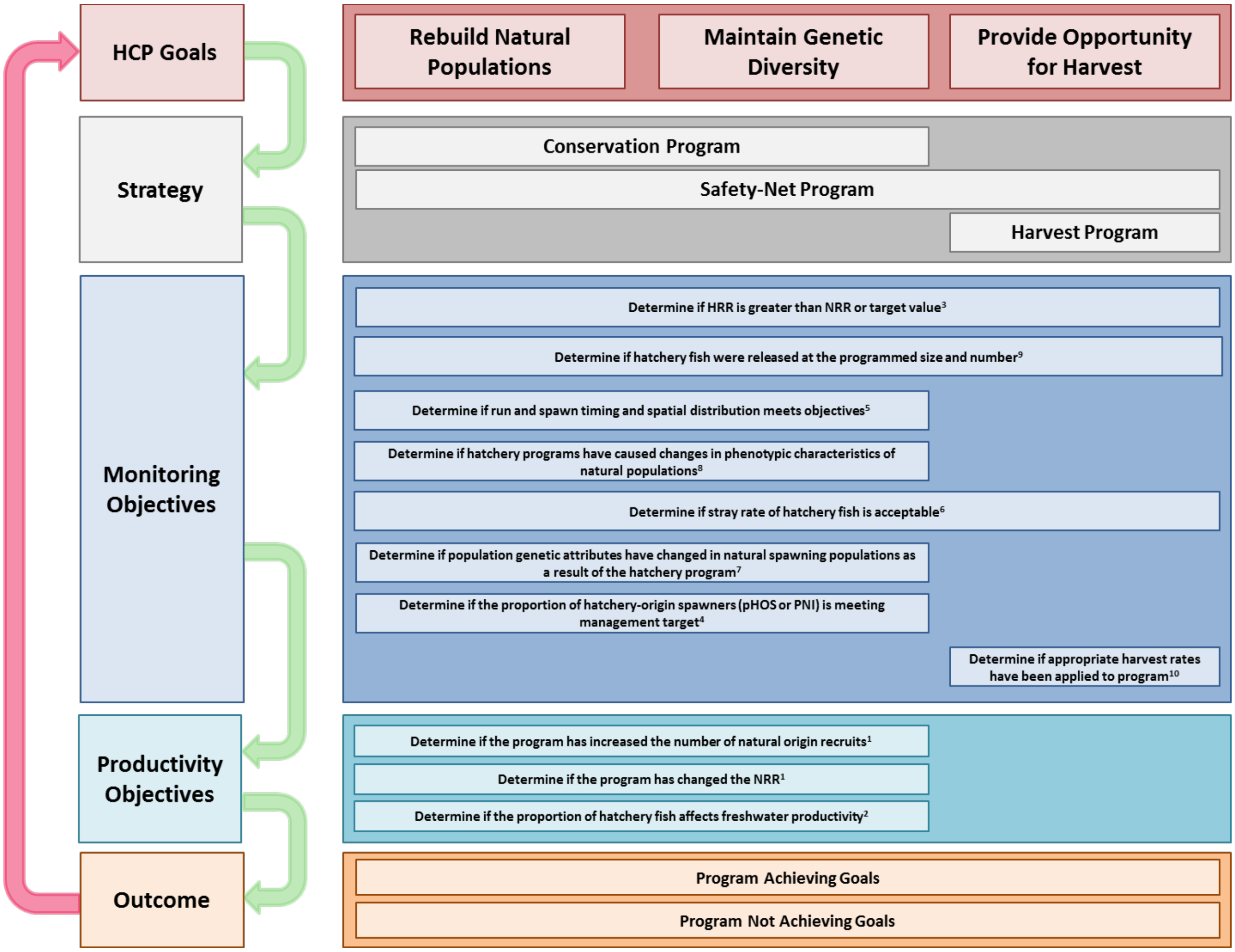


Figure 3. Adaptive management flow chart depicting HCP goals, associated strategies to meet the goals, the monitoring and evaluation objectives (indicated in superscript), and the adaptive management feedback cycle. The strategies, objectives, and outcomes are aligned vertically under the corresponding goals.

Table 3. Hatchery programs in the mid-Columbia River Basin, 2012. Funding entities included Douglas PUD (D), Chelan PUD (C), Grant PUD (G), Bonneville Power Administration (B), Bureau of Reclamation (O), and Army Corps of Engineers (A) and are listed in order of contribution. Total artificial production targets in the mid-Columbia River exceeds 20 million juveniles annually.

Program	Species	Basin	Purpose	Funding Entity	Production
Methow[5]	Spring Chinook [1]	Methow	NNI/Conservation	G, C, D	223,765
Chief Joseph[7]	Spring Chinook	Okanogan	Reintroduction/Harvest	B, G, C, D	900,000
Chiwawa[5]	Spring Chinook [1]	Wenatchee	NNI/Conservation	C	144,026
Nason[5]	Spring Chinook [1]	Wenatchee	NNI/Conservation	G	223,670
Winthrop[7]	Spring Chinook [2]	Methow	Safety-Net	O	400,000
Leavenworth	Spring Chinook [2]	Wenatchee	Harvest	O	1,200,000
Wells[5]	Steelhead [1]	Columbia	Inundation/Safety-Net	D	160,000
Winthrop[7]	Steelhead [1]	Methow	Conservation	O	100,000-200,000
Wells[5]	Steelhead [1]	Methow	Inundation/Safety-Net	D	100,000
Wells/Omak[5, 6]	Steelhead [1]	Okanogan	NNI/Conservation	G	100,000
Wells[5]	Steelhead [1]	Twisp	Inundation/Conservation	D	40,000
Wells[5]	Steelhead [1]	Twisp	NNI/Conservation	D	8,000
Chiwawa[5]	Steelhead [1]	Wenatchee	NNI/Conservation	C	22,000
Chiwawa[5]	Steelhead [1]	Wenatchee	Inundation/Harvest	C	165,000
Chiwawa[5]	Steelhead [1]	Wenatchee	Species trade	C	60,300
Ringold	Steelhead [9]	Columbia	Harvest	Mitchell Act	180,000
Wells[5]	Summer Chinook [2, 3]	Columbia	Inundation/Harvest	D	484,000
Chief Joseph[7]	Summer Chinook [3]	Okanogan	NNI/Cons./Harvest	B, C, D	700,000
Chelan Falls[5]	Summer Chinook [2]	Chelan	Inundation/Harvest	C	400,000
Chelan Falls[5]	Summer Chinook [2]	Chelan	NNI/Conservation	C	176,000
Wells[5]	Summer Chinook [2]	Columbia	Inundation/Harvest	D	320,000
Entiat	Summer Chinook	Entiat	Harvest	O	400,000
Carlton[5]	Summer Chinook	Methow	NNI/Conservation	G	200,000
Chief Joseph[7]	Summer Chinook	Okanogan	NNI/Cons./Harvest	B, G, C, D	1,300,000
Dryden[5]	Summer Chinook	Wenatchee	NNI/Conservation	C, G	500,000
Priest[5]	Fall Chinook[3]	Columbia	Inundation/Harvest	G	5,000,000
Priest[5]	Fall Chinook[3]	Columbia	NNI/Harvest	G	325,543
Priest[5]	Fall Chinook[4]	Columbia	Fry loss/Harvest	G	273,961
Priest[5, 7]	Fall Chinook[3]	Columbia	Harvest	A	1,700,000
Ringold[7]	Fall Chinook[3]	Columbia	Harvest	A	3,500,000
Yakama Nation	Coho	Wenatchee	Reintroduction/Harvest	B, G, C, D	1,000,000
Yakama Nation[8]	Coho	Methow	Reintroduction/Harvest	B, G, C, D	500,000
Skaha	Sockeye	Okanogan	Reintroduction/Harvest	C, G	≤ 5 M eggs

[1] Species listed under the Endangered Species Act.

[2] Segregated program.

[3] Sub-yearling production.

[4] Fry production.

[5] Program covered by this M&E Plan.

[6] Program also partially covered by CCT M&E Plan.

[7] Program affects PUD-funded programs covered by this plan.

[8] Planned to increase to 1,000,000.

[9] Part of the Mitchell Act suite of mitigation programs under the FCRPS BiOp.

SECTION 2: ADULT PRODUCTIVITY

2.1 Natural Replacement Rates of Supplemented Populations[1]

Objective 1: Determine if conservation programs have increased the number of naturally spawning and naturally produced adults of the target population and if the program has reduced the natural replacement rate (NRR) of the supplemented population.

At the core of a conservation program is the objective of increasing the number of spawning adults (i.e., the combined number of naturally produced and hatchery fish) in order to affect a subsequent increase in the number of returning naturally produced fish or natural-origin recruits (NOR). In order for the natural population to remain stable or to increase, the Natural Replacement Rate (NRR), or the ratio of NORs to the parent spawning population, must be at a level where parents are being replaced by their offspring as spawners in the next generation. It is possible to affect an increase in natural-origin spawners through supplementation with a stable or decreasing NRR. However, if the NRR is below replacement (NRR<1.0), termination of the supplementation program will result in a declining natural population should that state of NRR persist. The proportion of the hatchery-origin spawners (pHOS) that will increase natural production without creating adverse effects to the genetic diversity or reproductive success rate of the natural population is unknown, and may be dependent on how individual hatchery programs are operated, as well as available spawning and rearing habitat. Some programs restrict pHOS to reduce the risk to the natural population with the intent of optimizing productivity, concomitantly reducing the overall number of spawners. All other objectives of the M&E Plan either directly support this objective or seek to minimize negative effects of the conservation programs on non-target stocks of concern.

Differences in carrying capacities of supplemented and non-supplemented streams can confound the analysis of the effects of supplementation on total number of spawners returning to the streams. For example, if the supplemented population is at carrying capacity and the non-supplemented population is not, the total number of spawners returning to the non-supplemented population may show an increasing trend over time, while the supplemented population would show no increasing trend. To avoid concluding that the supplementation program has no effect or perhaps a negative effect on total spawners, density corrections should be included in the analyses. Hypotheses that may require density corrections are noted under each monitoring question. Appendix 1 describes methods for estimating carrying capacities.

Monitoring Questions:

Q1.1.1 Has the supplementation program changed the adult productivity (NRRs) of the supplemented populations?[2]

Target Species/Populations:

---

[1] Supplementation programs may include a safety-net component.

[2] Because adult productivity is affected by the abundance of the population (i.e., productivity decreases with increasing abundance), the goal of supplementation is to increase or maintain productivity, but not decrease it.

- Q1.1.1 applies to all conservation and safety-net stocks.

Statistical Hypotheses 1.1.1[3]:

- $Ho_{1.1.1.1}$: Slope in NRRs before supplementation ≤ slope in NRRs after supplementation.
- $Ho_{1.1.1.2}$: Differences in slopes in NRRs between supplemented and reference populations before supplementation ≤ differences in slopes in NRRs between supplemented and reference populations after supplementation.
- $Ho_{1.1.1.3}$: Mean NRRs before supplementation ≤ mean NRRs after supplementation.
- $Ho_{1.1.1.4}$: Mean ratio scores in NRRs before supplementation ≤ Mean ratio scores in NRRs during supplementation.
- $Ho_{1.1.1.5}$: Mean ratio scores in NRRs (adjusted for density dependence) before supplementation ≤ Mean ratio scores in NRRs (adjusted for density dependence) during supplementation. [This hypothesis adjusts NRRs for density-dependent effects (see Hillman et al. 2012 for details; Appendix 6).]
- $Ho_{1.1.1.6}$: There is no association between the proportion of hatchery-origin spawners (pHOS) and the residuals from the smooth hockey stick stock-recruitment curve; rho = 0. [If there is a significant negative association between pHOS and the residuals, then hatchery fish may be reducing the productivity of the wild population.]

Measured Variables:

- Number of hatchery and naturally produced fish on spawning grounds
- Number of naturally produced fish harvested

Derived Variables:

- Number of naturally produced recruits by brood year for both naturally produced parents and hatchery parents (≥age-3).
- NRRs (calculated as NORs/spawner).
- Stock-recruit models, parameters, and residuals.
- Includes ratio scores of NRRs (requires reference population[s]).
- Includes calculation of ratios NORs (requires reference population).
- Appendix 1: Spawning escapement and carrying capacity information (as applicable)

Spatial/Temporal Scale:

- Calculated annually based on brood year.
- Time series.

Possible Statistical Analysis:

---

[3] Quality and quantity of data will determine which hypotheses are evaluated. See Hillman et al. 2012 (Appendix 6) for details.

- These analyses shall be performed every 5-years. Use graphic analyses, trend analyses, t-tests, Aspin-Welch tests, and randomization tests to evaluate the statistical hypotheses (see Hillman et al. 2012; Appendix 6). The specific analysis used will depend on the availability of reference conditions.
- Correlation analysis will examine associations between hatchery adult composition and NRRs.
- On a five-year period, correlate productivity with extraneous factors such as ocean productivity indices.

Analytical Rules:

- This is a productivity indicator that will be used to assess the success of the supplementation program.
- Type I Error of 0.05.

2.2 Natural-Origin Recruits of Supplemented Populations

Monitoring Questions:

Q1.2.1: Has the supplementation program changed the abundance of NORs within the supplemented population?

Target Species/Populations:

- Q1.2.1 applies to all supplemented or safety net stocks.

Statistical Hypotheses 1.2.1[4]:

- $Ho_{1.2.1.1}$: Slope in NORs[5] before supplementation ≥ slope in NORs after supplementation.
- $Ho_{1.2.1.2}$: Differences in slopes in NORs between supplemented and reference populations before supplementation ≥ differences in slopes in NORs between supplemented and reference populations after supplementation.
- $Ho_{1.2.1.3}$: Mean NORs before supplementation ≥ mean NORs after supplementation.
- $Ho_{1.2.1.4}$: Mean ratio scores in NORs before supplementation ≥ Mean ratio scores in NORs during supplementation.
- $Ho_{1.2.1.5}$: Mean ratio scores in NORs/Maximum Recruitment before supplementation ≥ Mean ratio scores in NORs/Maximum Recruitment during supplementation. [This hypothesis adjusts NORs for the capacity of the habitat; it tests the fraction of the habitat saturated with NORs (see Hillman et al. 2012 for details).]
- $Ho_{1.2.1.6}$: There is no association between the proportion of hatchery-origin spawners (pHOS) and NORs; rho = 0. [If there is a significant negative association between

---

[4] Quality and quantity of data will determine which hypotheses are evaluated. See Hillman et al. 2012 (Appendix 7) for details.

[5] "Slope in NORS" refers to abundance of NORs across time (years).

pHOS and NORs, then hatchery fish may be reducing the reproductive success of the wild population.]

Measured Variables:

- Number of hatchery and naturally produced fish on spawning grounds.
- Number of hatchery and naturally produced fish taken for broodstock.
- Number of hatchery and naturally produced fish taken in harvest (if recruitment is to the Columbia).

Derived Variables:

- NORs (number of naturally produced recruits (total recruits) by brood year for both naturally produced parents and hatchery parents [≥age-3]).
- Stock-recruit models, parameters, and residuals.
- Includes ratio scores of NORs (requires reference population[s]).
- Estimates of carrying capacity (see Appendix 1).

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- These analyses shall be performed every 5-years. Use graphic analyses, trend analyses, t-tests, Aspin-Welch tests, and randomization tests to evaluate the statistical hypotheses (see Hillman et al. 2012). The specific analysis used will depend on the availability of reference conditions.
- Correlation analysis will examine associations between hatchery adult composition and NORs.
- On a five-year period, correlate NORs with extraneous factors such as ocean productivity indices.

Analytical Rules:

- This is a productivity indicator that will be used to assess the success of the supplementation program.
- Type I Error of 0.05.

SECTION 3: JUVENILE PRODUCTIVITY

3.1 Freshwater Juvenile Productivity

Objective 2: Determine if the proportion of hatchery fish on the spawning grounds affects the freshwater productivity of supplemented stocks.

Out-of-basin effects (e.g., smolt passage through the hydro system, harvest, and ocean productivity, etc.) influence the survival of smolts after they migrate from the tributaries. These effects introduce substantial variability into the adult-to-adult survival rates (NRRs and HRRs) and may mask in-basin effects (e.g., habitat quality, density-dependent mortality, and differential reproductive success of hatchery and naturally produced fish). Therefore, an estimate of freshwater productivity may help inform the performance of hatchery and natural-origin spawners.

The objective of estimating freshwater productivity in the Upper Columbia ESU/DPS is to estimate the survival from egg to a critical juvenile life stage(s) of target stocks. Smolt or juvenile production models generated from the information obtained through these programs will provide a level of predictability with greater sensitivity to in-basin effects than spawner-recruitment models that consider all effects.

Differences in the current carrying capacities of supplemented and non-supplemented streams can confound the effects of supplementation on numbers of juveniles per redd. For example, if the supplemented population is at or above carrying capacity and the non-supplemented population is not, numbers of juveniles per redd in the non-supplemented population may be significantly greater than the number of juveniles per redd in the supplemented population. In addition, pHOS may be correlated with overall spawner abundance. In these cases, it is difficult or impossible to separate density-dependent effects from the influence of pHOS on freshwater productivity. To avoid concluding that the supplementation program has no effect or perhaps a negative effect on juveniles per redd, the capacity of the habitats must be included in the analyses. The Supplementary Hypotheses presented below are designed to address the confounding effects of different densities on the analyses.

Monitoring Questions:

Q2.1.1: Has the supplementation program changed the number of juveniles (smolts, parr, and/or emigrants) per redd within the supplemented population?

Q2.2.1: Does the number of juveniles per redd decrease as the proportion of hatchery spawners increases?[6]

Target Species/Populations:

- Both Q2.1.1 and Q2.2.1 apply to all conservation stocks.

Statistical Hypotheses for 2.1.1[7]:

---

[6] Information is needed to estimate the effects of density dependence on these questions. Consider spatial distribution of redds.

[7] Quality and quantity of data will determine which hypotheses are evaluated. See Hillman et al. (2012) for details.

- $Ho_{2.1.1.1}$: Slope in juveniles/redd before supplementation $\leq$ slope in juveniles/redd after supplementation.
- $Ho_{2.1.1.2}$: Differences in slopes in juveniles/redd between supplemented and reference populations before supplementation $\leq$ differences in slopes in juveniles/redd between supplemented and reference populations after supplementation.
- $Ho_{2.1.1.3}$: Mean juveniles/redd before supplementation $\leq$ mean juveniles/redd after supplementation.
- $Ho_{2.1.1.4}$: Mean ratio scores in juveniles/redd before supplementation $\leq$ Mean ratio scores in juveniles/redd during supplementation.
- $Ho_{2.1.1.5}$: Mean ratio scores in juveniles/redd (adjusted for density dependence) before supplementation $\leq$ Mean ratio scores in juveniles/redd (adjusted for density dependence) during supplementation. [This hypothesis adjusts juveniles/redd for density-dependent effects (see Hillman et al. 2012 for details; Appendix 6).]
- $Ho_{2.1.1.6}$: There is no association between the proportion of hatchery-origin spawners (pHOS) and the residuals from the smooth hockey stick stock-recruitment curve; rho = 0. [If there is a significant negative association between pHOS and the residuals, then hatchery fish may be reducing the productivity of the wild population.]

Statistical Hypotheses for 2.2.1:

- $Ho_{2.2.1.1}$: There is no association between the proportion of hatchery-origin spawners (pHOS) and the residuals from the smooth hockey stick stock-recruitment curve; rho = 0. [If there is a significant negative association between pHOS and the residuals, then hatchery fish may be reducing the productivity of the wild population.]
- $Ho_{2.2.1.2}$: The slope between proportion of hatchery spawners and juveniles/redd is ≥ 0.

Measured Variables:

- Number of hatchery and naturally produced fish on spawning grounds.
- Numbers of redds.
- Number of juveniles (smolts, parr [where appropriate], and emigrants).

Derived Variables:

- Number of juveniles per spawner.
- Number of juveniles per redd.
- Carrying capacity (see Appendix 1).

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- These analyses shall be performed every five-years. Use graphic analyses, trend analyses, t-tests, Aspin-Welch tests, and randomization tests to evaluate the statistical hypotheses (see Hillman et al. 2012; Appendix 6). The specific analysis used will depend on the availability of reference conditions.
- Correlation analysis will examine associations between hatchery adult composition and juveniles/redd.

Analytical Rules:

- This is a productivity indicator that will be used to assess the success of the supplementation program.
- Type I Error of 0.05.

SECTION 4: NATURAL ENVIRONMENT MONITORING INDICATORS

4.1 Hatchery Replacement Rates (HRRs)

Objective 3: Determine if the hatchery adult-to-adult survival (i.e., hatchery replacement rate, HRR) is greater than the natural adult-to-adult survival (i.e., natural replacement rate, NRR) and the target hatchery survival rate.

The survival advantage from the hatchery (i.e., egg-to-smolt) must be sufficient to produce a greater number of returning adults than if broodstock were left to spawn naturally. If a hatchery program cannot produce a greater number of adults than naturally spawning fish, then the program should be modified or discontinued. Production levels were initially developed using historical run sizes and smolt-to-adult survival rates (BAMP 1998). Using the stock specific NRR and agreed upon target values (see Appendix 2), comparisons to actual survival rates will be made to ensure the expected level of survival has been achieved.

Monitoring Questions:

Q3.2.1: Is the adult-to-adult survival rate of hatchery fish (HRR) greater than or equal to the adult-to-adult survival rate (NRR) of naturally produced fish?

Q3.2.2: Is the adult-to-adult survival rate of hatchery fish (HRR) greater than or equal to the Target Value identified in Appendix 2[8]?

Target Species/Populations:

- Q3.2.1 applies to all conservation stocks.
- Q3.2.2 applies to all stocks.

Statistical Hypothesis 3.2.1:

- $Ho_{3.2.1.1}$: $HRR_{\text{Year x}} \geq NRR_{\text{Year x}}$

Statistical Hypothesis 3.2.2:

- $Ho_{3.2.2.1}$: HRR ≥ Target Value identified in Appendix 2

Measured Variables:

- Number of hatchery and naturally produced fish on spawning grounds.
- Number of hatchery and naturally produced fish harvested.
- Number of hatchery and naturally produced fish collected for broodstock.
- Number of broodstock used by brood year (hatchery and naturally produced fish).

Derived Variables:

- Number of hatchery and naturally produced adults by brood year (≥age-3).

---

[8] Target values may be adjusted by the hatchery committees.

- HRR (number of returning adults per brood year/broodstock)
- NRR (from Objective 1)
- Appendix 2: HRR targets identified in Appendix 2

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- For Q3.2.1 use graphic analysis and paired-sample quantile tests to compare HRR to NRR
- For Q3.2.2 use graphic analysis and one-sample quantile tests to compare HRR to the target value.
- On a five-year period, correlate HRRs with extraneous factors such as ocean productivity indices.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

4.2 Proportion of Hatchery-Origin Spawners (pHOS or PNI)

Objective 4: Determine if the proportion of hatchery-origin spawners (pHOS or PNI) is meeting management target.

Certain hatchery programs have pHOS or PNI targets, while other do not. HGMPs and permits inform the selection of targets, which are identified in Appendix 3.

Monitoring Questions:

Q4.1.1: Is the estimated proportion of hatchery-origin spawners (pHOS) less than or equal to the management target, and/or, is the estimated Percent Natural Influence (PNI) greater than or equal to the management target identified in Appendix 3?

Target Species/Populations:

- Q4.1.1 applies to all conservation and safety-net stocks that have a defined pHOS or PNI target or sliding scale (see Appendix 3).

Statistical Hypothesis 4.1.1:

- $Ho_{4.1.1.1}$: pHOS > target value or $PNI_{\text{Supplemented population}}$ < target value identified in Appendix 3

Measured Variables:

- Number of hatchery and naturally produced fish on spawning grounds

Derived Variables:

- pHOS or PNI
- Appendix 3: PNI and pHOS targets and sliding scales identified in Appendix 3

Spatial/Temporal Scale:

- Calculate annually.
- Analyzed as time series.

Possible Statistical Analysis:

- Use graphic analysis and summary statistics to compare pHOS or PNI to the target value in Appendix 3.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.

4.3 Run Timing, Spawn Timing, and Spawning Distribution

<u>Objective 5</u>: Determine if the run timing, spawn timing, and spawning distribution of the hatchery component is similar to the natural component of the target population or is meeting program-specific objectives.

Strategies for conservation programs typically intend that hatchery and natural-origin fish spawn together and in similar locations. However, in some cases, strategies may differ from this paradigm (e.g., summer Chinook salmon in the Wenatchee and Methow rivers; see Appendix 4). Phenotypic plasticity or selection resulting from the hatchery environment (i.e., domestication) may affect run (migration) timing, spawn timing, and spawning distribution. If conservation programs do not adequately represent the genetic diversity of the natural population, and if phenotypic traits in supplementation fish related to fitness deviate from the naturally produced spawning population, the goals of supplementation may not be achieved. Hatchery adults that migrate and/or spawn at different times or are spatially segregated from natural-origin fish may be subject to reduced fitness. Hatchery adults that spawn at different times or locations than natural-origin fish would be reproductively isolated from the natural population. The extent of such isolation, ranging from no isolation to substantial isolation, may be exploited for management purposes in some cases.

Migration Timing

Monitoring Questions:

Q5.1.1: Is the migration timing of hatchery and natural-origin fish from the same age class similar?

Target Species/Populations:

- Q5.1.1 applies to all conservation stocks.

Statistical Hypotheses 5.1.1:

- $Ho_{5.1.1.1}$: Migration timing $_{\text{Hatchery Age X}}$ = Migration timing $_{\text{Naturally produced Age X}}$

- $Ho_{5.1.1.2}$: The cumulative frequency of migration timing of hatchery-origin fish = the cumulative frequency of migration timing of natural-origin fish.
- $Ho_{5.1.1.3}$: The 10th percentile, 50th percentile (mode), 90th percentile, and mean migration timing of hatchery-origin fish = the 10th percentile, 50th percentile (median), 90th percentile, and mean migration timing of natural-origin fish.

Measured Variables:

- Ages of hatchery and natural-origin fish sampled via PIT tags or stock assessment monitoring.
- Time (Julian date) of arrival at mainstem projects and within tributaries (e.g., traps, PIT arrays) with the intent to identify biologically significant differences.

Derived Variables:

- Mean Julian date for a given age class.

Spatial/Temporal Scale:

- Calculate annually based on return year and age class.
- Time series.

Possible Statistical Analysis:

- Use graphic analyses (cumulative frequency polygons), paired t-tests, Aspin-Welch tests, and randomization tests.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Spawn Timing

Monitoring Questions:

Q5.2.1: Is the timing of spawning similar for conservation hatchery and natural-origin fish?

Target Species/Populations:

- Q5.2.1: Applies to all semelparous species and populations supplemented by conservation programs. Steelhead can only be assessed for natural spawning in situations where hatchery and natural-origin fish can be appropriately marked and detected.

Statistical Hypotheses 5.2.1:

- $Ho_{5.2.1.1}$: The cumulative frequency of spawn timing of hatchery-origin fish = the cumulative frequency of spawn timing of natural-origin fish.
- $Ho_{5.2.1.2}$: The 10th percentile, 50th percentile (mode), 90th percentile, and mean spawn timing of hatchery-origin fish = the 10th percentile, 50th percentile (mode), 90th percentile, and mean spawn timing of natural-origin fish.

- $Ho_{5.2.1.3}$: The relationship between elevation and spawn timing of hatchery-origin fish = the relationship between elevation and spawn timing of natural-origin fish.

Measured Variables:

- Time (Julian date) and elevation (m) of hatchery and natural-origin salmon carcasses or marked steelhead detected on spawning grounds within defined reaches.
- Time (Julian date) of ripeness of hatchery and natural-origin steelhead captured for broodstock.

Derived Variables:

- Mean Julian date.

Spatial/Temporal Scale:

- Calculate annually based on return year.
- Time series.

Possible Statistical Analysis:

- Use graphic analyses (cumulative frequency polygons), paired t-tests, Aspin-Welch tests, and randomization tests.
- Use graphic analyses, ANCOVA, and/or regression analysis to assess relationships between elevation and spawn timing.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Spatial Distribution of Redds

Monitoring Questions:

Q5.3.1: Is the distribution of redds similar for conservation hatchery and natural-origin fish?

Q5.3.2: Is the distribution of redds similar to defined management targets (see Appendix 4)?

Target Species/Populations:

- Q5.3.1 applies to all conservation program stocks.
- Q5.3.2 applies only to conservation program stocks with specific spawning distribution targets (Carlton and Dryden summer Chinook programs; Appendix 4).

Statistical Hypothesis 5.3.1:

- $Ho_{5.3.1.1}$: The distribution of hatchery-origin redds (hatchery females) = the distribution of natural-origin redds (natural-origin females).

Statistical Hypothesis 5.3.2:

- $Ho_{5.3.2.1}$: The distribution of hatchery-origin redds (hatchery females) = the target distribution identified in Appendix 4.

Measured Variables:

- Location (GPS coordinate) of female salmon carcasses observed on spawning grounds. The distribution of hatchery and naturally produced steelhead redds may be evaluated if marking or tagging efforts provide reasonable results.

Derived Variables:

- Location of female salmon carcass at the historic reach scale and at the 0.1 km scale.
- Calculate percent overlap in distribution across available spawning habitat or historical reaches.
- Appendix 4: Management targets for spatial distribution of spawners or redds (as applicable).

Spatial/Temporal Scale:

- Calculate annually based on return year.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and Yates' Chi-square analysis for both Q5.3.1 and Q5.3.2.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

4.4 Stray Rates

Objective 6: Determine if the recipient stray rate of hatchery fish is below the acceptable levels to maintain genetic variation among stocks.

Maintaining locally adapted traits among independent fish populations requires that returning hatchery fish have a high rate of site fidelity to the target population or stream. Hatchery practices (e.g., imprinting on water sources at key life history stages, release methodology, release location, age at return, broodstock used, spawner density, spawning habitat quality and access, and environmental conditions) are the main variables thought to affect stray rates. Regardless of the magnitude of homing of adult returns, if adult hatchery fish do not contribute to the natural population, the program will not meet the basic condition of a supplementation program.

Independent populations are populations that are genetically differentiated from other populations. In some cases, genetic differentiation may be assumed based on phenotypic traits or geographic isolation when molecular genetics analyses are not available. When populations are not independent, straying among them does not pose a risk of genetic homogenization. In addition, stray rates of hatchery-origin fish cannot be expected to be lower than for natural-origin fish. When estimates of stray rates for natural-origin fish are available and if they exceed the 5% among population stray rate or 10% within population stray rate thresholds identified in this plan, analysis and interpretation of stray rates must take into account the concept that hatchery programs may be held to unattainable standards based on the natural stray rate. Current criteria established by the ICBTRT (2005) and the Upper Columbia Spring Chinook Salmon and Steelhead Recovery Plan (2007) indicate that fish that do stray to other non-target populations should not comprise greater than 5% of the non-target spawning population. Likewise, fish that stray into non-target spawning areas within an independent population should not comprise greater than 10% of the non-target spawning aggregate (see Tables 6.1 and 6.2).

This plan identifies three stray rate metrics; brood-year stray rate, among population return-year stray rate, and within population return-year stray rate. The return-year stray rates have specific targets that are from the ICBTRT (2005) and Upper Columbia Spring Chinook Salmon and Steelhead Recovery Plan (2007), and are linked to extinction risk. That is, hatchery strays from other populations cannot make up more than 5% of the spawning escapement within a non-target, recipient population. In addition, hatchery strays from other spawning aggregations within a population (e.g., Chiwawa spring Chinook) cannot make up more than 10% of the spawning escapement within a non-target, recipient spawning aggregate (e.g., White River). Brood-year stray rate, on the other hand, is not discussed in the Upper Columbia Spring Chinook Salmon and Steelhead Recovery Plan (2007) or ICBTRT (2005) and therefore has no specific target. Nevertheless, it is important to track brood-year stray rates to determine if hatchery operations affect the homing and straying of specific brood years. These data support the return-year stray metrics and are used to inform possible changes in genetic variation among stocks.

Brood-Year Stray Rates

Monitoring Questions:

Q6.1.1: What is the brood-year stray rate of hatchery fish?

Target Species/Populations:

- Q6.1.1 applies to all hatchery stocks.

Statistical Hypothesis 6.1.1:

- $Ho_{6.1.1.1}$: None.

Measured Variables:

- Number of hatchery carcasses found in non-target and target spawning areas or number of returning spawners counted via PIT-tag detection or at weirs in close temporal proximity to spawning areas.
- Number of hatchery fish collected for broodstock.
- Number of hatchery fish taken in fishery.
- Locations of live and dead strays (used to tease out overshoot).

Derived Variables:

- Total number of hatchery carcasses and take in fishery estimated from expansion analysis.
- Percent of the total brood return that strays.

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- Use graphical analysis to track brood-year stray rates over time.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.

Among-Population Return-Year Stray Rates

Monitoring Questions:

Q6.2.1: Do hatchery strays make up less than 5% of the spawning escapement within their non-target independent populations?

Target Species/Populations:

- Q6.2.1 applies to all hatchery stocks.

Statistical Hypothesis 6.2.1:

- $Ho_{6.2.1.1}$: Stray hatchery fish make up ≥ 5% of the spawning escapement (based on run year) within other independent populations [9]

Measured Variables:

- Number of hatchery carcasses (PIT-tagged steelhead) found in non-target and target spawning areas or number of returning spawners counted via PIT-tag detection or at weirs in close temporal proximity to spawning areas.

Derived Variables:

- Total number of hatchery salmon carcasses (PIT-tagged steelhead, spawners counted at weirs) estimated from expansion analysis.
- Percent of the non-target population that is made up of hatchery strays.

Spatial/Temporal Scale:

- Calculate annually based on return year.
- Time series.

Possible Statistical Analysis:

---

[9] This stray rate is suggested based on a literature review and recommendations by the ICBTRT (2005) and is identified in the Upper Columbia Spring Chinook Salmon and Steelhead Recovery Plan (2007). It can be re-evaluated as more information on naturally-produced Upper Columbia salmonids becomes available. This will be evaluated on a species and program specific basis and decisions made by the HCP HC and PRCC HSC. It is important to understand the actual spawner composition of the population to determine the potential effect of straying.

- Use graphical analysis and one-sample quantile tests to compare the estimated stray rate with the target (5%) stray rate.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Within-Population Return-Year Stray Rates

Monitoring Questions:

Q6.3.1: Do hatchery strays make up less than 10% of the spawning aggregate within non-target spawning areas within the target population?

Target Species/Populations:

- Q6.3.1 applies to all hatchery stocks.

Statistical Hypothesis 6.3.1:

- $Ho_{6.3.1}$: Stray hatchery fish make up ≥ 10% of spawning escapement (based on run year) within non-target spawning areas within the target population

Measured Variables:

- Number of hatchery carcasses (possibly PIT-tagged steelhead) found in non-target and target spawning aggregates or number of returning spawners counted via PIT-tag detection or at weirs in close temporal proximity to spawning areas.

Derived Variables:

- Total number of hatchery salmon carcasses (possibly PIT-tagged steelhead or spawners counted at weirs) estimated from expansion analysis.
- Percent of the non-target spawning aggregate that is made up of hatchery strays.

Spatial/Temporal Scale:

- Calculate annually based on return year.
- Time series.

Possible Statistical Analysis:

- Use graphical analysis and one-sample quantile tests to compare the estimated stray rate with the target (10%) stray rate.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

4.5 Population Genetics

Objective 7: Determine if genetic diversity, population structure, and effective population size have changed in natural spawning populations as a result of the conservation and safety-net hatchery programs and assess genetic changes of hatchery-origin returns.

The genetic component of the M&E Plan specifically addresses the potential for changes in genetic diversity in natural populations as a result of safety-net and conservation hatchery programs. The long-term fitness of populations is assumed to be related to maintaining the genetic diversity of natural populations. However, hatchery programs select a subset of individuals from the population to pass on genetic material to the next generation. This is often a relatively small number of individuals that produce a large number of offspring and can result in changes in allele frequencies and reductions of effective population size. Therefore, it is important to monitor the genetic status of the natural populations to determine if there are signs of changes in genetic distance among populations, changes in allele frequencies, linkage disequilibrium, and to estimate effective population size. Assessing the genetic effects of the hatchery program on natural populations does not require annual sampling but does require regular sampling at generational time scales. Additionally, genetic status of hatchery-origin returns should be monitored to further evaluate potential genetic risks to the natural population. Beginning with brood years 2017 and 2018 (and then at subsequent ten year intervals), testing statistical hypotheses associated with genetic components (Hypotheses 7.1.1, 7.2.1, 7.3.1, and 7.4.1) will be conducted with natural-origin baseline samples (the earliest genetic samples available for each program), natural-origin contemporary samples, and hatchery-origin contemporary samples.

Alternatively, genetic divergence between the hatchery-origin population and natural-origin population is expected with segregated programs. Monitoring of genetic risks associated with utilizing multi-generations of hatchery-origin broodstock is important. Beginning with brood years 2017 and 2018 (and then at subsequent ten-year intervals), testing statistical hypotheses associated with genetic components (Hypotheses 7.1.2 and 7.2.2) will be conducted with hatchery-origin baseline samples (the earliest genetic samples available for each program) and hatchery-origin contemporary samples.

An alternative analysis to statistical hypotheses testing is equivalence testing, which requires determination of biologically relevant effects. To date, biologically relevant effects for measured differences in genetic metrics have not been determined; when they are determined, equivalence testing will also be conducted.

Allele Frequency

Monitoring Questions:

Q7.1.1: Is the contemporary allele frequency of natural-origin fish and hatchery-origin fish similar to the baseline allele frequency of natural-origin fish over time?

Q7.1.2: Is the contemporary allele frequency of hatchery-origin broodstock similar to the baseline allele frequency of hatchery-origin broodstock over time?

Target Species/Populations:

- Q7.1.1 Applies to all conservation and safety net programs. .

- Q7.1.2 Applies to all segregated programs.

Statistical Hypotheses 7.1.1 and 7.1.2:

- $Ho_{7.1.1.1}$: Allele frequency natural-origin$_{t=baseline}$ = Allele frequency natural-origin$_{t=contemporary}$ = Allele frequency hatchery-origin$_{t=contemporary}$
- $Ha_{7.1.1.1}$: Allele frequency natural-origin$_{t=baseline}$ ≠ Allele frequency natural-origin$_{t=contemporary}$ ≠ Allele frequency hatchery-origin$_{t=contemporary}$
- $Ho_{7.1.1.2}$: Allele frequency hatchery-origin$_{t=baseline}$ = Allele frequency hatchery-origin$_{t=contemporary}$
- $Ha_{7.1.1.2}$: Allele frequency hatchery-origin$_{t=baseline}$ ≠ Allele frequency hatchery-origin$_{t=contemporary}$

Measured Variables:

- SNP genotypes

Derived Variables:

- Allele frequency

Spatial/Temporal Scale:

- Analyze as a time series, initially comparing contemporary samples to baseline samples, and then at a ten-year frequency.
- Compare conservation and safety-net program samples within subpopulations (when applicable), populations, and the upper Columbia.

Possible Statistical Analysis:

- Population differentiation tests, analysis of molecular variance (AMOVA), relative genetic distances, or suitable equivalence tests.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Linkage Disequilibrium

Monitoring Questions:

Q7.2.1.: Is contemporary linkage disequilibrium (LD) of natural-origin and hatchery-origin fish similar to the baseline LD of natural-origin fish over time?

Q7.2.2.: Is contemporary linkage disequilibrium (LD) of hatchery-origin broodstock similar to the baseline LD of hatchery-origin broodstock over time?

Target Species/Populations:

- Q7.2.1 Applies to all safety-net and conservation programs.

- Q7.2.2 Applies to all segregated programs.

Statistical Hypotheses 7.2.1 and 7.2.2:

- $Ho_{7.2.1.1}$: LD natural-origin$_{t=baseline}$=LD natural-origin$_{t=contemporary}$ = LD hatchery-origint=contemporary
- $Ha_{7.2.1.1}$: LD natural-origin$_{t=baseline}$ ≠ LD natural-origin$_{t=contemporary}$ ≠ LD hatchery-origin$_{t=contemporary}$
- $Ho_{7.2.1.2}$: LD hatchery-origin$_{t=baseline}$ = LD hatchery-origin$_{t=contemporary}$
- $Ha_{7.2.1.2}$: LD hatchery-origin$_{t=baseline}$ ≠ LD hatchery-origin$_{t=contemporary}$

Measured Variables:

- SNP genotypes

Derived Variables:

- Pairwise by locus tests of LD
- Counts or percentages of pairwise tests with statistically significant LD before and after correction for multiple tests

Spatial/Temporal Scale:

- Analyze as a time series, initially comparing contemporary samples to baseline samples, and then at a ten year frequency.
- Compare conservation and safety-net program samples within subpopulations (when applicable), populations, and the upper Columbia.

Possible Statistical Analysis:

- Probability testing for pairwise by locus LD, Chi-squared tests, Wilcoxon signed rank test, or suitable equivalence tests.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Genetic Distance Between Subpopulations and Populations

Monitoring Questions:

Q7.3.1: Does the genetic distance among supplemented subpopulations (where applicable) and populations remain the same over time?

Target Species/Populations:

- Q7.3.1 applies to all safety-net and conservation programs.

Statistical Hypothesis 7.3.1:

- $Ho_{7.3.1.1}$: Genetic distance between sub-populations$_{Year\ t=contemporary}$ = Genetic distance between subpopulations$_{Year\ t=baseline}$

- $Ho_{7.3.1.1}$: Genetic distance between $\text{populations}_{\text{Year } t=\text{contemporary}} \neq$ Genetic distance between $\text{populations}_{\text{Year } t=\text{baseline}}$

Measured Variables:

- SNP genotypes

Derived Variables:

- Allele frequencies

Spatial/Temporal Scale:

- Analyze comparing contemporary samples to baseline samples.
- Compare samples between sub-populations (where applicable) and populations.

Possible Statistical Analysis:

- Population differentiation tests, AMOVA, relative genetic distances, or suitable equivalence tests.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Effective Spawning Population

Monitoring Questions:

Q7.4.1: Is the ratio of effective population size (N_e) to spawning population size (N) constant or increasing over time?

Target Species/Populations:

- Q7.4.1 applies to all programs.

Statistical Hypothesis 3.3:

- $Ho_{7.4.1.1}$: $(N_e/N)_{t=\text{contemporary}} \geq (N_e/N)_{t=\text{baseline}}$ for each population.
- $Ha_{7.4.1.1}$: $(N_e/N)_{t=\text{contemporary}} < (N_e/N)_{t=\text{baseline}}$ for each population.

Measured Variables:

- SNP genotypes and estimates of N

Derived Variables:

- Allele frequencies

Spatial/Temporal Scale:

- Compare contemporary samples to baseline samples.

Possible Statistical Analysis:

- Statistics to calculate effective population size (e.g., harmonic means), Wilcoxon signed rank tests, or suitable equivalence tests.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

4.6 Phenotypic Traits

Objective 8: Determine if hatchery programs have caused changes in phenotypic characteristics of natural populations.

Fitness, or the ability of individuals to survive and pass on their genes to the next generation in a given environment, includes genetic, physiological, and behavioral components.[10] Maintaining the long-term fitness of supplemented populations requires a comprehensive evaluation of genetic and phenotypic characteristics. Evaluation of some phenotypic traits (i.e., run timing, spawn timing, spawning location, and stray rates) is addressed under Objective 5. Objective 8 assess the potential effects of domestication, including size at maturity, age at maturity, sex ratio, and fecundity. Age and size at maturity shall be assessed for both fish arriving in the Columbia system, and those recovered on the spawning grounds. Size (or age) selective mortality during migration through the Columbia system, such as through fisheries, could alter the age and size of fish on the spawning grounds.

Age at Maturity

Monitoring Questions:

Q8.1.1: Is the age at maturity of hatchery and natural-origin fish similar at the time they enter the Columbia River and when they spawn?

Target Species/Populations:

- Q8.1.1 applies to all conservation program stocks.

Statistical Hypotheses 8.1.1:

- $Ho_{8.1.1.1}$: Age at Maturity $_{\text{Hatchery produced spawners Gender X}}$ = Age at Maturity $_{\text{Naturally produced spawners Gender X}}$
- $Ho_{8.1.1.2}$: Age at Maturity $_{\text{All hatchery produced adults Gender X}}$ = Age at Maturity $_{\text{All naturally produced adults Gender X}}$

Measured Variables:

- Total and salt (ocean) age of hatchery and natural-origin salmon carcasses collected on spawning grounds.
- Total and salt age of broodstock.
- Total and salt age of fish at stock assessment locations (e.g., Dryden, Tumwater, Wells, Priest Rapids).

---

[10] These metrics are difficult to measure, and phenotypic expression of these traits may be all we can measure and evaluate.

- Whenever possible, age at maturity will be measured at weirs or dams near the spawning stream to avoid the size-related carcass recovery bias on spawning grounds (carcass sampling).
- Assess age of fish, including harvested fish.

Derived Variables:

- Total age and saltwater age
- Age of fish entering the Columbia River.

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and Yates' Chi-square.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Size at Maturity

Monitoring Questions:

Q8.2.1: Is the size (length) at maturity of a given age and sex of hatchery fish similar to the size at maturity of a given age and sex of natural-origin fish?

Target Species/Populations:

- Q8.2.1 applies to all conservation and safety-net stocks.

Statistical Hypothesis 8.2.1:

- Ho$_{8.2.1.1}$: Size (length) at Maturity $_{\text{Hatchery Age X and Gender Y}}$ = Size (length) at Maturity $_{\text{Naturally produced Age X and Gender Y}}$
- Ho$_{8.2.1.2}$: Size (length) at Maturity $_{\text{All hatchery adults Gender X}}$ = Size (length) at Maturity $_{\text{All naturally produced adults Gender X}}$

Measured Variables:

- Size (length), age, and gender of hatchery and natural-origin salmon carcasses collected on spawning grounds.
- Size (length), age, and gender of broodstock.
- Size (length), age, and gender of fish at stock assessment locations (e.g., Priest Rapids, Dryden, Tumwater, Wells, Twisp Weir).

- Whenever possible, size at maturity will be measured at weirs or dams near the spawning stream to avoid the size-related carcass recovery bias on spawning grounds (carcass sampling).

Derived Variables:

- Total age and saltwater age

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and three-way ANOVA by origin, gender, and age

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Fecundity at Size[11]

Monitoring Questions:

Q8.3.1: Is the fecundity vs. size relationship of hatchery and natural-origin fish similar?

Q8.3.2: Is the gonadal mass vs. size relationship of hatchery and natural-origin fish similar?

Target Species/Populations:

- Both Q8.3.1 and Q8.3.2 apply to all conservation stocks using both natural- and hatchery-origin broodstock.

Statistical Hypothesis 8.3.1:

- $Ho_{8.3.1.1}$: Slope of Fecundity vs. Size $_{\text{Hatchery}}$ = Slope of Fecundity vs. Size $_{\text{Naturally produced}}$

Statistical Hypothesis 8.3.2:

- $Ho_{8.3.2.1}$: Gonadal Mass vs. Size $_{\text{Hatchery}}$ = Gonadal Mass vs. Size $_{\text{Naturally produced}}$

Measured Variables:

- Length, weight, and age (covariate) of hatchery and natural-origin broodstock after eggs have been removed.
- Number and weight of eggs

Derived Variables:

- Total age and saltwater age.

---

[11] May not apply to all programs.

- Mean weight per egg.

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis, regression, t-test, and ANCOVA.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Sex Ratio

Monitoring Questions:

Q8.4.1: Is the sex ratio of hatchery and natural-origin fish similar?

Target Species/Populations:

- Q8.4.1 applies to all conservation stocks.

Statistical Hypothesis 8.4.1:

- $Ho_{8.4.1.1}$: Sex Ratio $_{\text{Hatchery}}$ = Sex Ratio $_{\text{Naturally produced}}$

Measured Variables:

- Age and sex of hatchery and natural-origin salmon carcasses collected on spawning grounds or sampled at dams or weirs.
- Whenever possible sex ratio will be measured at weirs or dams near the spawning stream to avoid the size-related carcass recovery bias on spawning grounds (carcass sampling or ultrasound on live fish).

Derived Variables:

- Ratio of sexes based on brood year returns

Spatial/Temporal Scale:

- Calculate annually based on brood year.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and Yates'Chi-square.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

SECTION 5: HATCHERY ENVIRONMENT MONITORING INDICATORS

5.1 Release Targets

Objective 9: Determine if hatchery fish were released at the programmed size and number.

The HCP outlines the number and size of fish that are to be released to meet NNI and inundation compensation levels. The size of the fish at release may be altered according to an adaptive management process in the Hatchery Committee(s), and the number of fish can be altered by survival study results and adjustment of hatchery production for population dynamics. Size of fish at release can affect survival, sex ratios, age at return, stray rate, and fecundity. In addition, the variation in size at release may affect performance of the fish.

The coefficient of variation (CV) will be evaluated to ascertain if program performance is related to variation in size at release. Note also that variation in a population is a natural condition and striving to control this variation could result in directional or stabilizing artificial selection that could have unforeseen long-term consequences. Attaining uniform or multi-modal growth in a hatchery environment may not be adaptive for fitness in the wild. Therefore, pursuit of a CV target should be seen as an informative exercise, but is not in itself indicative of success or failure of a hatchery program. Furthermore, growth regimes may prove to be important in affecting adult returns and age structure. Although many factors can influence both the size and number of fish released, past hatchery cultural experience with these stocks should assist in meeting program production levels. Appendix 5 presents the target size at release and CVs for the programs. These targets shall be assessed annually to ensure they are optimized to inform management decisions.

Size at Release of Hatchery Fish

Monitoring Questions:

Q9.1.1: Is the size (fish per pound; fpp) of hatchery fish released equal to the program goal identified in Appendix 5?

Target Species/Populations:

- Q9.1.1 applies to all hatchery stocks.

Statistical Hypothesis 9.1.1:

- $Ho_{9.1.1.1}$: Hatchery fish $_{\text{fpp at release}}$ = Programmed $_{\text{fpp at release}}$ (see Appendix 5)

Measured Variables:

- Fork length and weights of random samples of hatchery juveniles at release.

Derived Variables:

- Mean length (FL), mean weight, and fish per pound
- Appendix 5: Rearing targets

Spatial/Temporal Scale:

- Calculate annually.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and descriptive statistics to compare the estimated fpp of hatchery fish at time of release with the program goal.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.

Coefficient of Variation (CV) of Hatchery Fish Released

Monitoring Questions:

Q9.2.1: Is the CV of hatchery fish released equal to the program target identified in Appendix 5?

Target Species/Populations:

- Q9.2.1 applies to all hatchery stocks.

Statistical Hypothesis 9.2.1:

- $Ho_{9.2.1.1}$: Hatchery fish $CV_{\text{at release}}$ = Programmed CV in Appendix 5

Measured Variables:

- Length and weights of random samples of hatchery smolts.

Derived Variables:

- Coefficient of Variation: $cv = (1 + 1/4n) \times (s/x)$ (where s = standard deviation, x = estimated mean, n = sample size)
- Appendix 5: Rearing targets

Spatial/Temporal Scale:

- Calculate annually.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and descriptive statistics to compare the estimated CV of size of hatchery fish released with the program goal.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.

Condition Factor (K) of Hatchery Fish Released

Monitoring Questions:

Q9.3.1: Is the K of hatchery fish released equal to the program target identified in Appendix 5?

Target Species/Populations:

- Q9.3.1 applies to all hatchery stocks.

Statistical Hypothesis 9.3.1:

- $Ho_{9.3.1.1}$: Hatchery fish $K_{\text{at release}}$ = Programmed K identified in Appendix 5

Measured Variables:

- Monthly individual lengths and weights of random samples of hatchery juveniles.

Derived Variables:

- Condition Factor: $K = W/L^3 \times 10^5$

Spatial/Temporal Scale:

- Calculate annually.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and descriptive statistics to compare the estimated K of released hatchery fish with the program goal.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

Number of Hatchery Fish Released

Monitoring Questions:

Q9.4.1: Is the number of hatchery fish released equal to the program goal identified in Appendix 5?

Target Species/Populations:

- Q9.4.1 applies to all hatchery stocks.

Statistical Hypothesis 9.4.1:

- $Ho_{9.4.1.1}$: Hatchery Fish Number = Programmed Number identified in Appendix 5

Measured Variables:

- Numbers of smolts released from the hatchery.

Derived Variables:

- Appendix 5: Rearing targets

Spatial/Temporal Scale:

- Calculate annually.

- Time series.

Possible Statistical Analysis:

- Use graphic analysis and one-sample quantile tests to compare the estimated number of hatchery fish released with the program goal.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

SECTION 6: HARVEST MONITORING INDICATORS

6.1 Harvest Rates

Objective 10: Determine if appropriate harvest rates have been applied to conservation, safety-net, and segregated harvest programs to meet the HCP/SSSA goal of providing harvest opportunities while also contributing to population management and minimizing risk to natural populations.

Harvest will be applied to different types of programs in an effort to achieve the management objectives of those programs. Programs designed to augment harvest should routinely contribute to harvest at a rate that greatly reduces the incidence of straying to natural spawning grounds, but also allows the program to be sustained. Safety-net programs may be harvested as part of an adult management strategy to minimize excessive escapement of hatchery-origin fish to spawning grounds. Similarly, conservation programs may undergo harvest to manage returning adults, but the emphasis for these programs should be to achieve escapement goals. In all cases, harvest effort should not have the unintended consequence of removing excessive numbers of conservation or natural-origin fish. In years when the expected returns of hatchery adults are above the level required to meet program goals (i.e., supplementation of spawning populations and/or brood stock requirements), surplus fish may be available for harvest. The M&E Plan specifically addresses harvest and harvest opportunities upstream of Priest Rapids Dam. Harvest or removal of surplus hatchery fish from the spawning grounds may assist in reducing potential adverse ecological and genetic impacts to natural populations (e.g., loss of genetic variation within and between populations, loss of fitness, reduced effective population size, and density-dependent effects).

Monitoring Questions:

Q10.1.1: Conservation Programs: Is the harvest on conservation hatchery fish at an appropriate level to manage natural spawning of conservation hatchery fish but low enough to sustain the hatchery program?

Q10.1.2: Safety-Net Programs: Is the harvest on conservation hatchery fish at an appropriate level to manage natural spawning of safety-net hatchery fish but low enough to sustain the hatchery program?

Q10.1.3: Is the harvest on hatchery fish produced from harvest-augmentation programs high enough to manage natural spawning but low enough to sustain the hatchery program?

Q10.1.4: Is the escapement of fish from conservation and safety-net programs in excess of broodstock and natural production[12] needs to provide opportunities for terminal harvest?

Target Species/Populations:

- Q10.1.1 applies to conservation programs.
- Q10.1.2 applies to safety-net programs.

---

[12] The current best estimates of carrying capacity (maximum recruits) will be used, as available.

- Q10.1.3 applies harvest augmentation programs.
- Q10.1.4 applies to conservation and safety-net programs.

Statistical Hypothesis 10.1.1:

- $Ho_{10.1.1.1}$: Harvest rate ≤ Maximum level to meet program goals

Statistical Hypothesis 10.1.2:

- $Ho_{10.1.2.1}$: Harvest rate ≤ Maximum level to meet program goals

Statistical Hypothesis 10.1.3:

- $Ho_{10.1.3.1}$: Escapement ≤ Maximum level to meet supplementation goals

Statistical Hypothesis 10.1.4:

- $Ho_{10.1.4.1}$: Harvest rate ≤ Maximum level to meet program goals

Measured Variables:

- Numbers of hatchery fish taken in harvest.
- Numbers of natural-origin fish taken in harvest.

Derived Variables:

- Total harvest by fishery estimated from expansion analysis.

Spatial/Temporal Scale:

- Calculated annually.
- Time series.

Possible Statistical Analysis:

- Use graphic analysis and one-sample quantile tests to compare the estimated harvest of hatchery fish with the program goal.

Analytical Rules:

- This is a monitoring indicator that will be used to support management decisions.
- Type I Error of 0.05.

SECTION 7: REGIONAL OBJECTIVES

Hatchery programs have the potential to increase diseases that typically occur at low levels in the natural environment (Objective 9). In addition, hatchery fish can reduce the abundance, size, or distribution of non-target taxa through ecological interactions (Objective 10). In this section, we address incidence of disease and non-target taxa of concern.

7.1 Incidence of Disease

Objective 11: Determine if the incidence of disease has increased in the natural and hatchery populations.

The hatchery environment has the potential to amplify diseases that are typically found at low levels in the natural environment. Amplification could occur within the hatchery population (i.e., vertical and horizontal transmission) or indirectly from the hatchery effluent or commingling between infected and non-infected fish (i.e., horizontal transmission). Potential impacts to natural populations have not been extensively studied, but should be considered for programs in which the hatchery fish are expected to commingle with natural fish. This is particularly important for supplementation type programs. Specifically, the causative agent of bacterial kidney disease (BKD), *Renibacterium salmoninarum* (Rs), could be monitored at selected acclimation ponds, both in the water and fish, in which the risk and potential for transmission from the hatchery is highest. Although it is technologically possible to measure the amount of Rs in water or Rs DNA in smolts and adults non-lethally sampled, the biological meaning of these data are uncertain. Currently, the only metric available for M & E purposes is measuring the antigen level from kidney/spleen samples (i.e., ELISA, PCR). When available, non-lethal sampling may replace or be used in concert with lethal sampling.

Implementation of this objective will be conducted in a coordinated approach within the hatchery and natural environment. BKD management within the hatchery population (e.g., broodstock or juveniles) has the potential to reduce the prevalence of disease through various actions (e.g., culling or reduced rearing densities). BKD management must also take into account and support other relevant objectives of the M & E program (e.g., Hatchery Return Rate [HRR], number of smolts released). Hence, the goal of BKD management is to decrease the prevalence of disease and maintain hatchery production objectives (i.e., number and HRR).

As previously discussed, disease transmission from hatchery to naturally produced fish may occur at various life stages and locations. Of these, horizontal transmission from hatchery effluent, vertical transmission on the spawning grounds, and horizontal transmission in the migration corridor have been identified as disease interactions that could be examined under this objective, although others may also be relevant. Experimental designs addressing this objective may require technology not yet available, although in some instances samples may be collected, but not analyzed until a link can be established between bacteria levels in samples and disease prevalence.

Developing a complete set of questions and hypotheses statements for this objective may not be practical at this time, because there is currently no BKD Management Plan. However, while developing experimental designs for this objective, it may be feasible to incorporate both hatchery and natural environment monitoring under a single study design. Integration of the different aspects of the objective would likely result in a more robust approach into understanding the effectiveness of disease management strategies.

Proposed Tasks:

T1: Assemble fish health data for fish used as brood (e.g., ELISA results).

T2: Conduct data exploration exercise to identify potential relationships between pathogen profiles and likely causative variables (e.g., rearing conditions and management actions).

T3: Develop hypotheses for potential testing to meet objective.

7.2 Non-Target Taxa of Concern (NTTOC)

Objective 12: Determine if the release of hatchery fish affects non-target taxa of concern (NTTOC) within acceptable limits.

Hatchery programs have the potential to affect non-target taxa through various types of interactions (e.g., competition and predation). These interactions can reduce the distribution, size, and abundance of non-target species. The non-target taxa of concern (NTTOC) ecological risk assessment was developed as a regional objective that would addressed ecological interactions on non-target taxa.

In 2008, the Wells HCP, Rocky Reach HCP, Rock Island HCP Hatchery Committees, and the Priest Rapids Hatchery Sub-Committee agreed to an approach to evaluate the potential effects of hatchery programs on NTTOC. The committees originally planned to convene a panel of experts to conduct a preliminary evaluation of the potential effects of Plan supplemented species on NTTOC. At the 15 October 2008 Hatchery Committees meeting, the members agreed to convene an expert panel to conduct a preliminary evaluation of potential effects of supplemented Plan Species on non-target taxa using an approach similar to that used in the Yakima Basin (Pearsons and Hopley 1999; Ham and Pearsons, 2001). The Committees agreed to convene the panel in spring or early summer 2009, and focus this initial effort on HCP Plan Species and the two non-Plan Species, westslope cutthroat trout and lamprey. The Committees identified species interactions, containment objectives for non-target species, and fisheries professionals who possessed the expertise to contribute as panel members. The Committees directed the Hatchery Evaluation Technical Team (HETT) to pursue assessment of the hatchery programs potential effects on NTTOC.

The HETT evaluated methods to conduct a risk assessment on NTTOC, and proposed using a combined modeling and a Delphi panel approach, whereby the modeling results would be compared and correlated with the Delphi panel results. The HETT identified the PCD Risk 1 model (Busack et al., 2005; Pearsons and Busack, 2012) to conduct the modeling evaluation. The PCD Risk 1 model is a data intensive, individual-based stochastic model. The HETT determined that the assembled data to be used as inputs for the PCD Risk 1 model would also serve to provide expert panelists the necessary data for them to conduct risk assessments. Hence, the HETT embarked on an extensive effort to gather, organize, and extract the required data from existing datasets, literature, and biologists familiar with the programs and/or particular NTTOC. Ultimately, the input data were assembled in a relational database that allowed the data to be output in user-friendly formats for modeling or Delphi panel use. The database also served to hold the modeling results, which could be extracted and summarized as needed. Following the modeling work, the Committees decided not to assemble the expert panel, because the panel would not be able to evaluate adequately the very large number of possible interactions.

A report titled *Ecological Risk Assessment of Upper-Columbia Hatchery Programs on Non-Target Taxa of Concern* was drafted in 2013 and finalized in 2014, which included the modeling results to date. The results in the report represent a very extensive effort to model the risk of all the upper Columbia hatchery programs for the identified NTTOC for which data and model runs were available. Should new information become available, the Committees agreed to assess the suitability of the data as it relates to conducting future NTTOC evaluations as a regional objective.

SECTION 8: ADAPTIVE MANAGEMENT

One of the challenges of evaluating PUD hatchery programs is that hatchery programs are modified resulting in hatchery treatments that are uneven throughout the duration of the hatchery program. Modifications occur as a result of recalculating hatchery release numbers every 10 years and also through adaptive management. To solve this evaluation challenge, we propose to conduct two scales of analysis. First, the entire duration of the program will be analyzed using the entire data set. This evaluation will analyze whether the overall adaptively managed program achieved objectives. Second, where appropriate, analyses will be compared across periods or programs to determine if major program changes have resulted in hypothesized changes to key response variables. We acknowledged that partitioning data into shorter periods will likely result in reduced statistical power so only the biggest changes will be evaluated. In the future, the hatchery committees will develop a table or figure that identifies major program changes in fish culture or M&E.

In the past, hatchery programs have been evaluated at the hatchery program scale (e.g., Nason Creek, Carlton summer Chinook). In some cases, it may be worthwhile to evaluate supplementation programs at different spatial scales. For example, the Nason Creek spring Chinook salmon program can be evaluated at the scale of Nason Creek, the combined effects of spring Chinook hatchery programs in the Wenatchee basin at the Wenatchee basin scale, and then all of the spring Chinook programs in the upper Columbia at the upper Columbia basin scale.

Comparisons of supplemented populations (treatments) to in-basin reference populations are the best way to evaluate whether treatments have caused changes to variables such as natural-origin recruits or productivity. Many suitable out-of-basin references are available (see Appendix 6), but these references do not control for unique factors that may be happening in the upper Columbia or areas outside the upper Columbia. For example, large fires that occur in the Upper Columbia may not occur at similar times in areas outside of this area. Candidate in-basin reference populations are not ideal for spring Chinook salmon because they are small and are above a lake (e.g., Little Wenatchee River) or they have had a long history of hatchery stocking (e.g., Entiat River). Every population of upper Columbia summer and fall Chinook is supplemented so in-basin references are not currently available. Without a suitable number of in-basin reference populations that are similar in size and distribution to treated populations, it will be difficult to unambiguously assess hatchery effects on certain variables. Although not ideal, the only way to increase in-basin reference comparisons is to strategically reduce the number of places where hatchery fish are released such as was done for the Entiat River.

Previous stocking history will lessen the value of reference populations; however, they can still be of value. For instance, the Committees can still test whether NORs are increased under supplementation compared to periods when other populations are not supplemented (i.e., a reverse BACI analysis).

SECTION 9: REFERENCES

Busack, C. A., K. P. Currens, T. N. Pearsons, and L. M. Mobrand. 2005. Tools for evaluating ecological and genetic risks in hatchery programs. Final report to Bonneville Power Administration (project 2003-058-00, contract BPA00016399).

Ham, K. D., and T. N. Pearsons. 2001. A practical approach for containing ecological risks associated with fish stocking programs. Fisheries 25(4):15-23.

Hays, S., T. Hillman, T. Kahler, R. Klinge, R. Langshaw, B. Lenz, A. Murdoch, K. Murdoch, and C. Peven. 2006. Analytical framework for monitoring and evaluating PUD hatchery programs. Prepared for the Habitat and Conservation Plans Hatchery Committee. Wenathcee, Washington.

Hillman, T., M. Miller, J. Murauskas, S. Hays, J. Miller, A. Murdoch, and T. Miller. 2012. Monitoring and evaluation of the Chelan County PUD hatchery programs. Prepared for the HCP Hatchery Committee, Wenatchee, WA.

Hillman, T., A. Murdoch, T. Pearsons, M. Miller, and G. Mackey. 2011. Methods for identifying reference populations and testing differences in abundance and productivity between reference populations and supplemented populations: Chiwawa spring Chinook case study. Prepared for the HCP Hatchery Committee, Wenatchee, WA.

ICBTRT (Interior Columbia Basin Technical Recovery Team). 2005. Interior Columbia Basin TRT: viability criteria for application to interior Columbia Basin salmonid ESUs. National Marine Fisheries Service, Northwest Fisheries Science Center. Seattle, WA. July 2005.

Mackey, et al. 2014. Ecological Risk Assessment of Upper-Columbia Hatchery Programs on Non-Target Taxa of Concern. Final Report to HCP Hatchery Committees.

Pearsons, T. N., and R. B. Langshaw. 2009a. Monitoring and evaluation plan for Grant PUDs salmon and steelhead supplementation Programs. Grant County Public Utility District, Ephrata, Washington.

Pearsons, T. N., and R. B. Langshaw. 2009b. Priest Rapids Hatchery Fall Chinook Salmon monitoring and evaluation plan. Grant County Public Utility District, Ephrata, Washington.

Pearsons, T. N. and C. A. Busack. 2012. PCD Risk 1: A tool for assessing and reducing ecological risks of hatchery operations in freshwater. Environmental Biology of Fishes 94:45-65. DOI:10.1007/s10641-011-9926-8.

Pearsons, T. N., and C. W. Hopley. 1999. A practical approach for assessing ecological risks associated with fish stocking programs. Fisheries 24(9):16-23.Murdoch, A. and C. Peven. 2005. Conceptual approach to monitoring and evaluating the Chelan County Public Utility District hatchery programs. Prepared for the Chelan PUD Habitat Conservation Plan's Hatchery Committee. Wenatchee, Washington.

Murdoch, A., C. Snow, C. Frady, A. Repp, M. Small, S. Blankenship, T. Hillman, M. Miller, G. Mackey, and T. Kahler. 2012. Evaluation of Hatchery Programs Funded by Douglas County PUD, 5-Year Report, 2006-2010. Prepared for the Wells HCP Hatchery Committee.

SECTION 10: GLOSSARY

Adult-to-Adult survival (Ratio)	The number of parent broodstock relative to the number of returning adults.
Age at maturity	The age of fish at the time of spawning (hatchery or naturally).
Augmentation	A hatchery strategy where fish are released for the sole purpose of providing harvest opportunities.
Broodstock	Adult salmon and steelhead collected for hatchery fish egg harvest and fertilization.
Donor population	The source population for supplementation programs before hatchery fish spawned naturally.
Effective population size (Ne)	The number of reproducing individuals in an ideal population (i.e., Ne = N) that would lose genetic variation due to genetic drift or inbreeding at the same rate as the number of reproducing adults in the real population under consideration (Hallerman 2003).
ESA	Endangered Species Act passed in 1973. The ESA-listed species refers to fish species added to the ESA list of endangered or threatened species and are covered by the ESA.
Expected value	The number of smolts or adults derived from survival rates agreed to in the Biological Assessment and Management Plan (BAMP 1998).
Extraction rate	The proportion of the spawning population collected for broodstock.
Genetic diversity	All the genetic variation within a species of interest, including both within and between population components.
Genetic stock structure	A type of assortative mating, in which the gene pool of a species is composed of a group of subpopulations, or stocks, that mate panmictically within themselves.
Genetic variation	All the variation due to different alleles and genes in an individual, population, or species.
HCP	Habitat Conservation Plan is a plan that enables an individual or organization to obtain a Section 10 Permit which outlines what will be done to “minimize and mitigate” the impact of the permitted take on a listed species.

HCP-HC	Habitat Conservation Plan Hatchery Committee is the committee that directs actions under the hatchery program section of the HCP's for Chelan and Douglas PUDs.
HRR	Hatchery Replacement Rate is the ratio of the number of returning hatchery adults relative to the number of adults taken as broodstock, both hatchery and naturally produced fish (i.e., adult-to-adult replacement rate).
Long-term fitness	Long-term fitness is the ability of a population to self-perpetuate over successive generation.
Naturally produced	Progeny of fish that spawned in the natural environment, regardless of the origin of the parents.
Mean Ratio	The ratio between a treatment and control population, with the mean taken across a time period, such as years. Used in analysis in Before-After-Control-Impact studies.
Ne	Effective population size.
Non-target taxa of concern (NTTOC)	Species, stocks, or components of a stock with high value (e.g., stewardship or utilization) that may suffer negative effects because of a hatchery program.
NRR	Natural replacement rate is the ratio of the number of returning naturally produced adults relative to the number of adults that naturally spawned, both hatchery and naturally produced.
NTTOC	Non-target taxa of concern.
pHOS	Proportion of Hatchery Origin Spawners.
PNI	Proportionate Natural Influence.
pNOB	Proportion of Natural Origin Broodstock.
PRCC HSC	Priest Rapids Coordinating Committee Hatchery Subcommittee.
Productivity	The capacity in which juvenile fish or adults can be produced.
Reference population	A population in which no directed artificial propagation is currently directed, although may have occurred in the past. Reference populations are used to monitor the natural variability in survival rates and out of basin impacts on survival.
Smolt-to-adult survival rate (SAR)	Smolt-to-adult survival rate is a measure of the number of adults that return from a given smolt population.
Segregated	A type of hatchery program in which returning adults are spatially or temporally isolated from other populations.

Size at maturity	The length or weight of a fish at a point in time during the year in which spawning will occur.
Smolts per redd	The total number of smolts produced from a stream divided by the total number of redds from which they were produced.
SNP or single-nucleotide polymorphism	A single-nucleotide polymorphism is a variation in a single nucleotide that occurs at a specific position in the genome, where each variation is present to some appreciable degree within a population.
Spawning Escapement	The number of adult fish that survive to spawn.
Stray rate	The rate at which fish spawn outside of natal rivers or the stream in which they were released.
Supplementation	A hatchery strategy where the main purpose is to increase the relative abundance of natural spawning fish without reducing the long-term fitness of the population.
Target population	A specific population in which management actions are directed (e.g., artificial propagation, harvest, or conservation).

APPENDIX 1: ESTIMATION OF CARRYING CAPACITY

In the ecological literature, carrying capacity is often defined as the maximum population size that can be supported indefinitely by the environment (Cain et al. 2014). Said another way, carrying capacity is the maximum number or biomass of a species that a given habitat can support. This maximal environment load is often referred to as "habitat capacity" and is identified with the letter "C." In contrast, the carrying capacity parameter "K" in population models (e.g., logistic equation, Beverton-Holt model, hockey stick model, and the Ricker model) defines a maximum equilibrium population size. Maximum equilibrium population size is generated from density dependent mechanisms that reduce population growth rates as population size increases (negative density dependence). This is referred to as compensation. Population size fluctuates about the maximum equilibrium size because of variability in vital rates that are unrelated to density (density independent factors) and measurement error. Maximum equilibrium population size is often referred to as "population capacity." The two capacities (habitat capacity and population capacity) are related but not identical and therefore should not be confused. Habitat capacity will usually be greater than population capacity.

Estimation of carrying capacity is important because hatchery managers use it to inform supplementation programs, harvest managers use it to set appropriate harvest and escapement levels, modelers use it in life-cycle models to predict the effects of different recovery scenarios, and restoration practitioners use it to guide restoration actions. The purpose of this paper is to describe methods that can be used to estimate carrying capacity for stocks within the Upper Columbia River basin. We apply these methods to Wenatchee and Chiwawa River spring Chinook salmon.[13] Data used in this exercise are shown in Tables 1 and 2 and come from Hillman et al. (2017). We begin by identifying simple methods used to detect density dependence. We then describe the use of population models to estimate population capacity. We also discuss the use of habitat models and quantile regression to estimate habitat capacity. We end by comparing results of different methods and offering recommendations for estimating carrying capacity.

Table 1. Numbers of redds, adult spawners (estimated from redd surveys), eggs (estimated as the number of redds times average brood-year fecundity), summer parr (estimated using snorkel surveys), and yearling smolts (estimates using a rotary screw trap) by brood year for spring Chinook salmon in the Chiwawa River watershed. Smolts represent the number of yearling Chinook produced entirely within the Chiwawa River watershed. Date are from Hillman et al. (2017). NS = not sampled.

Brood year	Numbers of Chiwawa spring Chinook				
	Redds	Spawners	Eggs	Parr	Smolts
1991	104	242	478,400	45,483	42525
1992	302	676	1,570,098	79,113	39723
1993	106	233	556,394	55,056	8662
1994	82	184	485,686	55,241	16472
1995	13	33	66,248	5,815	3830

---

[13] Technically, Wenatchee River spring Chinook are one population. Chiwawa River spring Chinook are a subgroup of the Wenatchee spring Chinook population.

Brood year	Numbers of Chiwawa spring Chinook				
	Redds	Spawners	Eggs	Parr	Smolts
1996	23	58	106,835	16,066	15475
1997	82	182	374,740	68,415	28,334
1998	41	91	218,325	41,629	23,068
1999	34	94	166,090	NS	10,661
2000	128	346	642,944	114,617	40,831
2001	1,078	1,725	4,984,672	134,874	86,482
2002	345	707	1,605,630	91,278	90,948
2003	111	270	648,684	45,177	16,755
2004	241	851	1,156,559	49,631	72,080
2005	332	599	1,436,564	79,902	69,064
2006	297	529	1,284,228	60,752	45,050
2007	283	1,296	1,256,803	82,351	25,809
2008	689	1,158	3,163,888	106,705	35,023
2009	421	1,347	1,925,233	128,220	30,959
2010	502	1,094	2,165,628	141,510	47,511
2011	492	2,032	2,157,420	103,940	37,185
2012	880	1,478	3,716,240	149,563	34,334
2013	714	1,378	3,367,224	121,240	39,396
2014	485	999	1,961,825	111,224	37,170
2015	543	967	2,631,921	140,172	

Table 2. Numbers of redds, adult spawners (estimated from redd surveys), eggs (estimated as the number of redds times average brood-year fecundity), and yearling smolts (estimates using a rotary screw trap) by brood year for spring Chinook salmon in the Wenatchee River basin. Smolts represent the number of yearling Chinook produced entirely within the Wenatchee River basin. Date are from Hillman et al. (2017). NS = not sampled.

Brood year	Numbers of Wenatchee spring Chinook			
	Redds	Spawners	Eggs	Smolts*
2000	350	830	1,758,050	76,643
2001	2,109	3,217	8,674,624	243,516
2002	1,139	1,965	5,300,906	165,116
2003	323	673	1,887,612	70,738
2004	574	1,686	2,663,445	55,619
2005	830	1,484	3,587,083	302,116
2006	588	1,000	2,542,512	85,558
2007	466	2,035	2,069,506	60,219
2008	1,411	2,278	6,479,312	82,137
2009	733	2,299	NS	NS

Brood year	Numbers of Wenatchee spring Chinook			
	Redds	**Spawners**	**Eggs**	**Smolts***
2010	968	1,921	NS	NS
2011	872	3,139	3,823,720	89,917
2012	1,704	2,720	7,195,992	67,973
2013	1,159	2,133	5,512,204	58,595
2014	885	1,600	3,894,000	36,752

* From 2000-2010 the smolt trap operated near the Town of Monitor; from 2013 to present the trap operated near the Town of Cashmere.

Evidence of Density Dependence

To calculate population capacity, the size of the population or stock must be influenced to a large degree by density-dependent factors. That is, population growth is affected by mechanisms whose effectiveness increases as population size increases. As population density increases, factors such as competition, predation, and disease (and parasites) cause birth rates to decrease, death rates to increase, and dispersal to increase. When densities decrease, the opposite occurs; birth rates increase and death and emigration rates decrease. In general, when the density of the population becomes high enough, density-dependent factors decrease population size because food or space are in short supply (Chapman 1966). In the ecological literature, this is referred to as "population regulation."

A simple way to determine if density-dependent factors regulate population size is to plot population growth rate (or appropriate surrogate) against population size. If population regulation is occurring, the relationship between population size and population growth rate decreases exponentially (decreases linearly if data are log-transformed). Surrogates for population growth rate include survival rates, natality (birth rates), productivity, recruits, individual growth rates, and movement. Figure 1 shows the relationship between productivity (parr/spawner and smolts/spawner) and spawning escapement for Wenatchee River and Chiwawa River spring Chinook. One could use redd counts as a surrogate for spawning abundance. Because most female spring Chinook construct only one redd (Murdoch et al. 2009), redd counts reflect the number of female spawners in the population. In this report, we use number of spawners (spawning escapement) because most management decisions are based on spawning escapement.

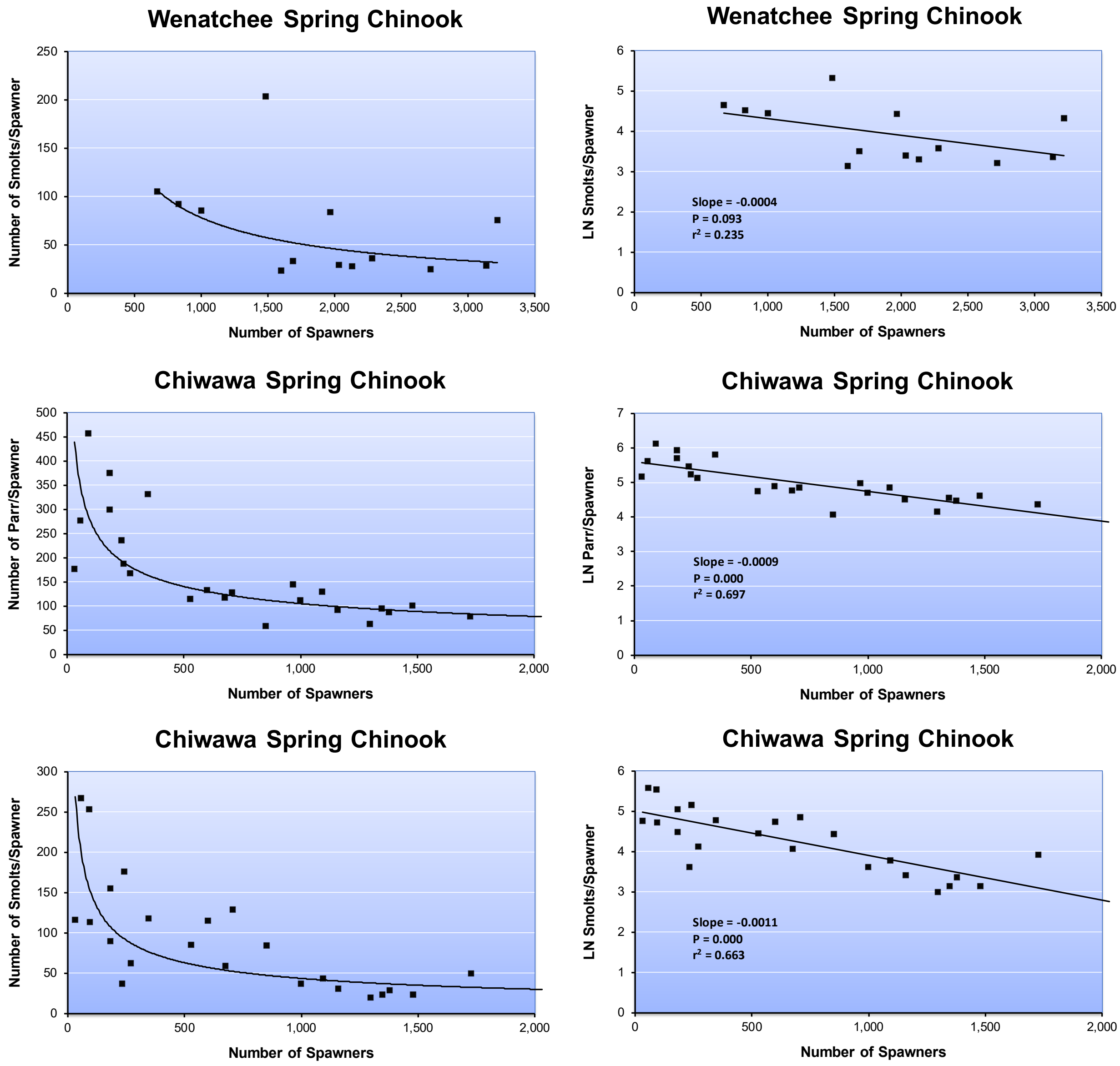


Figure 1. Relationship between spawner abundance and smolts/spawner for Wenatchee spring Chinook (top figures), spawner abundance and parr/spawner for Chiwawa spring Chinook (middle figures), and spawner abundance and smolts/spawner for Chiwawa spring Chinook (bottom figures). Figures on the right show natural log transformed productivity data.

The negative relationship between spawner abundance and juvenile productivity indicates the presence of density dependence in Chiwawa spring Chinook. Although there is a hint of density dependence in the Wenatchee River productivity data, the relationship was not significant statistically. This in part may be related to changes in sampling over the 13-year period. The negative relationship was significant for both summer parr and yearling smolts in the Chiwawa River watershed. We caution, however, that there may be a bias in the simple regression analysis presented in the figures. That is, the dependent (productivity) and independent (abundance) variables are not independent and this can produce a negative bias in regression estimates of slope. Nevertheless, the decline in juvenile productivity with increasing spawner abundance indicates the presence of density dependence. Given the presence of density dependence, we should be able to estimate population capacity.

Estimating Carrying Capacity

Several different methods can be used to estimate population capacity. For example, time series analyses, including the logistic or Gompertz functions, or stock-recruitment models can be used to estimate population capacity. Common stock-recruitment models include Ricker, Beverton-Holt, and smooth hockey stick models. These models incorporate environmental variability and can be used to estimate the size of the spawning population needed to produce the maximum number of recruits. Habitat capacity, on the other hand, can be estimated using fish-habitat models. In general, these models estimate habitat capacity as the product of habitat area and fish/habitat relationships. These range from simple models such as percent habitat saturation models to more complex models including habitat suitability, quantile regression forest models, dynamic food-web models, and bioenergetic or net rate of energy intake models. In this report, we explore the use of stock-recruitment models to estimate population capacity. We apply quantile regression to stock-recruitment models to estimate habitat capacity and compare those results to a habitat model, the quantile regression forest model.

Population Capacity

To estimate population capacity, we evaluated the fit of three different stock-recruitment models to Chiwawa and Wenatchee River spring Chinook data: Ricker, Beverton-Holt, and smooth hockey stick models. In using these models, we assume:

- *Density-dependent mortality*—For some time period before recruitment, the brood instantaneous mortality rate is proportional to the number of parent spawners (Ricker 1954).
- *Lognormal variation*—At any particular spawning stock size, the variation in recruitment is log-normally distributed about its average, and acts multiplicatively (Quinn and Deriso 1999).
- *Measurement error*—Error in spawning stock size estimates (measurement error) is small relative to the range of spawning stock sizes observed (Hilborn and Walters 1992). Variation in realized recruitment at any particular spawning stock size (process error) dominates recruitment measurement error.
- *Stationarity*—The average stock-recruitment relationship is constant over time (Hilborn and Walters 1992). That is, environmental conditions randomly affect survival independent of stock size or time.

In general, the methods we used to fit the models to the data followed those outlined in Hilborn and Walters (1992) and Froese (2008). The Ricker model, which assumes that the number of recruits increases to a maximum and then declines as the number of spawners increases, takes the form:

$$\boldsymbol{E(R) = \alpha S e^{-\beta S}}$$

where E(R) is the expected recruitment, S is spawner abundance, α is the number of recruits per spawner at low spawning levels, and β describes how quickly the recruits per spawner drop as the number of spawners increases. We estimated population capacity (K) as:

$$\boldsymbol{K = \left(\frac{\alpha}{\beta}\right) e^{-1}}$$

and the number of spawners (SP) needed to produce the maximum number of recruits as:

$$\boldsymbol{SP = \frac{1}{\beta}}$$

The Beverton-Holt model assumes that the number of recruits increases constantly toward an asymptote as the number of spawners increases. After the asymptote is reached, the number of recruits neither increases nor decreases. The asymptote represents the maximum number of recruits the system can support (i.e., population capacity for the system; K). The Beverton-Holt curve takes the form:

$$\boldsymbol{E(R) = \frac{(\alpha S)}{(\beta + S)}}$$

where E(R) and S are as above, α is the maximum number of recruits produced (i.e., α = K), and β is the number of spawners needed to produce (on average) recruits equal to one-half the maximum number of recruits. The number of spawners needed to produce the maximum number of recruits is ∞ in the Beverton-Holt model.

Like the Beverton-Holt model, the smooth hockey stick model assumes that the number of recruits increases toward an asymptote (population capacity; K) as the number of spawners increases. After the carrying capacity is reached, the number of recruits neither increases nor decreases. The carrying capacity represents the maximum equilibrium number of recruits the system can support. This curve takes the form (Froese 2008):

$$\boldsymbol{E(R) = R_{\infty}\left(1 - e^{-\left(\frac{\alpha}{R\infty}\right)S}\right)}$$

where E(R) and S are as above, α is the slope at the origin of the spawner-recruitment curve, and R_{∞} is the carrying capacity of recruits (i.e., $R_{\infty} = K$). There is no direct estimate of SP in the smooth hockey stick model. Therefore, we estimated SP as the number of spawners needed to produce 0.95(K).

We used non-linear regression to fit the three models to spawner-recruitment data. Before fitting the models, we transformed recruitment data using natural logs. We estimated bias and uncertainty measures (95% CI) for the model parameters using bootstrap procedures, which assumed that the {R, S} sample represented or approximated the population. The number of bootstrap samples was 3,000. We computed and stored the non-linear regression results for each bootstrap sample. We then calculated the bootstrap 95% CI by arranging the 3,000 bootstrap parameter values in sorted order and selected the 2.5 and 97.5 percentiles from the list.

We used Akaike's Information Criterion for small sample size (AIC_c) to determine which model(s) best explained the relationship between spawners and recruitment in the supplemented and reference populations. AIC_c was estimated as:

$$\boldsymbol{AIC_c = -2log\big(£(\theta|data)\big) + 2K + \left(\frac{2K(K+1)}{n-K-1}\right)}$$

where *log(£(θ|data))* is the maximum likelihood estimate, *K* is the number of estimable parameters (structural parameters plus the residual variance parameter), and *n* is the sample size (Burnham and Anderson 2002). We used least-squares methods to estimate *log(£(θ|data))*, which was calculated as $log(\sigma^2)$, where σ^2 = residual sum of squares divided by the sample size ($\sigma^2 = RSS/n$). AIC_c assessed model fit in relation to model complexity (number of parameters). The model with the smallest AIC_c value represented the "best approximating" model within the model set. Remaining models were ranked relative to the best model using AIC_c difference scores (ΔAIC_c), Akaike weights (w_i), and evidence ratios. Models with ΔAIC_c values less than 2 indicated that there is substantial support for these models as being the best-fitting models within the set (Burnham and Anderson 2002). Models with values greater than 2 had less support. Akaike weights are probabilities estimating the strength of the evidence supporting a particular model as being the best model within the model set. Models with small w_i values are less plausible as competing models (Burnham and Anderson 2002). If no single model could be specified as the best model, a "best subset" of competing models was identified using (1) AIC_c differences to indicate the level of empirical support each model had as being the best model, (2) evidence ratios based on Akaike weights to indicate the relative probability that any model is the best model, and (3) coefficients of determination (R^2) assessing the explanatory power of each model.

Chiwawa River Spring Chinook Parr

We successfully fit the three stock-recruitment curves to the Chiwawa spring Chinook parr data (Figure 2).

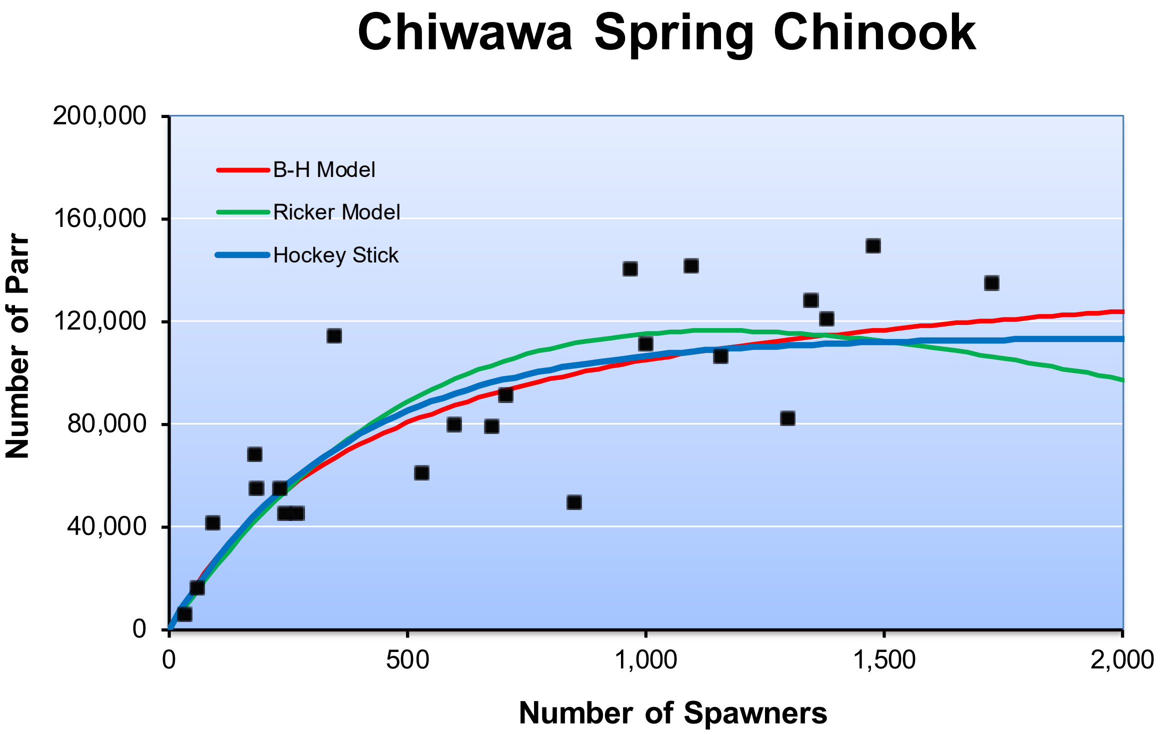


Figure 2. Relationship between numbers of spring Chinook parr and numbers of spawners in the Chiwawa River watershed, 1992-2016 (no sampling occurred in 2000). Figure shows the fit of the Beverton-Holt, Ricker, and smooth hockey stick models to the data.

For summer parr, the use of AIC_c indicated that the Beverton-Holt model best approximated the information in the productivity data. The estimated structural parameters for this model were:

$$Parr = \frac{(150{,}902 \times Spawners)}{(438 + Spawners)}$$

where the bootstrap estimated standard errors for the two parameters were 21,142 and 145, respectively. The adjusted $R^2 = 0.812$.

The second-best model was the smooth hockey stick model, which was 0.245 AIC_c units from the best model. The estimated parameters for this model were:

$$LN(Parr) = 11.6 + LN\left(1 - e^{-\left(\frac{312.9}{113{,}801}\right)Spawners}\right)$$

where the bootstrap estimated standard errors of the two parameters were 0.097 and 57.578, respectively, and the $R^2 = 0.810$.

The AIC_c difference scores, Akaike weights, and evidence ratios indicated that there was substantial support for both the Beverton-Holt and smooth hockey stick models. There was less support for the Ricker model, which was > 2 AIC_c units from the best models. This was further supported by the fact that, relative to the best models, the Ricker model had an evidence ratio greater than 3.

Depending on the stock-recruitment model used, population capacity ranged from 113,801 to 150,902 parr (Table 3). The Beverton-Holt model estimated the highest capacity, while the smooth hockey stick model estimated the lowest. The number of spawners needed to produce the population capacity of parr ranged from 1,089 to 1,163 (Table 3).

Table 3. Estimates of Beverton-Holt, smooth hockey stick, and Ricker model parameters, parr capacity (K), parr productivity (parr per spawner), and the number of spawners needed to produce the maximum number of parr for Chiwawa River spring Chinook.

Model	Parameter		Population capacity (K)	Intrinsic productivity	Spawners
	A	B			
Beverton-Holt	150,902.145	437.655	150,902	345	∞
Smooth Hockey Stick	11.642	312.913	113,801	313	1,089
Ricker	272.696	0.0009	116,650	273	1,163

It is important to note that the population capacity estimates are based on the number of parr counted in the Chiwawa River watershed during August. There are spring Chinook fry and parr that move out of the Chiwawa River watershed during spring and early summer (Hillman et al. 2017). It is unknown if these fish leave because of density-dependent pressures, they are flushed out during high flows, it is a life-history characteristic, or a combination of these. Regardless of the mechanism or reason, some of these fish may survive and rear in the Wenatchee or Columbia rivers. These emigrants are not included in the capacity estimates shown in Table 3.

The capacity estimates for spring Chinook parr apply only to the Chiwawa River watershed, a watershed within the Wenatchee River basin. Estimating parr capacity for the entire Wenatchee River basin using stock-recruitment models is difficult because there is no long-term time series of parr data for the entire basin. However, we can extrapolate parr capacity estimates from the

Chiwawa River watershed to the entire Wenatchee River basin using intrinsic potential (IP). Multiplying the parr capacity per intrinsic potential within the Chiwawa River watershed by the total intrinsic potential within the Wenatchee River basin yields an estimate of parr capacity for the Wenatchee River basin (Table 4). The Interior Columbia Basin Technical Recovery Team estimated IP based on wetted width, valley width (confinement), and gradient (see Cooney and Holzer 2006). They used sedimentation and temperature to refine IP for each 200-m long reach. We used the total stream area (km^2) weighted by intrinsic potential and temperature limited to extrapolate parr capacity to the entire Wenatchee River basin.

Table 4. Estimates of Wenatchee River basin parr capacity based on intrinsic potential (IP). The amount of IP within the Chiwawa River watershed is 0.481 km^2; the total amount of IP within the Wenatchee River basin is 1.798 km^2.

Model	Chiwawa parr capacity	Chiwawa parr/IP	Wenatchee parr capacity
Beverton-Holt	150,902	313,726	564,079
Smooth Hockey Stick	113,801	236,593	425,395
Ricker	116,650	242,516	436,043

Using this simple method, we estimate the Wenatchee River basin supports about 425,395-564,079 parr depending on which model is used. An important assumption of this simple method is that each unit of IP supports the same number of parr. This is clearly not true given that the quality of habitat within each unit of IP can vary greatly. That is, one unit of IP may contain more habitat structure (e.g., wood and cover) than another unit of IP. Importantly, the ratio of parr to IP comes from the Chiwawa River watershed, which contains some of the highest quality habitat within the Wenatchee River basin. Therefore, the estimated total parr capacity for the entire Wenatchee River basin is likely biased high. If habitat conditions throughout the Wenatchee River basin are enhanced to conditions similar to those in the Chiwawa River watershed, we may expect parr abundance to approach those estimated with this simple method.

Chiwawa River Spring Chinook Smolts

We successfully fit the three stock-recruitment curves to the Chiwawa spring Chinook smolt data (Figure 3). This information allows us to better understand the quality and quantity of overwintering habitat in the Chiwawa River basin.

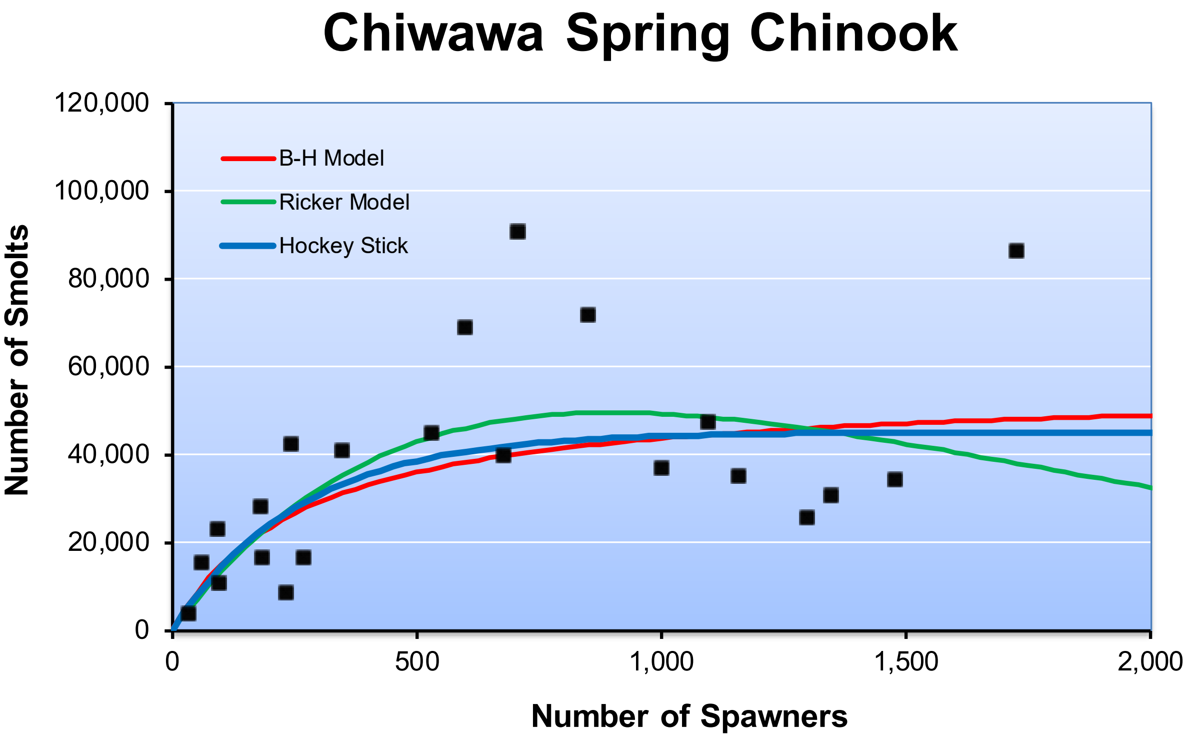


Figure 3. Relationship between numbers of spring Chinook smolts and numbers of spawners in the Chiwawa River watershed, 1992-2015. Figure shows the fit of the Beverton-Holt, Ricker, and smooth hockey stick models to the data.

For yearling smolts produced entirely within the Chiwawa River watershed, the use of AIC$_c$ indicated that the smooth hockey stick model best approximated the information in the productivity data. The estimated structural parameters for this model were:

$$LN(Smolts) = 10.7 + LN\left(1 - e^{-\left(\frac{174.1}{45,161}\right)Spawners}\right)$$

where the bootstrap estimated standard errors for the two parameters were 0.13 and 41.29, respectively. The adjusted $R^2 = 0.569$.

The second-best model was the Ricker model, which was 0.234 AIC$_c$ units from the best model. The estimated parameters for this model were:

$$Smolts = 149.45 \times Spawners(\mathrm{e}^{-0.00111 \times Spawners})$$

where the bootstrap estimated standard errors of the two parameters were 26.23 and 0.00018, respectively, and the $R^2 = 0.573$.

The third-best model was the Beverton-Holt model, which was 0.725 AIC$_c$ units from the best model. The estimated parameters for this model were:

$$Smolts = \frac{(55,702 \times Spawners)}{(273 + Spawners)}$$

where the bootstrap estimated standard errors of the two parameters were 10,421.9 and 123.0, respectively, and the $R^2 = 0.560$.

The AIC_c difference scores, Akaike weights, and evidence ratios indicated that there was substantial support for all three models. Relative to the best model, the other two models had evidence ratios less than 1.5.

Depending on the stock-recruitment model used, population capacity ranged from 45,161 to 55,702 smolts (Table 5). The Beverton-Holt model estimated the highest capacity, while the smooth hockey stick model estimated the lowest. The number of spawners needed to produce the population capacity of smolts ranged from 777 to 901 (Table 5).

Table 5. Estimates of Beverton-Holt, smooth hockey stick, and Ricker model parameters, smolt capacity (K), smolt productivity (smolts per spawner), and the number of spawners needed to produce the maximum number of smolts for Chiwawa River spring Chinook.

Model	Parameter		Population capacity (K)	Intrinsic productivity	Spawners
	A	B			
Smooth hockey stick	10.718	174.077	45,161	174	777
Ricker	149.452	0.00111	49,532	149	901
Beverton-Holt	55,702.281	273.910	55,702	203	∞

It is important to note that the population capacity estimates are based on the number of smolts produced entirely within the Chiwawa River watershed. As noted earlier, there are spring Chinook fry and parr that move out of the Chiwawa River watershed during spring, early summer, and fall (Hillman et al. 2017). Fall emigration is common and occurs even when densities of juveniles are very low, indicating that fall emigration is a life-history characteristic. Regardless of why the fish emigrate as fry and parr, some of these fish survive and rear in the Wenatchee or Columbia rivers. Some survive to smolt (unpublished WDFW data), but are not included in the smolt capacity estimates shown in Table 5.

As with parr, the capacity estimates for spring Chinook smolts apply only to the Chiwawa River watershed. As before, we can extrapolate smolt capacity estimates from the Chiwawa River watershed to the entire Wenatchee River basin using intrinsic potential (IP). In this case, we multiply the smolt capacity per intrinsic potential within the Chiwawa River watershed by the total intrinsic potential within the Wenatchee River basin. This yields an estimate of smolt capacity for the Wenatchee River basin (Table 6).

Table 6. Estimates of Wenatchee River basin smolt capacity based on intrinsic potential (IP). The amount of IP within the Chiwawa River watershed is 0.481 km^2; the total amount of IP within the Wenatchee River basin is 1.798 km^2.

Model	Chiwawa smolt capacity	Chiwawa smolts/IP	Wenatchee smolt capacity
Beverton-Holt	55,702	115,805	208,218
Smooth Hockey Stick	45,161	93,891	168,816
Ricker	49,532	102,976	185,152

Using this simple method, we estimate the population capacity for the Wenatchee River basin at 168,816-208,218 smolts depending on which model is used. Based on smolt trapping in the lower Wenatchee River over a 13-year period, total smolt abundance has ranged from 36,752 to 302,116 smolts (average = 107,300 smolts) (Table 2).[14] Thus, recent (2000-2014) smolt production appears to be below capacity estimates for most years but higher in some years.

An important assumption of this simple method is that each unit of IP supports the same number of smolts. As we noted earlier, this is not the case given that the quality of habitat within each unit of IP can vary greatly. Nevertheless, the ratio of smolts to IP comes from the Chiwawa River watershed, which contains some of the highest quality habitat within the Wenatchee River basin. Therefore, the estimated total smolt capacity for the entire Wenatchee River basin is likely biased high. If habitat conditions throughout the Wenatchee River basin are enhanced to conditions similar to those in the Chiwawa River watershed, we may expect smolt abundance to approach those estimated with this simple method.

Wenatchee River Spring Chinook Smolts

Rather than extrapolate results from the Chiwawa River watershed to the entire Wenatchee River basin, we can fit stock-recruitment models to the smolt data collected in the lower Wenatchee River and estimate population capacity directly from the population models. We successfully fit the three stock-recruitment curves to the Chiwawa spring Chinook smolt data; although, the models explained little of the variation in the stock-recruitment data ($R^2 < 0.05$) (Figure 3).

---

[14] It is important to point out that the trapping location has changed over time. During the period 2000-2008 and 2011-2012, the trap was located near the Town of Monitor. During the period 2013-present, the trap was located near the Town of Cashmere.

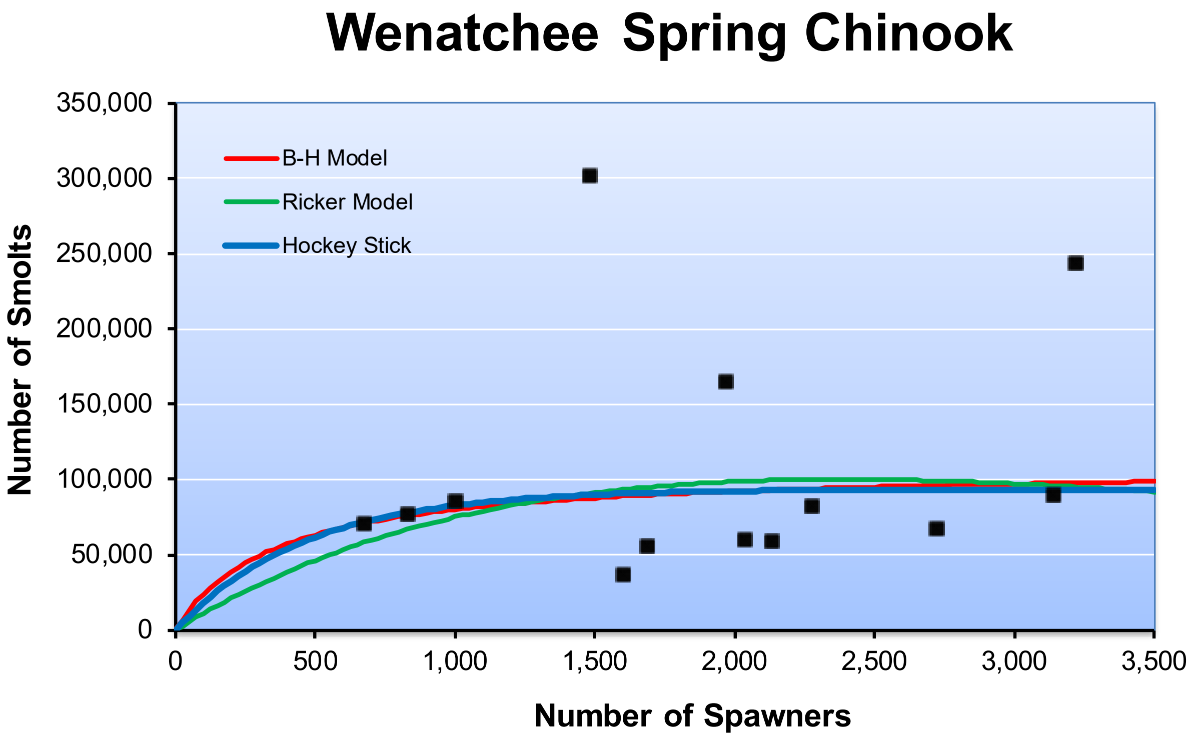


Figure 4. Relationship between numbers of spring Chinook smolts and numbers of spawners in the Wenatchee River basin, 2000-2014 (no data were collected in 2009 or 2010). Figure shows the fit of the Beverton-Holt, Ricker, and smooth hockey stick models to the data.

For yearling smolts produced within the Wenatchee River basin, the use of AIC$_c$ indicated that the Beverton-Holt model best approximated the information in the productivity data. The estimated structural parameters for this model were:

$$Smolts = \frac{(108{,}696 \times Spawners)}{(359 + Spawners)}$$

where the bootstrap estimated standard errors for the two parameters were 49,948 and 836, respectively. The adjusted $R^2 = 0.026$.

The second-best model was the smooth hockey stick model, which was 0.112 AIC$_c$ units from the best model. The estimated parameters for this model were:

$$LN(Smolts) = 11.4 + LN\left(1 - e^{-\left(\frac{20.72}{93{,}560}\right)Spawners}\right)$$

where the bootstrap estimated standard errors of the two parameters were 30.74 and 225.43, respectively, and the $R^2 = 0.017$.

The third-best model was the Ricker model, which was 0.0.808 AIC$_c$ units from the best model. The estimated parameters for this model were:

$$Smolts = 114.10 \times Spawners(\mathrm{e}^{-0.00042 \times Spawners})$$

where the bootstrap estimated standard errors of the two parameters were 56.16 and 0.00021, respectively, and the $R^2 = 0.001$.

The AIC$_c$ difference scores, Akaike weights, and evidence ratios indicated that there was substantial support for all three models. Relative to the best model, the other two models had evidence ratios less than 2.0.

Depending on the stock-recruitment model used, population capacity for the Wenatchee River basin ranged from 93,560 to 108,696 smolts (Table 7). The Beverton-Holt model estimated the highest capacity, while the smooth hockey stick model estimated the lowest. The number of spawners needed to produce the population capacity of smolts ranged from 1,389-2,381 (Table 7).

Table 7. Estimates of Beverton-Holt, smooth hockey stick, and Ricker model parameters, smolt capacity (K), smolt productivity (smolts per spawner), and the number of spawners needed to produce the maximum number of smolts for Wenatchee River spring Chinook.

Model	Parameter		Population capacity (K)	Intrinsic productivity	Spawners
	A	B			
Smooth hockey stick	11.446	201.724	93,560	202	1,389
Ricker	114.104	0.00042	99,944	114	2,381
Beverton-Holt	108,696.009	358.616	108,696	303	∞

The population capacity estimates reported here are based on the number of smolts produced within the Wenatchee River basin. It is likely that some juvenile spring Chinook rear in the Columbia River and survive to smolt. Those fish are not included in these estimates of capacity.

Habitat Capacity

Habitat capacity can be estimated using fish-habitat models and creative modeling of stock-recruitment data. As we noted earlier, there are several different fish-habitat models that can be used to estimate habitat capacity. In this paper, we explore the use of two different methods, quantile regression applied to stock-recruitment functions and the Quantile Regression Random Forest model. The former relies on simple stock and recruitment data, while the latter requires estimates of habitat quality and quantity, and functional relationships between maximum fish density and habitat conditions.

Quantile Regression Analysis of Stock-Recruitment Data

To estimate population capacity, we used non-linear regression techniques to fit stock-recruitment functions to the data. These techniques approximate the conditional mean of the recruitment data given the range of stock sizes. As such, the functions (curves) estimated from the analyses lie near the center of the distribution of data resulting in data points above and below the curve. Although this technique is useful for estimating population capacity, it is not appropriate for estimating habitat capacity. The fact that there are actual recruitment data above the estimated population capacity indicates that habitat capacity must be greater than the population capacity, or that measurement error is high. The former explanation is more likely than the latter.

One way to possibly estimate habitat capacity with stock-recruitment data is to fit stock-recruitment functions to the juvenile spring Chinook data using quantile regression techniques. Quantile regression estimates quantiles of the recruitment data given the range of stock sizes. Thus, we can use quantile regression to fit a stock-recruitment function to, say, the upper 90% or 95% of the recruitment distribution. In other words, we fit a stock-recruitment function to the upper limits of the recruitment data given the range of stock sizes. In this case, the resulting stock-recruitment curve is above most of the recruitment data and therefore few data points lie above the curve. Calculation of capacity from these functions should more closely represent habitat capacity, provided there is an adequate range of stock sizes. Quantile regression gives results similar to those obtained from calculating reference intervals (RI).

In this exercise, we calculated the upper 90% RI for the Beverton-Holt and Ricker functions. We assume the 90% RI will closely represent the habitat capacity for juvenile spring Chinook. We calculated the 90% RI only for the Beverton-Holt and Ricker models, because these functions can be transformed into linear function (see Hilborn and Walters 1992). RIs are easier to calculate on linear functions than non-linear functions. We were unable to transform the smooth hockey stick model into a linear function and therefore we did not calculate RIs for this function.

Chiwawa River Spring Chinook Parr—We calculated 90% RIs for Chiwawa Chinook parr data for both the Ricker and Beverton-Holt models (Figure 5). The estimated parameters for the 90% RI for the Ricker model were:

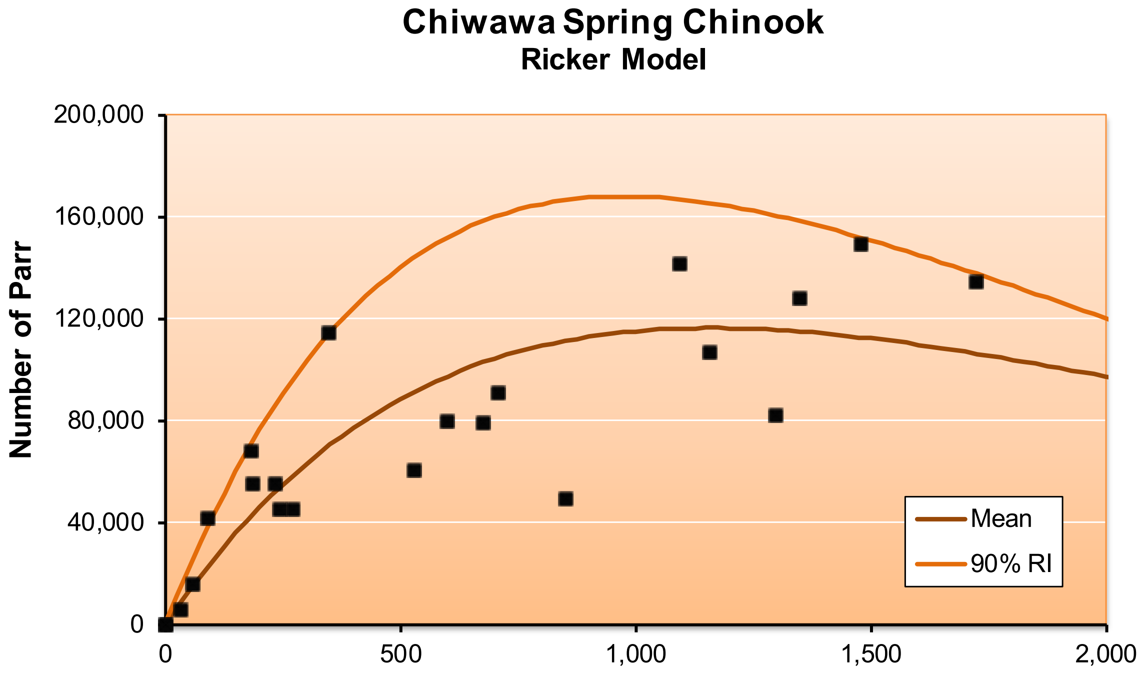
$$\log\left(\frac{Parr}{Spawners}\right) = 6.152 - \frac{6.152}{5{,}984.436}(Spawners)$$

This resulted in an estimated habitat capacity of 168,071 parr, which is about 1.4 times greater than the population capacity estimated with the Ricker model.

The estimated parameters for the 90% RI for the Beverton-Holt model were:

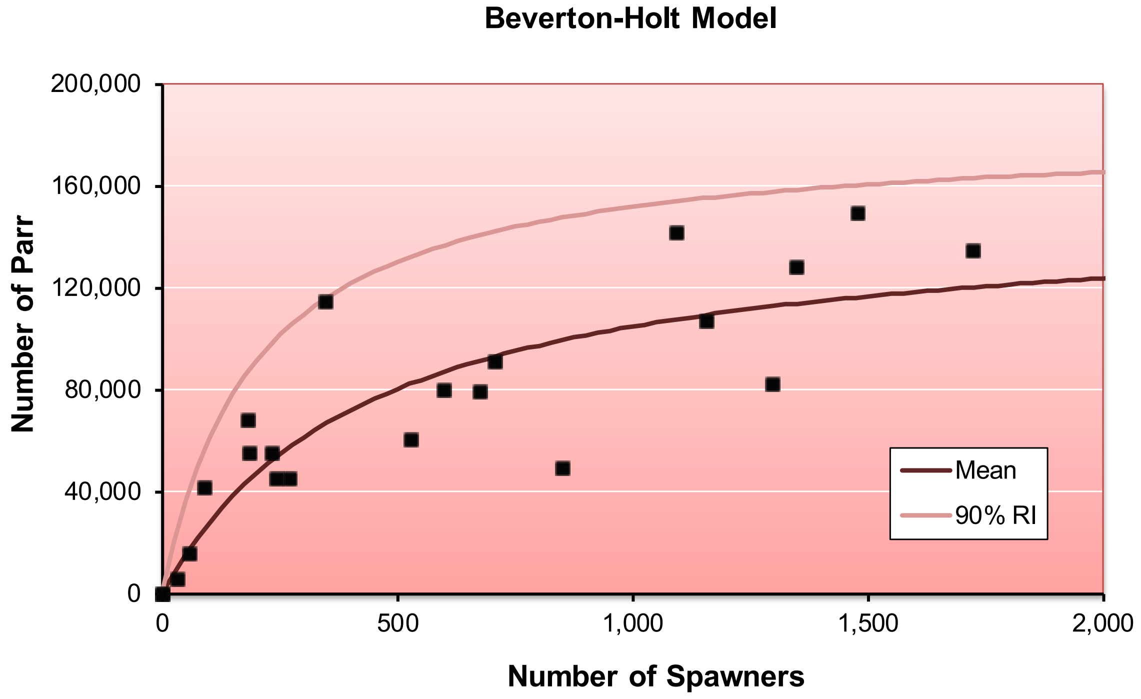
$$\frac{Spawners}{Parr} = \frac{196.91}{181{,}818} + \frac{1}{181{,}818}(Spawners)$$

This function resulted in an estimated habitat capacity of 181,818 parr, which was about 1.2 times greater than the population capacity estimated with the Beverton-Holt model.



Chiwawa Spring Chinook

Ricker Model



Beverton-Holt Model

Figure 5. Relationship between numbers of spring Chinook parr and numbers of spawners in the Chiwawa River watershed, 1992-2016. Upper figure shows the fit of the Ricker model and its 90% reference interval to the data; lower figure shows the fit of the Beverton-Holt model and its 90% reference interval.

If we extrapolate the habitat capacity estimates for Chiwawa spring Chinook parr to the entire Wenatchee River basin (using the IP method described earlier), we estimate the habitat capacity for the Wenatchee River basin to be 628,256 parr from the Ricker model and 679,645 parr from the Beverton-Holt model.

<u>*Chiwawa River Spring Chinook Smolts*</u>—As with parr, we calculated 90% RIs for Chiwawa Chinook smolt data for both the Ricker and Beverton-Holt models (Figure 6). The estimated parameters for the 90% RI for the Ricker model were:

$$\log\left(\frac{Smolts}{Spawners}\right) = 5.687 - \frac{5.687}{4{,}687.964}(Spawners)$$

This resulted in an estimated habitat capacity of 89,425 smolts, which is about 1.8 times greater than the population capacity estimated with the Ricker model.

The estimated parameters for the 90% RI for the Beverton-Holt model were:

$$\frac{Spawners}{Smolts} = \frac{102.129}{64{,}516} + \frac{1}{64{,}516}(Spawners)$$

This function resulted in an estimated habitat capacity of 64,516 smolts, which was about 1.2 times greater than the population capacity estimated with the Beverton-Holt model.

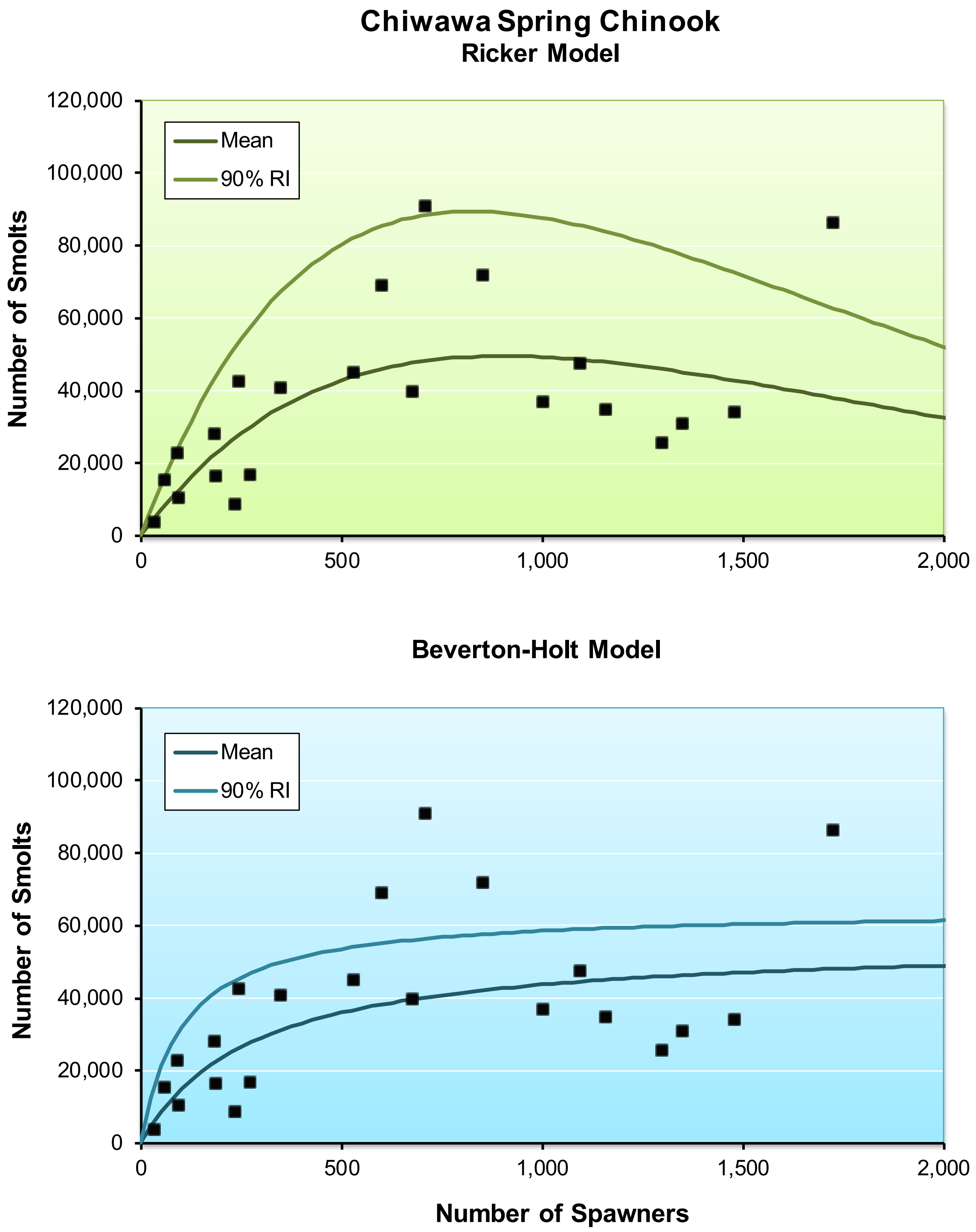


Figure 6. Relationship between numbers of spring Chinook smolts and numbers of spawners in the Chiwawa River watershed, 1992-2015. Upper figure shows the fit of the Ricker model and its 90% reference interval to the data; lower figure shows the fit of the Beverton-Holt model and its 90% reference interval.

If we extrapolate the habitat capacity estimates for Chiwawa spring Chinook smolts to the entire Wenatchee River basin (using the IP method described earlier), we estimate the habitat capacity for the Wenatchee River basin to be 334,276 smolts based on the Ricker model and 241,164 smolts from the Beverton-Holt model.

<u>*Wenatchee River Spring Chinook Smolts*</u>—We calculated 90% RIs for Wenatchee River Chinook smolt data for both the Ricker and Beverton-Holt models (Figure 7). The estimated parameters for the 90% RI for the Ricker model were:

$$\log\left(\frac{Smolts}{Spawners}\right) = 5.320 - \frac{5.320}{16{,}642.420}(Spawners)$$

This resulted in an estimated habitat capacity of 235,131 smolts, which is about 2.4 times greater than the population capacity estimated with the Ricker model.

The estimated parameters for the 90% RI for the Beverton-Holt model were:

$$\frac{Spawners}{Smolts} = \frac{357.593}{186{,}567} + \frac{1}{186{,}567}(Spawners)$$

This function resulted in an estimated habitat capacity of 186,567 smolts, which was about 1.7 times greater than the population capacity estimated with the Beverton-Holt model.

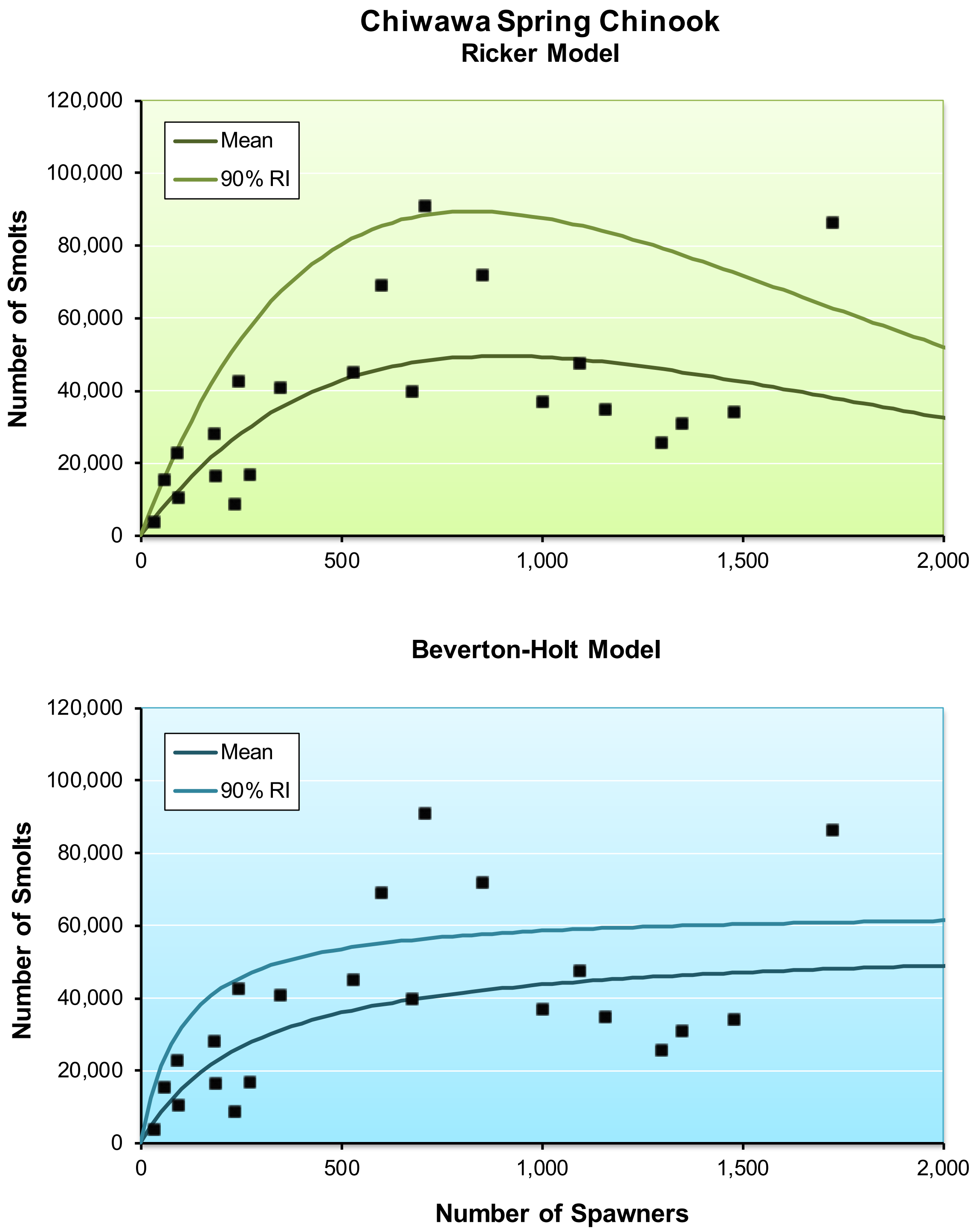


Figure 7. Relationship between numbers of spring Chinook smolts and numbers of spawners in the Wenatchee River basin, 2000-2015 (no data were collected in 2009 or 2010). Upper figure shows the fit of the Ricker model and its 90% reference interval to the data; lower figure shows the fit of the Beverton-Holt model and its 90% reference interval.

Quantile Regression Random Forest Model

Researchers with the Integrated Status and Effectiveness Monitoring Program (ISEMP) developed a model that estimates Chinook parr habitat capacity based on fish-habitat relationships (ISEMP/CHaMP 2015). Based on extensive sampling throughout the Columbia River basin, these researchers developed relationships between maximum densities of Chinook parr (summer estimates) and various habitat variables. Quantile regression forest (QRF) models use these relationships to estimate carrying capacities for juvenile Chinook. Very simply, QRF analysis develops non-linear relationships between fish density and different habitat variables. In this case, however, QRF analysis predicts the 90% quantile of fish density rather than the mean or median density. The researchers assume that the 90% quantile represents habitat capacity. This is important because the numbers of fish counted in some field sampling sites may not have been at maximum capacity. That is, it is likely that not all sites sampled were fully “seeded” with Chinook salmon. Thus, using the mean or median (50% quantile) would not represent habitat capacity, but some level below habitat capacity.

Researchers fit the QRF model to parr density data and 12 habitat variables that were collected from 227 sites within the distribution of Chinook throughout the Columbia River basin (within CHaMP/ISEMP watersheds). These variables were selected to represent a variety of types of habitat variables (e.g., substrate, riparian, complexity, temperature, etc.), contain the most "fish information," and be as uncorrelated as possible (ISEMP/CHaMP 2015). The 12 habitat variables and their relative importance are shown in Figure 8.

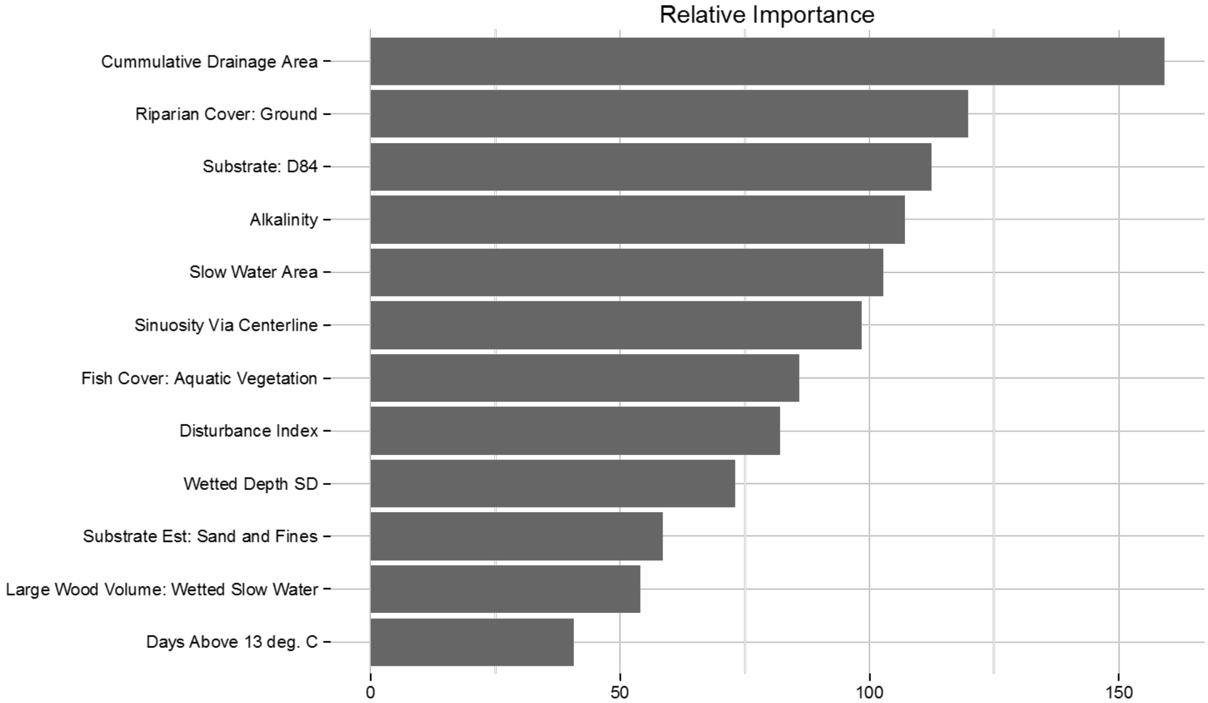


Figure 8. Relative importance of habitat variables included in juvenile Chinook salmon quantile regression forest models (Figure is from ISEMP/CHaMP 2015).

As a way of testing the model, ISEMP researchers used their QRF model to estimate Chinook parr capacities in different watersheds, including the Chiwawa River watershed, and compared their estimates to those generated from fish population data using stock-recruitment modeling. Figure 9 shows the relationship between the QRF model results and population model results for the Chiwawa River watershed. The red curve was generated using the QRF model and the blue curve was generated using the Beverton-Holt model. At the time of this analysis, the Beverton-Holt model was fit to 21 years of parr data, not the 24 years of data used in the analyses above.

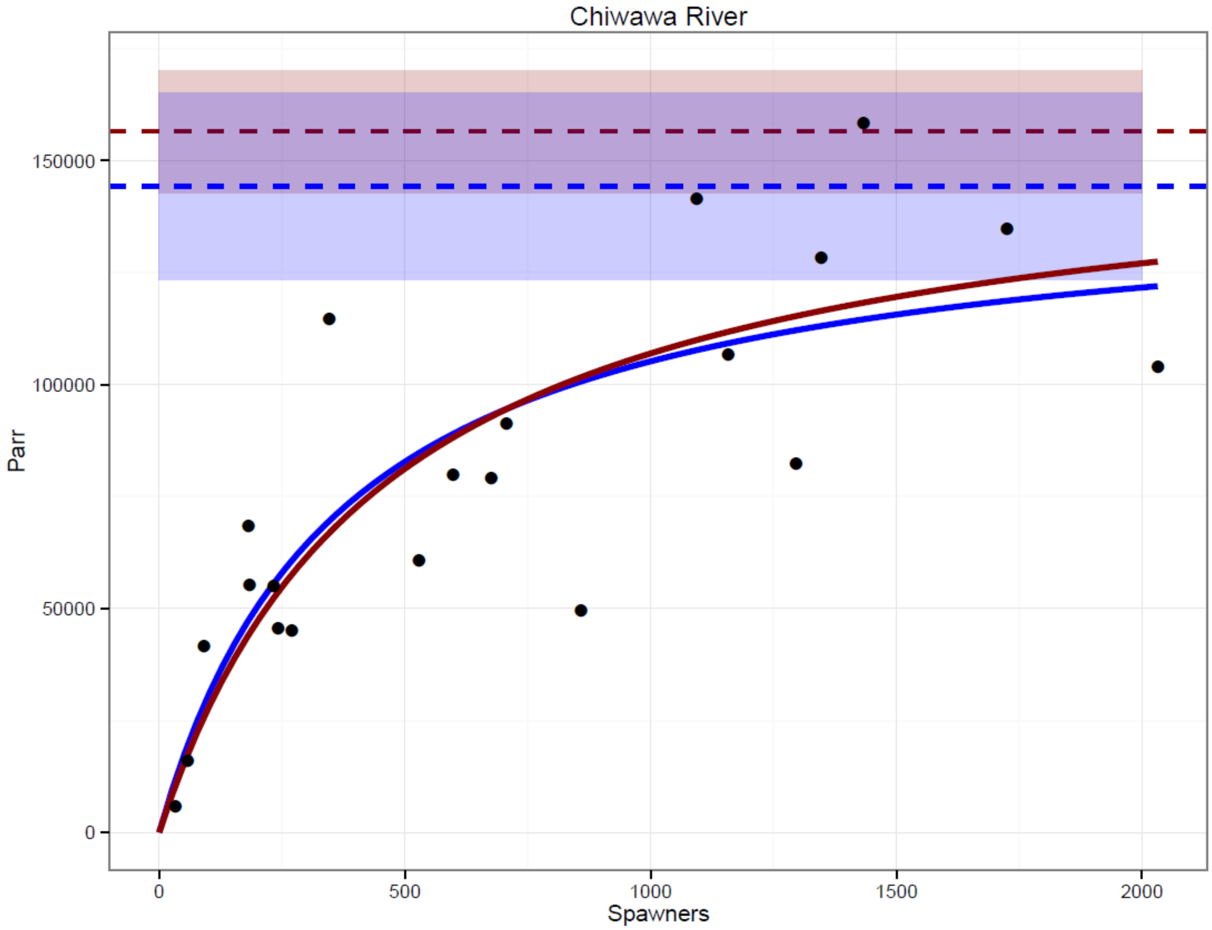


Figure 9. Comparison of productivity curves for Chiwawa spring Chinook parr generated from the QRF model (red line) and Beverton-Holt model (blue line). Dashed horizontal lines represent carrying capacity estimates. Shading about the capacity estimates represent the 95% confidence bounds. Figure is from ISEMP/CHaMP (2015).

The comparison shows that although the curves are very similar, the carrying capacity estimates (dashed horizontal lines) differed, with the habitat capacity generated from the QRF model being larger than the population capacity generated from the population data. That is, the QRF model estimated a habitat capacity of about 164,000 spring Chinook parr, while the population model estimated a population capacity of about 145,000 parr. Including more recent parr data in the Beverton-Holt model indicates that the population capacity estimate is about 151,000 parr for the Chiwawa River watershed. The 90% RI for the Beverton-Holt model estimated a habitat capacity of about 182,000, which is 1.1 times greater than the estimate from the QRF model. Note that the 90% RI for the Ricker model estimated a habitat capacity of about 168,000, which is close to the QRF model estimate.

Comparing Results

We estimated capacities for both spring Chinook parr and smolts for the Chiwawa River watershed and the entire Wenatchee River basin using different analytical tools. In this section, we compare the results from the different approaches.

Parr Capacity

Depending on the population model used, population capacity estimates for the Chiwawa River watershed ranged from 113,801 to 150,902 parr (Table 8). Not surprisingly, the Beverton-Holt model generally predicts the highest capacity estimates, while the smooth hockey stick model predicts the lowest. As expected, the population capacity estimates for Chiwawa parr were less than the habitat capacity estimates for parr. Habitat capacity estimates were about 1.2 to 1.5 times greater than the population capacity estimates (Table 8). Importantly, the fish-habitat model (QRF model) calculated a habitat capacity estimate that was close to that estimated from calculating 90% RI for the population models. Extrapolating Chiwawa capacity estimates to the entire Wenatchee River basin resulted in population capacities of 425,395 to 564,079 parr and habitat capacity estimates of 613,040 to 679,645 parr (Table 8).

Table 8. Comparison of spring Chinook parr capacity estimates for the Chiwawa River watershed and the Wenatchee River basin. Population capacities were estimated directly from the stock-recruitment functions; habitat capacities were estimated by calculating 90% reference intervals (using quantile regression; QR) for the stock-recruitment models and using a fish-habitat model (Quantile Regression Forest Model; QRF). Capacities for the Wenatchee River basin were estimated by extrapolating Chiwawa capacities using intrinsic potential.

Capacity type	Model	Chiwawa parr capacity	Wenatchee parr capacity
Population capacity	Beverton-Holt	150,902	564,079
	Smooth Hockey Stick	113,801	425,395
	Ricker	116,650	436,043
Habitat capacity	QR Beverton-Holt	181,818	679,645
	QR Ricker	168,071	628,256
	QRF Model	164,000	613,040

The number of spawners needed to achieve parr capacity also varied depending on the population model used (Table 9). For the Chiwawa River watershed, maximum spawners needed to achieve population capacity for parr ranged from 1,089 to 1,163 adults. Extrapolating Chiwawa results to the entire Wenatchee River basin resulted in maximum spawner estimates of 4,070 to 4,347 adults. We were able to estimate habitat capacity only with the Ricker model (Table 9). Using quantile regression to calculate the 90% RI for the Ricker model resulted in a maximum spawner abundance of 973 adults, which is less than the number needed to achieve population capacity. This is because the 90% RI for the Ricker function estimates a higher intrinsic productivity, which shifts the "hump" of the curve to the left resulting in a higher capacity estimate but a lower maximum spawner estimate (see Figure 5).

Table 9. Comparison of the number of spawners needed to achieve parr capacities in the Chiwawa River watershed and the Wenatchee River basin. For the Chiwawa River watershed, maximum spawners were estimated directly from the stock-recruitment functions. Maximum spawners for the entire Wenatchee River basin were estimated as the product of the extrapolated parr numbers times the ratio of maximum spawners to parr capacity for Chiwawa spring Chinook. Because of the nature of the Beverton-Holt model, no maximum spawners can be calculated from that model.

Capacity type	Model	Spawners need to achieve parr capacity	
		Chiwawa	Wenatchee
Population capacity	Smooth Hockey Stick	1,089	4,070
	Ricker	1,163	4,347
Habitat capacity	QR Ricker	973	3,636

Smolt Capacity

As with parr estimates, population capacity estimates for smolts varied depending on the population model used. For Chiwawa spring Chinook smolts, population capacities ranged from 45,161 to 55,702 smolts, with the smooth hockey stick providing the lowest estimate and the Beverton-Holt model providing the highest (Table 10). The population capacity estimates were about 55 to 86% of the habitat capacity estimates. Extrapolating Chiwawa capacity estimates to the entire Wenatchee River basin resulted in population capacities of 168,816 to 208,218 smolts and habitat capacity estimates of 241,164 to 334,276 smolts (Table 10). These were greater than those estimated using smolt and spawner data for the entire Wenatchee River basin. Fitting population models to smolt and spawner data for the entire basin resulted in population capacities of 93,560 to 108,696 smolts and habitat capacities of 186,567 to 235,131 smolts (Table 10).

Table 10. Comparison of spring Chinook smolt capacity estimates for the Chiwawa River watershed and the Wenatchee River basin. Population capacities were estimated directly from the stock-recruitment functions; habitat capacities were estimated by calculating 90% reference intervals (using quantile regression; QR) for the stock-recruitment models. Capacities for the Wenatchee River basin were estimated by extrapolating Chiwawa capacities using intrinsic potential and by fitting population models to the smolt and spawner data for the entire basin.

Capacity type	Model	Chiwawa smolt capacity	Wenatchee smolt capacity	
			Chiwawa extrapolation	Wenatchee data
Population capacity	Beverton-Holt	55,702	208,218	108,696
	Smooth Hockey Stick	45,161	168,816	93,560
	Ricker	49,532	185,152	99,944
Habitat capacity	QR Beverton-Holt	64,516	241,164	186,567
	QR Ricker	89,425	334,276	235,131

The number of spawners needed to achieve smolt capacity varied depending on the population model used (Table 11). For the Chiwawa River watershed, maximum spawners needed to achieve population capacity for smolts ranged from 777 to 901 adults. Note that the maximum number of adults needed to achieve population capacity for smolts is less than those needed to achieve population capacity for parr. Extrapolating Chiwawa results to the entire Wenatchee River basin resulted in maximum spawner estimates of 2,904 to 3,368 adults. These estimates are considerably higher than those estimated from fitting population models to Wenatchee River basin data. The latter estimated maximum spawners ranging from 1,389 to 2,381 adults. We were able to estimate habitat capacity only with the Ricker model (Table 11). Using quantile regression to calculate the 90% RI for the Ricker model resulted in a maximum spawner abundance of 824 adults for the Chiwawa River watershed and 3,129 adults for the entire Wenatchee River basin. Extrapolating Chiwawa results to the entire Wenatchee River basin resulted in a maximum spawner estimate of 3,080, which is close to the estimate generated by fitting the model to Wenatchee River basin data.

Table 11. Comparison of the number of spawners needed to achieve smolt capacities in the Chiwawa River watershed and the Wenatchee River basin. Maximum spawners were estimated directly from the stock-recruitment functions. Maximum spawners for the entire Wenatchee River basin were also estimated as the product of the extrapolated smolt numbers times the ratio of maximum spawners to smolt capacity for Chiwawa spring Chinook. Because of the nature of the Beverton-Holt model, no maximum spawners can be calculated from that model.

Capacity type	Model	Spawners need to achieve smolt capacity		
		Chiwawa	Wenatchee	
			Chiwawa extrapolation	Wenatchee data
Population capacity	Smooth Hockey Stick	777	2,904	1,389
	Ricker	901	3,368	2,381
Habitat capacity	QR Ricker	824	3,080	3,129

As an additional exercise, we calculated smolt capacities and maximum spawners generated from fitting population models to smolt and spawner data in the Chiwawa River, Nason Creek, and White River watersheds, and compared the sum of those estimates to the Wenatchee River basin estimates. Only the Ricker model could be fit to the White River and Nason Creek data (see Hillman et al. 2017). Estimated population capacities from the Ricker model were 49,532 smolts in the Chiwawa, 4,412 smolts in Nason Creek, and 4,659 smolts in the White River, resulting in a cumulative population capacity of 58,603 smolts (1,550 spawners are needed to achieve this cumulative smolt capacity). The cumulative population capacity estimate is nearly 60% of the total population capacity calculated from fitting the Ricker model to the entire Wenatchee River basin data. If these estimates are correct, this means that about 40% of the current Wenatchee River basin smolt capacity is outside the Chiwawa River, Nason Creek, and White River watersheds. Hillman et al. (2017) report that over the period 1989 to 2016, on average, 76% of spring Chinook spawning occurs in the three watersheds. Thus, a large percentage of smolt capacity is generated outside the major spawning areas. We believe this highlights the importance of the mainstem Wenatchee River as a rearing area for juvenile spring Chinook.

Recommendations

Based on the simple analyses conducted in this report, we offer the following recommendations:

1. Where sufficient stock and recruitment data are available, and the data have sufficient contrast, then use population (stock-recruitment) modeling as the primary method to calculate population capacity and the number of spawners needed to produce the maximum number of recruits under current or average habitat conditions. Select the best fitting stock-recruitment model based upon AIC_c, unless other factors suggest otherwise, such as evidence for a biological mechanism. A biological mechanism supporting a Ricker function, for example, would be that there is a stock-dependent effect on the mortality of eggs and juveniles (i.e., mortality is proportional to the initial cohort size). When AIC_c values are not appreciably different, then select the model that is most useful (e.g., Ricker and smooth hockey stick models are easier to work with than the Beverton-Holt model).
2. Adult-to-adult data are the most relevant because they account for all life stages and delayed effects in freshwater (e.g., small size at migration), but they are also the most variable (i.e., low R^2). Therefore, adult-to-juvenile data (e.g., parr, yearling smolts, total migrants) are likely the most useful for determining freshwater population capacity. Where data are available, pre-spawn adult to spawning adult survival can also be assessed using population models to evaluate density dependence and pre-spawn adult capacity.
3. The population models used to estimate population capacity should also be used in reference streams so one can make comparisons of carrying capacities and density-corrected productivities. Unless there are good reasons for selecting a different juvenile life-stage, the default should be to use yearling smolts because they represent the capacity of the tributaries to produce yearlings and it is also a clear identification and quantification of a migrant life-stage.
4. In the absence of fish-habitat models, quantile regression can be used to estimate habitat capacity by calculating reference intervals for the population models. The percentage of the reference interval should be set using the error in the estimation of the recruits and the level of desire to exclude anomalous data. For example, if the 95% confidence interval is approximately 10% of the recruitment estimate, then the reference interval should be set at 90% (e.g., RI = 100% - C.I.%).
5. Where sufficiency conditions in (1) are not met, use habitat-based expansion of density at capacity for the most ecologically similar population. For example, use Twisp capacity estimates for habitat-based expansions in the Methow. The habitat expansion metric should be "total stream area weighted by intrinsic potential and temperature limited," unless there are good reasons for a different expansion. The primary idea is to exclude areas that are known to not produce fish because of passage, temperature, or other limitations.
6. Capacity estimates should be described within the context of the information that was used to derive estimates. For example, spawner distribution of hatchery-origin fish could influence estimates of capacity if they are within poor habitat. However, the capacity estimates do reflect the historic and current hatchery practices. It is unknown how the capacity estimates would change if a different hatchery program that produced different spawning distributions was to be implemented. However, if those data do become available, then capacity estimates can be revised. Similarly, significant enhancements (e.g.,

improved passage) or degradations (e.g., fire) in habitat can also change capacity and can be incorporated into future estimates of capacity.

7. Regardless of the method used to estimate capacity, always describe the limitations of the data and assumptions of the models. Note where assumptions are violated and how these violations could affect the results of the analysis.

References

Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, N.Y.

Cain, M., W. Bowman, and S. Hacker. 2009. Ecology. Sinauer Associates, Inc., Sunderland, Massachusetts.

Chapman, D. 1966. Food and space as regulators of salmonid populations in streams. American Naturalist 100:345-357.

Cooney T. and D. Holzer. 2006. Appendix C: Interior Columbia Basin Stream Type Chinook Salmon and Steelhead Populations: Habitat Intrinsic Potential Analysis. Interior Columbia Basin Technical Recovery Team.

Froese, R. 2008. The continuous smooth hockey stick: a newly proposed spawner-recruitment model. Journal of Applied Ichthyology 24:703-704.

Hilborn, R. and C. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York, N.Y.

Hillman, T., M. Miller, M. Johnson, C. Moran, J. Williams, M. Tonseth, C. Willard, S. Hopkins, B. Ishida, C. Kamphaus, T. Pearsons, and P. Graf. 2017. Monitoring and evaluation of the Chelan and Grant County PUDs hatchery programs: 2016 annual report. Report to the HCP and PRCC Hatchery Committees, Wenatchee and Ephrata, WA.

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, Portland, OR.

Murdoch, A., T. Pearsons, and T. Maitland. 2009. The number of redds constructed per female spring Chinook salmon in the Wenatchee River Basin. North American Journal of Fisheries Management 29:441-446.

Quinn, T. and R. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, New York, N.Y.

Ricker, W. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada 11:559-623.

APPENDIX 2: HATCHERY REPLACEMENT RATES

Based on ideas developed by the HETT, in February 2016, the HCP Hatchery Committees and PRCC Hatchery Subcommittee agreed to the following rules and HRR targets:

1. Use the estimated 40% HRR Target during 5-year statistical evaluation periods.
2. Use varying degrees of action depending on the numbers of years that annual HRR deviates from Target.
    a. Green Light (below Target for ≤ 2 years).
    b. Red Light (below Target for > 2 years).
3. Each program will have its own HRR target with the following exceptions.
    a. Nason Creek spring Chinook will use the Chiwawa Target (there are currently no data to calculate a target for Nason Creek spring Chinook).
    b. Methow and Chewuch spring Chinook will use the greater of their two Targets (they are MetComp stock and evaluated similarly).

Table 1. Release numbers and 5-year hatchery replacement rates (HRR) targets for Upper Columbia River Hatchery Programs.

Species	Owner	Program (Hatchery)	Basin (Purpose)	Smolts released[1]	5-Year HRR[2]
Steelhead	CPUD	Eastbank (Chiwawa)	Wenatchee (Conservation)	123,650	6.9
Steelhead	CPUD	Eastbank (Chiwawa)	Wenatchee (Safety Net)	123,650	6.9
Steelhead	DPUD	Wells (Wells)	Columbia (Safety Net)	160,000	26.5
Steelhead	DPUD	Wells (Wells)	Methow (Safety Net)	100,000	26.5
Steelhead	DPUD	Wells (Wells)	Twisp (Conservation)	48,000	26.5
Steelhead	GPUD	Wells (Omak)	Okanogan (Conservation)	100,000	7.3 [3]
SUM Chinook	CPUD	Eastbank (Chelan Falls)	Chelan (Conservation)	176,000	5.7
SUM Chinook	CPUD	Eastbank (Chelan Falls)	Chelan (Harvest)	400,000	5.7
SUM Chinook	CPUD, GPUD	Eastbank (Dryden)	Wenatchee (Conservation)	500,000	5.7
SUM Chinook	DPUD	Wells (Wells)	Columbia (Harvest)	320,000	3.0
SUM Chinook	GPUD	Eastbank (Carlton)	Methow (Conservation)	200,000	3.0
SUM Chinook	CCT	Chief Joseph	Okanogan (Harvest)	1,100,000	8.6
SPR Chinook	CPUD	Eastbank (Chiwawa)	Wenatchee (Conservation)	144,026	6.7
SPR Chinook	CPUD, DPUD, GPUD	Wells (Methow)	Methow (Conservation)	193,765	3.8
SPR Chinook	DPUD, GPUD	Wells (Twisp)	Methow (Conservation)	30,000	2.7
SPR Chinook	GPUD	Eastbank (Nason)	Wenatchee (Conservation)	223,670	6.7

[1] Release goal established by HCPs and adjusted by HC.
[2] Derived from Annual Reports.
[3] Harvest not included.

APPENDIX 3: PNI and pHOS Targets and Sliding Scales

Select CPUD, DPUD, and GPUD funded hatchery mitigation programs have PNI management targets, while others do not. Table 1 summarizes management strategies by species and population. Detailed information can be found in the sections that follow. Descriptions provided in the following sections are taken directly from HGMPs and/or issued and draft permits.

Table 1. Summary of management strategies by species and population.

Species	Population	Management Strategy	Comments
Spring Chinook	Wenatchee	Sliding Scale of PNI management	Details can be found in Section 2.0
	Methow	Two—population sliding scale PNI management	Details can be found in Section 3.0
	Okanogan	None Currently	Details can be found in Section 4.0
Steelhead	Wenatchee	Two-zone management.	Details can be found in 5.0
	Methow	In-development	Details forthcoming; Section 6.0
	Okanogan	None Currently	Details can be found in Section 7.0
Summer Chinook	Wenatchee	None Currently	Details can be found in Section 9.0
	Methow	None Currently	Details can be found in Section 10.0
	Okanogan	0.67; pHOS 0.30	Details can be found in Section 11.0
	Upper Columbia River (Wells and Chelan Falls)	None Currently	Details can be found in Section 12.0
Fall Chinook	Hanford Reach	0.67	Details can be found in Section 13.0

2.0 Wenatchee Spring Chinook

Wenatchee spring Chinook will be managed according to the sliding scale identified in the Wenatchee Spring Chinook Management Plan (2010) and Permit Numbers 18118 and 18121. The sliding scale is based upon the estimated number of natural origin spring Chinook over Tumwater Dam. As more information becomes available the sliding scale may be adjusted as a result of gaining a better understanding of the pre-spawn mortality rate and carrying capacity.

Table 2. Sliding scale of PNI goals based on natural origin spring Chinook run size expected to the Wenatchee River basin. Percentiles are based on adult returns observed between 1999 and 2008.

Percentile	NOR Run Size				PNI
	Chiwawa	Nason Creek	White	Wenatchee River (above TWD)	
>75th	>372	>350	>87	>910	≥ 0.80
50% - 75%	278-372	259-349	68-86	631-909	≥ 0.67
25% - 50%	209-277	176-258	41-67	525-630	≥ 0.50
10%-25%	176-208	80-175	20-40	400-524	≥ 0.40

<10th	<175	<80	<20	<400	Any PNI

3.0 Methow/ Chewuch Spring Chinook

The following sliding scale (Table 3) is presented in the April 14, 2016 draft Methow Hatchery Spring Chinook Section 10-Draft. It is anticipated that no further changes will be made to the sliding scale prior to issuance of the final permits.

Table 3. PUD PNI sliding scale calculations for a range of natural run sizes.

Natural Origin Returns	PUD pHOS	WNFH pHOS	PUD pNOB	2-Pop PNI	PUD PNI (equation)
<300	Ensure minimum of 500 total spawners				
300	0.40	0.2	0.75	0.67	0.67
500	0.40	0.2	0.80	0.68	0.76
900	0.30	0.15	1.00	0.78	0.80
1500	0.25	0.1	1.00	0.8	0.80
2000	0.25	0.1	1.00	0.8	0.80
2500	0.25	0.1	1.00	0.8	0.80

4.0 Okanogan Spring Chinook

The Okanogan spring Chinook program is a re-introduction effort implemented as a non-essential experimental population under ESA Section 10j to re-introduce spring Chinook into the Okanogan River. As a non-essential experimental population targeting re-introduction and establishment of a local population of spring Chinook, the Okanogan spring Chinook program will not conduct adult management actions to reduce the proportion of 10j hatchery fish on the spawning grounds or conduct broodstocking efforts in the Okanogan for a 10-year period (2014 – 2023), as such, no PNI or pHOS objectives have been identified for this program in this 10-year period.

CJH Program segregated production released into the mainstem Columbia River are non-listed Leavenworth stock released reared/acclimated/released at CJH. Although no PNI or pHOS targets are identified for the Okanogan 10j population, minimizing strays from the CJH segregated spring Chinook program is a program objective, as such, returning segregated program fish will be subject to directed harvest and aggressive adult surplusing at CJH to minimize straying to the Okanogan River Basin as well as other extant upper Columbia River spring Chinook populations. Stray targets for the segregated program are 5% or less stray rate (i.e. spawning contribution to other upper Columbia River spring Chinook populations).

5.0 Wenatchee Steelhead

Interim escapement goal for Wenatchee River steelhead will be 1,500 spawners with an additional goal of attaining an average PNI of 0.67 for the Wenatchee River basin population as a whole. To achieve the stated goal, the Wenatchee steelhead program will use a two-zone management approach wherein the upper basin (above TWD) will be managed for recovery using an integrated recovery program, a separate spawning escapement goal, and a PNI standard to achieve the overall basin goal of an average PNI over time of 0.67 (Table 4). Areas below TWD will be managed to minimize hatchery supplementation with a pHOS goal of < 0.10.

Steelhead returning upstream of TWD will be managed as an integrated recovery program with a pNOB goal of 1.0. The above TWD escapement goal will be 1,094 spawners. Working within this framework, pNOB will be maximized above TWD while pHOS will be minimized.

Table 4. Wenatchee steelhead two-zone management and PNI targets.

Location	Run Escapement Goal	pNOB Conservation Program	pNOB Safety Net Program	pHOS	PNI
Above TWD	1,094	1.0	0.0	Varies	Varies
Below TWD	406	N/A	N/A	< 0.10	< 0.67
Basin Total	1,500	N/A	N/A	Minimal	Average = 0.67

6.0 Methow Steelhead

Methow steelhead PNI targets are currently in development.

7.0 Okanogan Steelhead

Current program has no PNI goal. CTCR submitted an Okanogan steelhead HGMP to NOAA Fisheries on February 4, 2014. Within the HGMP provisions were included to allow a greater collection of natural-origin broodstock and multiple adult management strategies to address over-escapement of hatchery-origin steelhead to the spawning grounds. The HGMP also identified a near-term (1-4 years) and a long-term PNI objectives of 0.50 and > 0.67, respectively. Once NOAA has completed the consultation and issued a new permit, providing the opportunity to increase the proportion of natural-origin fish in the broodstock and additional adult management strategies, the program will adopt the PNI objectives and this Appendix can be amended accordingly.

8.0 Wells Columbia Mainstem Safety-net Steelhead

The Safety-Net Mainstem Columbia component released below Wells Dam will be managed primarily at the Wells Hatchery volunteer channel. The objective of the adult management of the Safety-Net Mainstem Columbia component is to prevent runs of this component from moving into natural spawning areas. This will be accomplished through in-river harvest and removal of volunteers at the Wells Hatchery outfall. There are no PNI goals for this component.

9.0 Wenatchee Summer Chinook

No PNI goals are established.

10.0 Methow Summer Chinook

No PNI goals are established.

11.0 Okanogan Summer Chinook

Okanogan summer/fall Chinook will be managed to achieve a 5-year rolling average PNI of 0.67 and pHOS of 0.30. Strategies to achieve that PNI target include up to 100% pNOB, aggressive removal of hatchery-origin Chinook in selective fisheries, at the Okanogan weir, and during surplusing at CJH ladder. Reduction in the number of juveniles released in the Okanogan River Basin (integrated program) is also a management option, should adult management actions be unable to control the proportion of hatchery fish on the spawning grounds to achieve that PNI target.

CJH segregated summer/fall Chinook program rears/acclimates/releases smolts into the mainstem Columbia River at CJH. Broodstock are 100% hatchery-origin, as such no PNI target for this production component. Stray rate (i.e. contribution to upper Columbia summer/fall Chinook populations) is 5% or less. Adult management on returning adults from the segregated program include fisheries, removal at the Okanogan weir, and removal at the CJH ladder.

12.0 Upper Columbia Summer Chinook (Chelan Falls and Wells)

No PNI goals are established. Chelan Falls and Wells FH summer Chinook programs are segregated harvest programs designed to provide opportunity for harvest. Adult returns are not intended to spawn naturally; therefore, there is no escapement goal for natural spawning areas. Adult returns will be managed to meet program objectives. Chelan Falls and Wells Hatchery summer Chinook are available for harvest in the ocean and Columbia River commercial, tribal, and recreational fisheries.

13.0 Priest Rapids Fall Chinook

The Hanford Reach fall Chinook population is intentionally supplemented by Grant PUD at the Priest Rapids Hatchery and the ACOE at the Priest Rapids and Ringold Springs hatcheries. Managers desire to achieve a population level PNI that includes all hatchery programs of $\geq$0.67. Grant PUD and the HSC do not have control over operation or expansion of the ACOE program and therefore will strive to operate the Priest Rapids Hatchery fall Chinook program in a way that does its fair share of achieving a population level PNI of 0.67.

APPENDIX 4: SPATIAL DISTRIBUTION OF SPAWNERS OR REDDS

Strategies for conservation programs typically intend that hatchery and naturally produced fish spawn together and in similar locations. However, in some cases, strategies may differ from this paradigm. In Table 1, conservation programs that have a spatial distribution management plan that deviates from similar to the natural spawning spatial distributions are presented. Otherwise, conservation programs are intended to have a spawning distribution similar to the natural origin spawning spatial distributions, as described by M&E Objective 5.3.

Table 1. Management targets for the spatial distribution of hatchery-origin redds for conservation programs that deviate from Objective 5.3.

Program	Target	Rational	Source
Carlton Summer Chinook	The observed spawning distribution of hatchery origin Methow summer Chinook from 2005-2010 represents the base-line spawner distribution for evaluating the performance of the hatchery program (i.e., M&E plan check-ins). It is acknowledged that this distribution is lower in the River than the spawning distribution of natural origin summer Chinook salmon.	Based upon an assessment of summer Chinook and ESA-listed spring Chinook abundance and spawner distribution, it was determined that an increase in summer Chinook spawning abundance in the upper most range of natural origin summer Chinook distribution or potentially above the current range may pose an unknown and potentially adverse impact to ESA listed spring Chinook. Due to the concern for spring Chinook, the HSC has endorsed an acclimation site in the Methow Basin that is lower in the basin than may be required to attain exact replication of natural and hatchery origin summer Chinook spawner distribution.	SOA 2011-02 Priest Rapids Coordinating Committee Hatchery Subcommittee Statement of Agreement on Monitoring & Evaluation (M&E) Objective for Spawning Distribution of Hatchery-Origin Summer Chinook
Dryden Summer Chinook	The observed spawning distribution of hatchery origin Wenatchee summer Chinook from 2008-2013 (previous 5 years to the current M&E check-in cycle) represents the base-line spawner distribution for evaluating the performance of the hatchery program (i.e., M&E plan check-ins).	The primary site endorsed by the HSC for Grant PUD overwinter acclimation of summer Chinook is the Dryden Pond, and is the current acclimation and release site for the existing summer Chinook supplementation program funded and owned by Chelan PUD. Because current data indicates that spawning distribution of hatchery summer Chinook from the existing program is lower in the Wenatchee River than natural origin spawners, expectations are that acclimation of Grant PUD's summer Chinook at Dryden Pond would continue to return hatchery origin summer Chinook that result in different spawning distributions for hatchery and natural summer Chinook.	Adapted from SOA 2011-02 Priest Rapids Coordinating Committee Hatchery Subcommittee Statement of Agreement on Monitoring & Evaluation (M&E) Objective for Spawning Distribution of Hatchery-Origin Summer Chinook

APPENDIX 5: WITHIN HATCHERY REARING TARGETS

Rearing Targets for Upper Columbia River Hatchery Programs. K-factor or fork length targets will be determined based on data from the pending "Five-Year Report."

Table 1. Numbers, fish per pound (fpp), coefficient of variation (CV), and condition factor (K) targets at release of Upper Columbia River Hatchery Programs.

Hatchery	Species	Life Stage	Basin	Release number	FPP	CV	K-factor
Methow	Spring Chinook	Yearling	Methow	193,765[1]	15	<10	TBD
Methow	Spring Chinook	Yearling	Twisp	30,000	15	<10	TBD
Chief Joseph	Spring Chinook	Yearling	Columbia	700,000	15	<10	TBD
Chief Joseph	Spring Chinook	Yearling	Okanogan	200,000	15	<10	TBD
Chiwawa	Spring Chinook	Yearling	Wenatchee	144,026	18	<10	TBD
Nason	Spring Chinook	Yearling	Wenatchee	223,670[3]	18-24	<10	TBD
Winthrop	Spring Chinook	Yearling	Methow	400,000	17	<10	TBD
Leavenworth	Spring Chinook	Yearling	Wenatchee	1.2 M	17	<10	TBD
Wells	Steelhead	Yearling	Columbia	160,000	6	<10	TBD
Wells	Steelhead	Yearling	Methow	100,000	6	<10	TBD
Wells	Steelhead	Yearling	Twisp	48,000	6	<10	TBD
Wells	Steelhead	Yearling	Omak	~100,000[4]	5-8	<10	TBD
Wells	Steelhead	Yearling	Okanogan	~100,000[4]	5-8	<10	TBD
Winthrop	Steelhead	Two year	Methow	200,000	4-6	<10	TBD
Chiwawa	Steelhead	Yearling	Wenatchee	247,300[5]	6	9.0	TBD
Wells	Summer Chinook	Subyearling	Columbia	480,000	50[6]	<7	TBD
Wells	Summer Chinook	Yearling	Columbia	320,000	10	<7	TBD
Chief Joseph	Summer Chinook	Subyearling	Columbia	400,000	50	<7	TBD
Chief Joseph	Summer Chinook	Subyearling	Okanogan	300,000	50	<7	TBD
Chelan Falls	Summer Chinook	Yearling	Chelan	576,000	13	9.0	TBD
Entiat	Summer Chinook	Yearling	Entiat	400,000	17	<10	TBD
Carlton	Summer Chinook	Yearling	Methow	200,000	13-17	<12	TBD
Chief Joseph	Summer Chinook	Yearling	Columbia	500,000	10	<7	TBD
Chief Joseph	Summer Chinook	Yearling	Okanogan	799,998[7]	10	<7	TBD
Dryden	Summer Chinook	Yearling	Wenatchee	500,001	18	9.0	TBD
Priest	Fall Chinook	Subyearling	Columbia	7.3 M[8]	50	<10	TBD
Ringold	Fall Chinook	Subyearling	Columbia	3.5 M	50	<10	TBD

[1] The total release includes the release of 108,249 into the Methow River at the Methow Fish Hatchery, 25,000 into the Methow River at the Goat Wall site, and 60,516 into the Chewuch River at the Chewuch Acclimation Facility.

[2] These fish come from Winthrop National Fish Hatchery (MetComp) eyed eggs.

[3] The total release includes 125,000 conservation fish and 98,670 safety net fish.

[4] The combined Okanogan and Omak steelhead release number is 100,000.

[5] The total release includes 66,771 fish into Nason Creek, 53,170 into the Chiwawa River, 102,359 into the Wenatchee River, and 25,000 into Blackbird Pond.

[6] The Wells subyearling Chinook are not reared to achieve a specific size target. The fish are released on a date to optimize survival and are grown to the largest size possible before release.

[7] The total release is divided equally among the Omak, Riverside, and Similkameen Acclimation Ponds.

[8] The total release consists of 5.6 m fall Chinook for the Grant PUD program and 1.7 M fall Chinook for the Army Corps of Engineers program.

APPENDIX 6: IDENTIFYING AND ANALYZING REFERENCE POPULATIONS

An important goal of supplementation is to increase spawning abundance and natural-origin recruitment of the supplemented population, and not reduce the productivity of the supplemented population. Indeed, a successful supplementation program must increase spawning abundance and natural-origin recruitment to levels above those that would have occurred without supplementation. There are several methods that can be used to test the effects of supplementation programs on these population metrics. One important method is to compare the performance of population metrics (e.g., spawning abundance, natural-origin recruitment, and productivity) in the supplemented population to those in un-supplemented (reference) populations. By comparing supplemented populations to reference populations, one can determine if the supplementation programs benefit, harm, or have no effect on the supplemented populations. These comparisons, however, are only valid if the performance of the reference populations is similar to the performance of the supplemented population prior to the period of supplementation. If the performance of the two populations differs significantly before any supplementation occurs, then any results from comparing the two populations after supplementation will be suspect. It is therefore important to select reference populations that are as similar as possible to the supplemented populations.

One of the goals of the Conceptual Approach to Monitoring and Evaluating the Chelan County PUD Hatchery Programs (Murdoch and Peven 2005) is to use reference populations to analyze the potential effects of hatchery supplementation programs on natural-origin salmon and steelhead spawner abundance and productivity[15]. Murdoch and Peven (2005) identified specific objectives to evaluate the performance of the program. For example, Objective 1 determines if the supplementation programs have increased the number of naturally spawning and naturally produced adults of the target population (supplemented population) relative to a reference population. Objective 7 determines if the proportion of hatchery fish on the spawning grounds affects the freshwater productivity (e.g., number of juveniles per redd) of supplemented streams when compared to reference streams. The relevant questions tested under each objective are as follows:

Objective 1:

- Is the annual change in the number of natural-origin recruits produced from the supplemented populations greater than or equal to the annual change in natural-origin recruits in an un-supplemented population?
- Is the change in natural replacement rates within the supplemented population greater than or equal to the change in natural replacement rates in an un-supplemented population?

Objective 7:

---

[15] Productivity is defined as adult recruits per spawner, where recruits are the number of adults produced from a given brood year (i.e., spawners plus adults harvested).

- Is the change in numbers of juveniles (smolts, parr, or emigrants) per redd in the supplemented population greater than or equal to that in an un-supplemented population?[16]

In this paper, we describe methods used to identify suitable reference streams and statistical techniques that can be used to compare reference populations with supplemented populations. Although we apply the methods described in this paper to Chiwawa spring Chinook salmon (hereafter referred to as Chinook), the methods should also apply to steelhead and other supplemented salmon stocks in the Upper Columbia Basin.

Identification of Reference Populations

Reference populations are an important component of an effectiveness monitoring design because they provide the standard by which treatment conditions are compared (ISRP and ISAB 2005; Murdoch and Peven 2005; Galbreath et al. 2008). Selecting appropriate reference areas and maintaining them over long periods of time is needed to establish the effectiveness of supplementation programs.

We developed a three-step process for identifying suitable reference populations (Figure 1). Each step serves as a filter. That is, potential reference populations are evaluated based on specific criteria under each step. Populations that pass through each step are considered suitable reference populations for a specific supplemented population.

---

[16] In this paper, we only address adult recruits, not juvenile recruits. This is because we were unable to find suitable reference populations for analysis of juveniles. However, the methods described in this paper would also apply to juveniles.

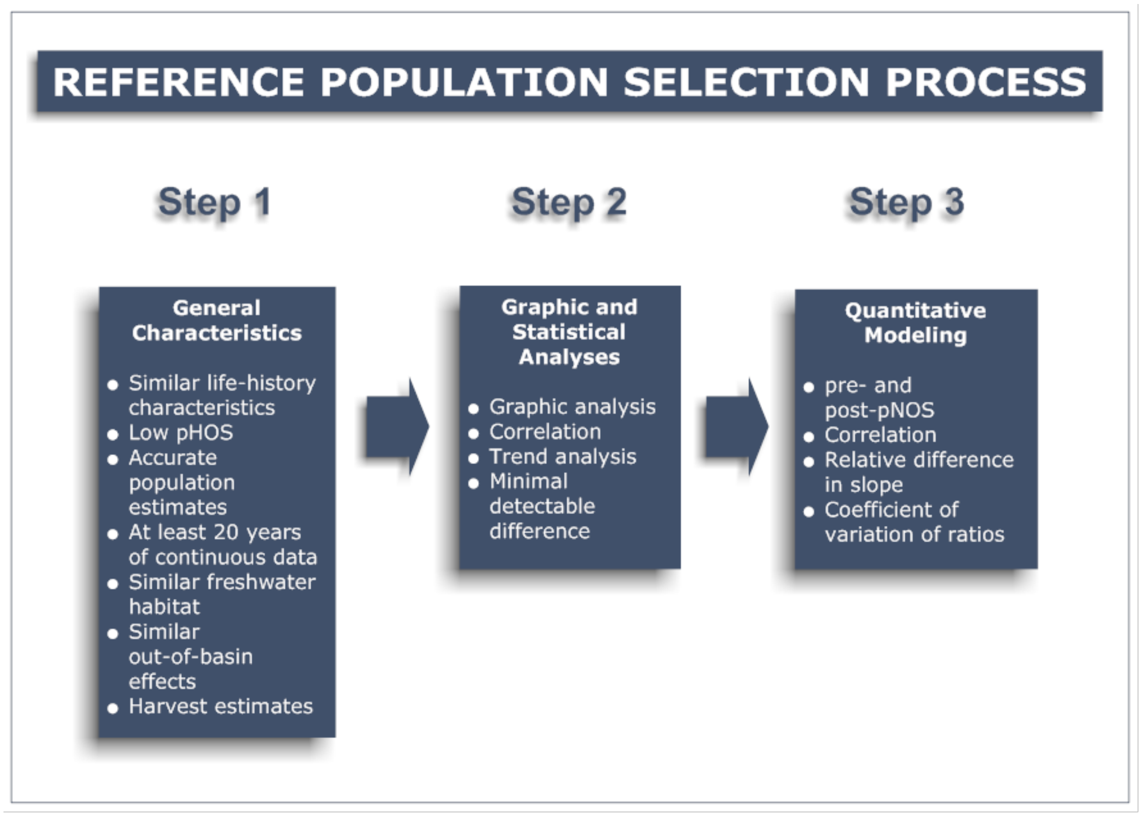


Figure 1. Criteria evaluated during each step in the process of identifying suitable reference populations.

Step 1: General Characteristics

Under step 1, potential reference populations are evaluated based on several general criteria. When compared to the supplemented population, potential reference populations should have:

- Similar life-history characteristics (e.g., run timing, migration characteristics, etc.).
- No or few hatchery fish in the reference area (pHOS < 10%).
- Accurate abundance estimates.
- Long time series of natural-origin abundance and productivity estimates (at least 20 years of continuous data).
- Similar trends in freshwater habitat.
- Similar out-of-basin effects (i.e., similar migration and ocean survivals).
- Harvest estimates for adjusting escapement estimates.

We used these criteria to begin the process of selecting suitable reference populations for the Chiwawa spring Chinook program. We began by identifying stream-type Chinook populations within the Columbia Basin. Galbreath et al. (2008; their Table 1) identified stream-type Chinook populations within the Columbia River Basin that may serve as suitable reference populations for hatchery programs. Supplementing their work with data from the NOAA Fisheries Salmon Population Summary Database, we identified 18 candidate stream-type Chinook populations that may serve as reference populations for the Chiwawa supplementation program (Table 1).

Table 1. Populations of stream-type Chinook salmon and their comparison to Chiwawa spring Chinook.

Population	Similar life-history	No or few hatchery fish	Accurate abundance estimates	Long time series (at least 20 years)	Similar freshwater habitat impairments	Similar out-of-basin effects	Comments
Deschutes River	Yes	Yes	Yes	Yes	No	No	
John Day mainstem	Yes	Yes	Yes	Yes	No	No	
Middle Fk John Day	Yes	Yes	Yes	Yes	No	No	
North Fk John Day	Yes	Yes	Yes	Yes	No	No	
Granite Creek	Yes	Yes	Yes	Yes	No	No	
Wenaha River	Yes	No	Yes	Yes	Yes	No	Hatchery strays (>10%)
Minam River	Yes	No	Yes	Yes	Yes	No	Hatchery strays (>10%)
Slate Creek	Yes	Yes	Yes	No	No	No	
Secesh River	Yes	Yes	Yes	Yes	Yes	No	
Middle Fk Salmon River	Yes	Yes	Yes	No	No	No	Fair productivity est.
Big Creek	Yes	Yes	Yes	Yes	No	No	
Camas Creek	Yes	Yes	Yes	Yes	No	No	Fair productivity est.
Loon Creek	Yes	Yes	Yes	Yes	No	No	Fair productivity est.
Sulphur Creek	Yes	Yes	Yes	Yes	No	No	
Bear Valley Creek	Yes	Yes	Yes	Yes	No	No	
Marsh Creek	Yes	Yes	Yes	Yes	Yes	No	
North Fk Salmon River	Yes	Yes	No	No	Yes	No	
Lemhi River	Yes	Yes	Yes	Yes	No	No	
East Fk Salmon River	Yes	No	Yes	Yes	No	No	Hatchery strays (>10%)
Valley Creek	Yes	No	Yes	Yes	No	No	Hatchery strays (>10%)
Chamberlain Creek	Yes	Yes	Yes	No	Yes	No	
Naches River	Yes	Yes	Yes	Yes	Yes	No	
Little Wenatchee River	Yes	No	Yes	Yes	Yes	Yes	Hatchery strays (>10%)
Entiat River	Yes	No	Yes	Yes	No	No	Hatchery release ending

We then assessed the accuracy and length of the series of abundance estimates. We assumed that abundance estimates generated from expanded redd counts or adjusted weir counts would compare well with estimates in the Chiwawa Basin, which were based on expanded redd counts. In addition, we looked for populations that had an abundance data series that extended from at least 1981 to present. Based on this analysis, we identified 18 populations with abundance estimates that could be compared to those from the Chiwawa Basin (Table 1).

Next, we determined if the potential reference populations came from watersheds with habitat conditions similar to those in the Chiwawa Basin. For this exercise, we searched recovery plans and draft recovery plans to identify tributary factors that limit Chinook abundance, productivity, and survival within the reference populations. We compared these factors with those limiting

Chinook salmon in the Chiwawa Basin. Based on this analysis, we identified eight populations with habitat impairments similar to those in the Chiwawa Basin (Table 1).

Finally, we examined the potential reference populations to see if they experienced out-of-basin effects similar to spring Chinook from the Chiwawa Basin. In this case, we compared the number of mainstem dams that each potential reference population passes during migration. Six of the potential reference populations pass less than six mainstem dams; the other populations pass eight mainstem dams (Table 1). Only the Little Wenatchee population passes seven dams, similar to the Chiwawa population.

In sum, there were no reference populations that matched the Chiwawa spring Chinook population on all the criteria identified above. Differential out-of-basin effects and freshwater habitat conditions prevented most reference populations from matching with Chiwawa spring Chinook. However, some of the potential reference populations were similar to the Chiwawa population on several criteria and warranted further investigation. We selected the following populations for further investigation: Sesech River, Marsh Creek, Naches River, Little Wenatchee, and Entiat River.

We included the Little Wenatchee because it is within the Wenatchee River basin and experiences similar out-of-basin effects and has the same climatic and environmental conditions as the Chiwawa. A confounding effect with the Little Wenatchee is that Chiwawa hatchery fish have strayed into the Little Wenatchee. However, straying of Chiwawa hatchery fish should decrease with the change in source water to the Chiwawa acclimation ponds in 2006. We also included the Entiat River because it is an adjacent basin to the Chiwawa and experiences similar climatic and environmental conditions. The spring Chinook hatchery program that has operated in the Entiat since 1975 has been discontinued. Therefore, this population offers a unique opportunity to compare the Chiwawa population to a population in which the hatchery program has been discontinued.

Step 2: Graphic and Statistical Analysis

Graphic Analysis

Although we were unable to find potential reference populations that matched with the Chiwawa population on all criteria considered under Step 1, spawner abundance, natural-origin recruits (NORs), and productivity of some of the potential reference populations may nevertheless track closely with the Chiwawa population. If the time series of abundance, NORs, and productivity of a potential reference population tracks closely with the abundance, NORs, and productivity of the Chiwawa population, the reference population may provide a reasonable reference condition for testing the effects of supplementation on the Chiwawa population.

Under Step 2, we used graphing techniques to examine the relationship of abundance, NORs, and productivity between the Chiwawa population and the five reference populations (Sesech River, Marsh Creek, Naches River, Little Wenatchee, and Entiat River). We compiled spawner abundance, NORs, and productivity data from local biologists and the NOAA Fisheries Salmon Population Summary Database. We then compared time series plots of spawner abundance, NORs, and productivity data of potential reference populations with the Chiwawa population (Figures 2, 3, and 4; plots on the left side of figures). The time series only included the period 1981 to 1992, which represented the period before supplementation of the Chiwawa population (pre-treatment period). We also plotted the relationship between the abundance, NORs, and productivity of each potential reference population to the Chiwawa population (Figures 2, 3, and 4; plots on right side of figures). These plots show whether the reference populations closely tracked the Chiwawa population. As a point of reference, data points that fall along the dashed line would represent a perfect relationship between the two populations (i.e., both populations have identical abundance, NORs, and productivity estimates). While a perfect relationship between two independent populations is unrealistic, a strong linear relationship between the two populations indicates populations with similar trends.

Based on analysis of spawner abundance, the Naches River time series tracked more closely with the abundance of Chiwawa spring Chinook than did the other potential reference populations. The poor relationship with the other potential reference streams was largely because of the relatively high abundance of Chiwawa spring Chinook during the mid-1980s. As with spawner abundance, analyses of NORs indicated a close relationship between the Naches and Chiwawa populations. The other potential reference populations tracked poorly with the Chiwawa. The analyses of productivity indicated close relationships between potential reference populations and the Chiwawa population. The Naches, Sesech, and Little Wenatchee populations tracked the closest with the Chiwawa population.

When analyzing the potential effects of a supplementation program on fish performance, it is common to transform the data to meet various assumptions of statistical analysis. The most common transformation used to adjust abundance, NORs, and productivity data is the natural logarithm (LN or $\log_e$). We therefore transformed the spawner abundance, NORs, and productivity data using LN and re-plotted the relationships between the potential reference populations and the Chiwawa population (Figures 5, 6, and 7). We added 1 to each observation before taking its logarithm to avoid taking the logarithm of 0, which is undefined (note that the LN of 1 is 0).

By transforming spawner abundance, NORs, and productivity data, most of the potential reference populations tracked more closely with the Chiwawa population. The Naches, Entiat, and Little Wenatchee abundance data tracked the closest with the Chiwawa abundance data (Figure 5). For NORs, Marsh Creek and the Little Wenatchee populations tracked the closest with the Chiwawa (Figure 6). For productivity, the Naches, Sesech, and Little Wenatchee tracked the closest with the Chiwawa (Figure 7).

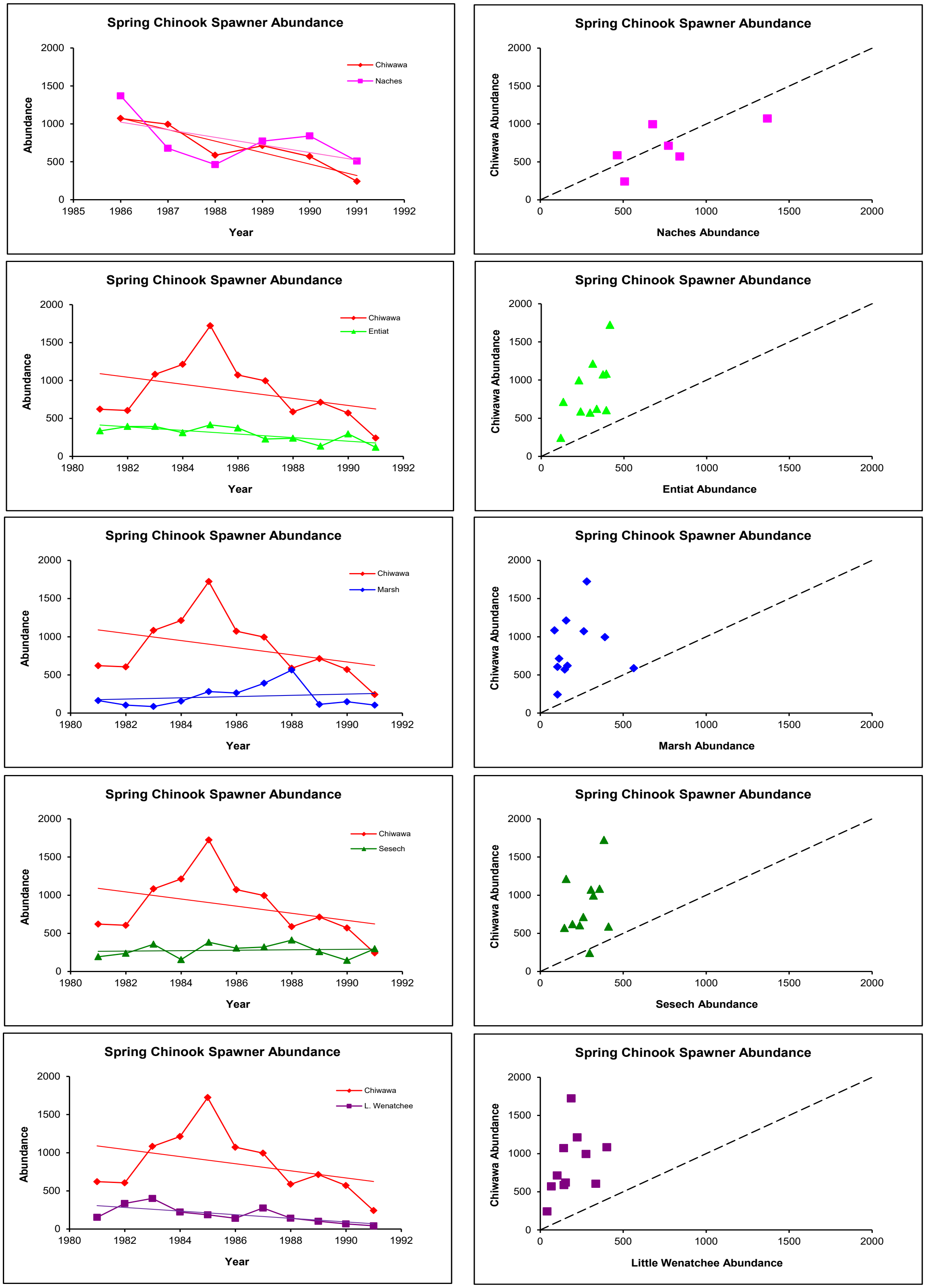


Figure 2. Time series of spawner abundance of potential reference populations and the Chiwawa spring Chinook population before the Chiwawa population was supplemented with hatchery fish.

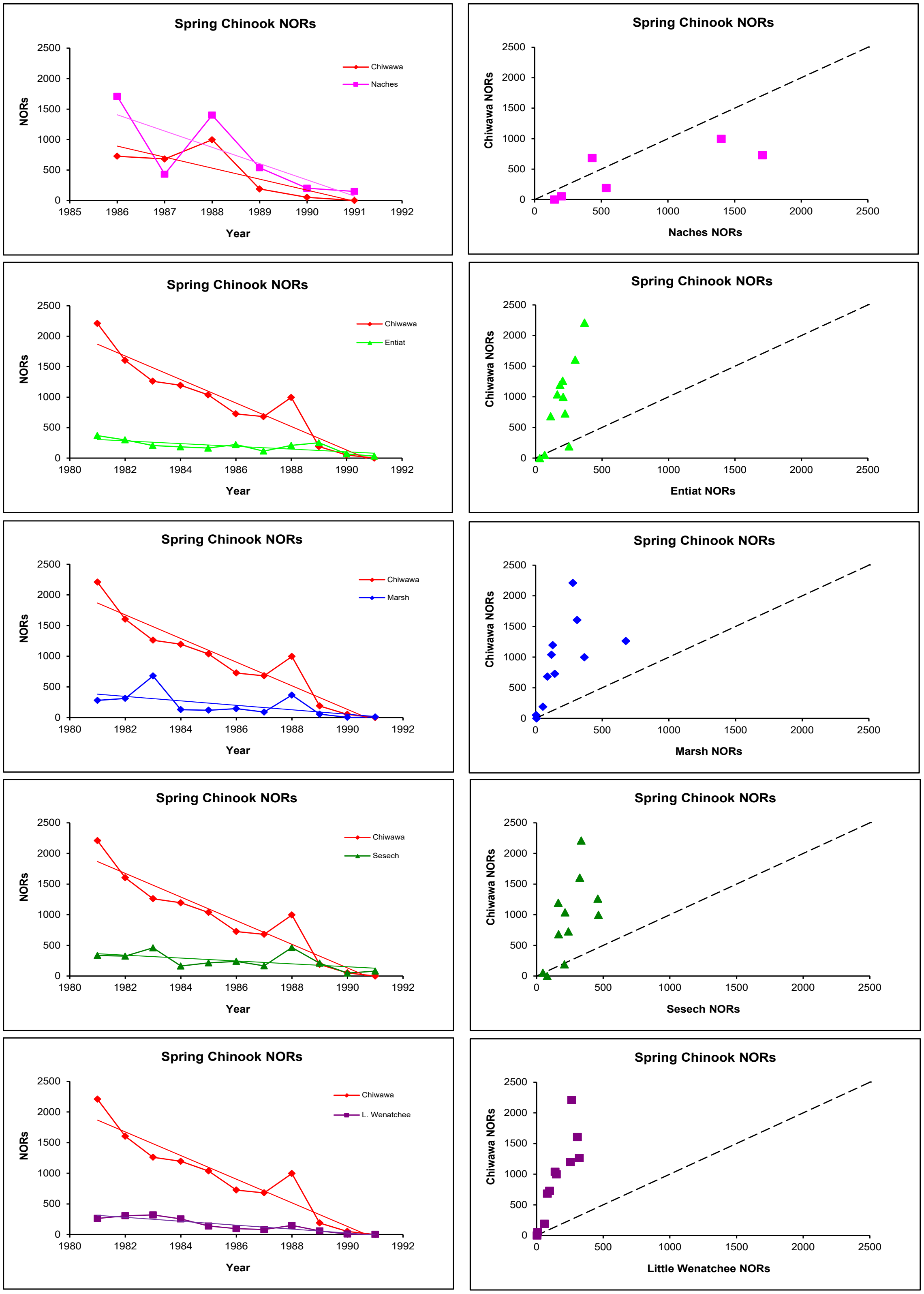


Figure 3. Time series of natural-origin recruits (NORs) of potential reference populations and the Chiwawa spring Chinook population before the Chiwawa population was supplemented with hatchery fish.

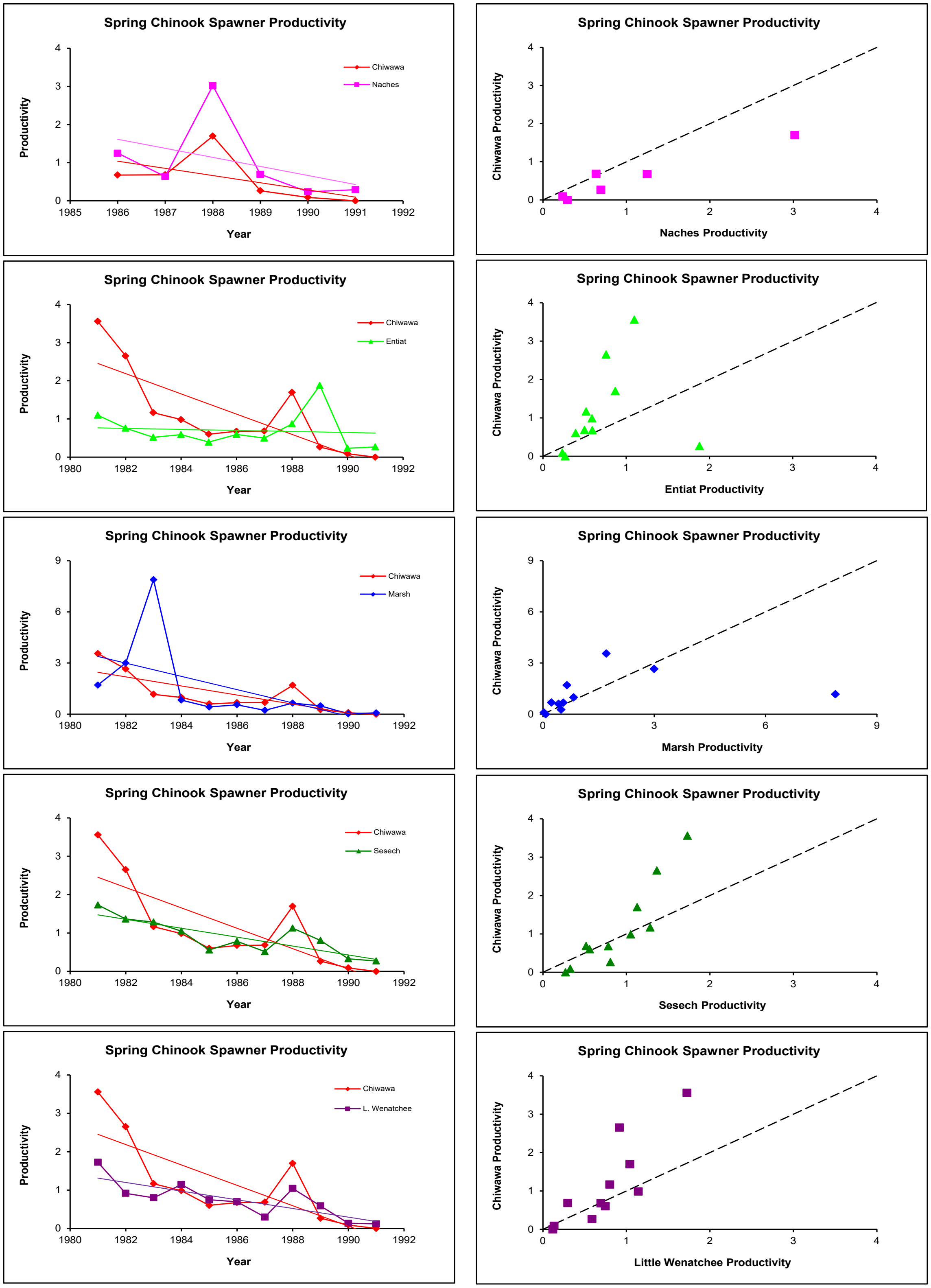


Figure 4. Time series of adult productivity of potential reference populations and the Chiwawa spring Chinook population before the Chiwawa population was supplemented with hatchery fish.

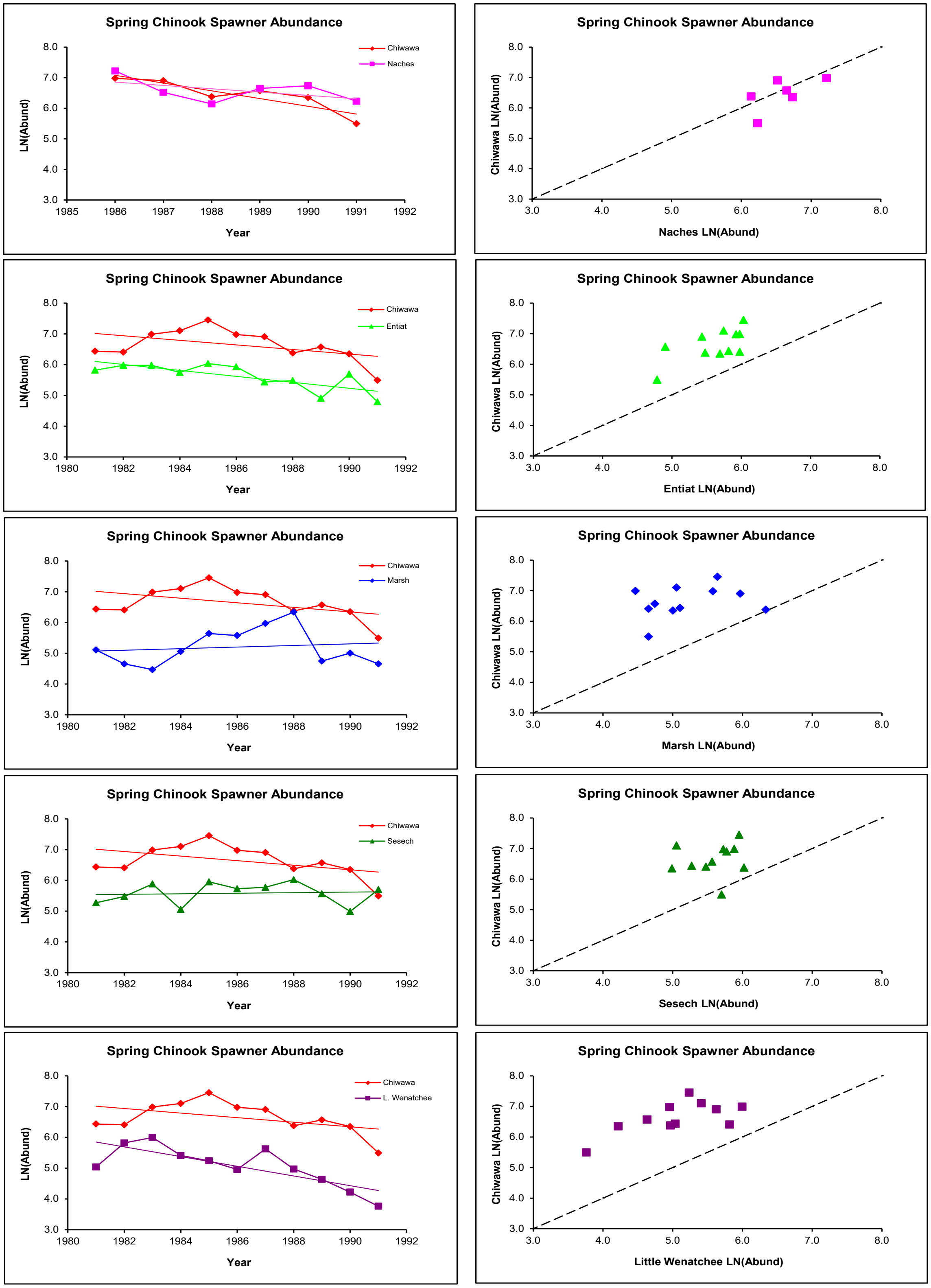


Figure 5. Time series of natural log spawner abundance of potential reference populations and the Chiwawa spring Chinook population before the Chiwawa population was supplemented with hatchery fish.

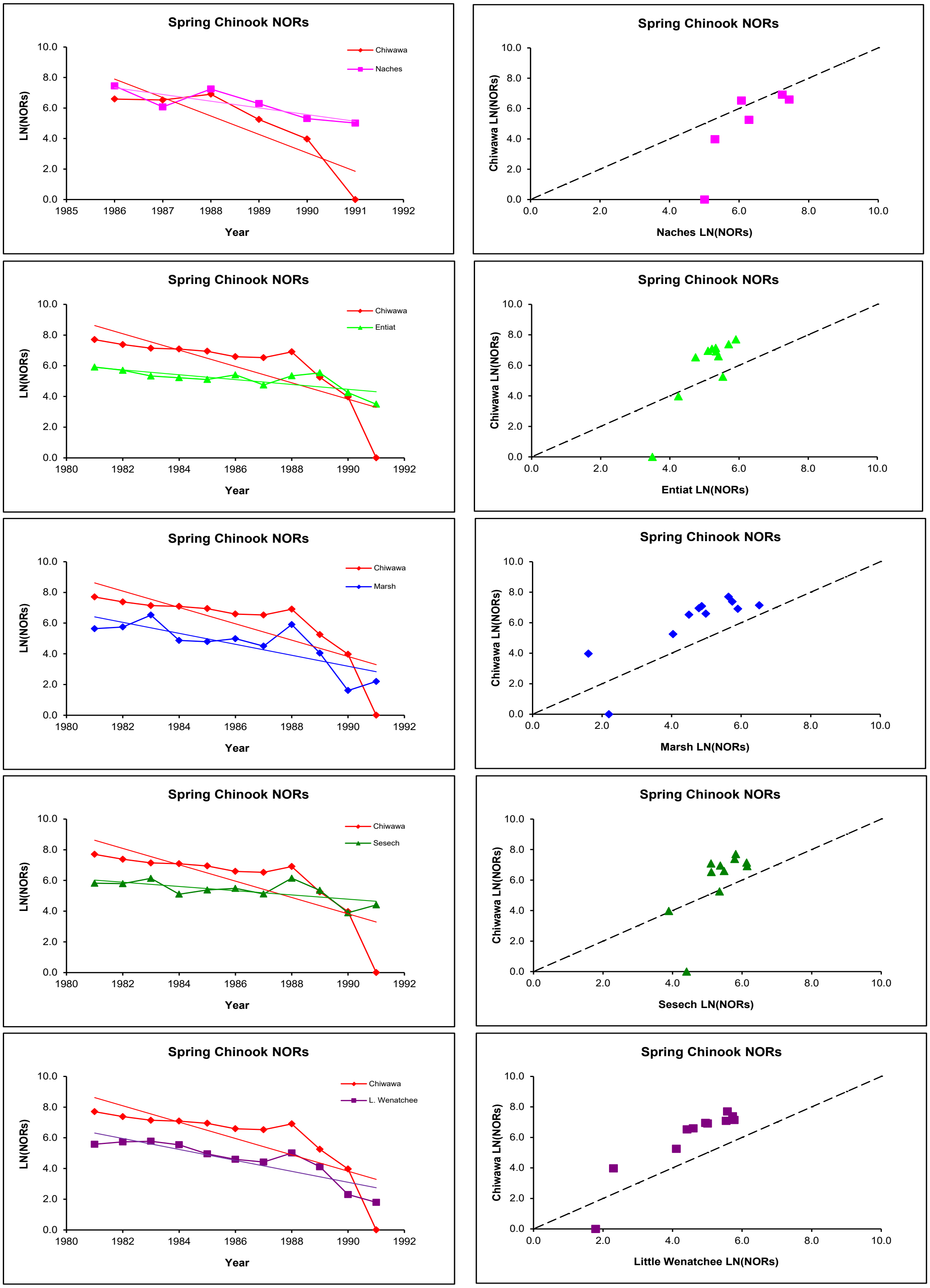


Figure 6. Time series of natural log natural-origin recruits (NORs) of potential reference populations and the Chiwawa spring Chinook population before the Chiwawa population was supplemented with hatchery fish.

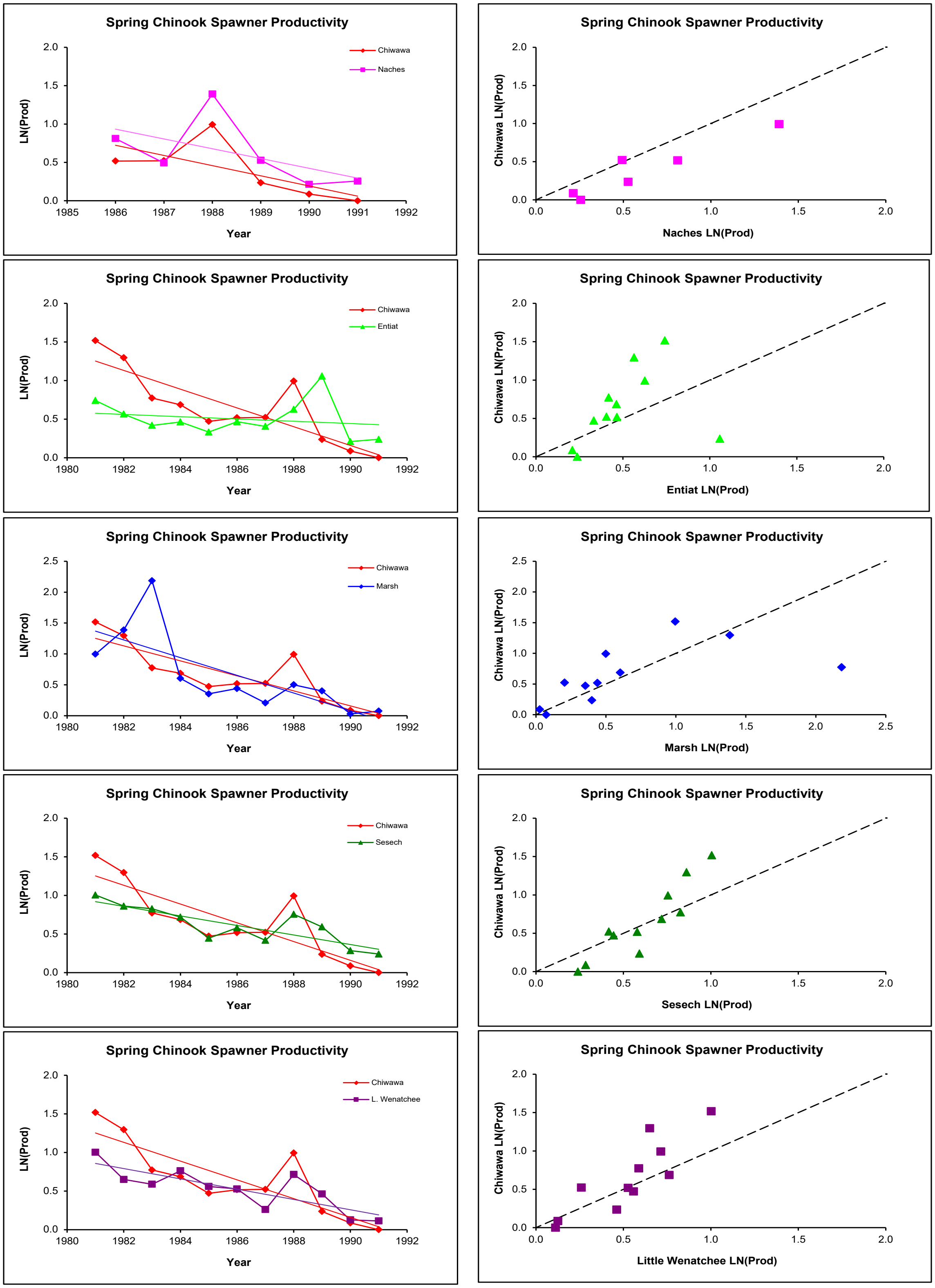


Figure 7. Time series of natural log adult productivity of potential reference populations and the Chiwawa spring Chinook population before the Chiwawa population was supplemented with hatchery fish.

Correlations and Trends

Other methods for evaluating the suitability of potential reference populations under Step 2 include correlation and trend analyses. For correlation analysis, we simply calculated the Pearson correlation coefficient, which is an index of the strength of the association between the potential reference populations and the Chiwawa population. The coefficient ranges from -1 to 1, where a value near 1 or -1 represents that strongest association between the populations. A value of 0 means no association. We used only spawner abundance, NORs, and productivity data during the pre-treatment period (1981-1992). We assumed that populations with coefficients greater than 0.6 represented reasonable reference conditions.

For trend analyses, we used least squares techniques to compute a straight-line trend through the spawner abundance and productivity data for the potential reference populations and the Chiwawa population. Trends were fit to the pre-treatment time series data (1981-1992). We then used t-tests to determine if the slopes of the trends between potential reference populations and the Chiwawa population differed significantly.

It is important to note that time-series trend analyses are susceptible to temporal correlations in the data. Autoregressive integrated moving average (ARIMA) models can be used to describe the correlation structure in temporal data (Gotelli and Ellison 2004). However, these models require a long time series (N > 40) and therefore we could not use them to model the spring Chinook data. As such, we were unable to correct for any temporal correlation that may exist within the time series.

Tests of correlation with spawner abundance data indicated that the Naches River closely correlated with the Chiwawa population (Table 2). There was no difference in abundance trends between the potential reference populations and the Chiwawa population (Table 2; Figure 2). For NORs, all potential reference populations correlated with the Chiwawa population (Table 2). However, trends in NORs of all reference populations, except Naches, differed significantly from the Chiwawa population (Table 2; Figure 3). For productivity, the Naches, Sesech, and Little Wenatchee correlated with the Chiwawa population (Table 2). Only the Entiat productivity trend differed significantly from the Chiwawa population trend (Table 2; Figure 4).

Table 2. Pearson correlation coefficients and t-test results comparing slopes of trends between potential reference populations and the Chiwawa spring Chinook population; d.f. = degrees of freedom and for correlation coefficients, an asterisk (*) indicates significance at $P < 0.05$.

Reference populations	Pearson correlation coefficient	t-test on slopes		
		t-value	d.f.	P-value
Spawner Abundance Data				
Naches	0.684*	-0.659	8	0.528
Entiat	0.598*	-0.596	18	0.559
Marsh	0.147	-1.341	18	0.197
Sesech	0.274	-1.265	18	0.222
Little Wenatchee	0.399	-0.591	18	0.562
Natural-Origin Recruits				
Naches	0.803*	0.666	8	0.524
Entiat	0.795*	-7.495	18	0.000

Reference populations	Pearson correlation coefficient	t-test on slopes		
		t-value	d.f.	P-value
Marsh	0.605*	-5.786	18	0.000
Sesech	0.648*	-6.874	18	0.000
Little Wenatchee	0.880*	-7.206	18	0.000
Productivity Data				
Naches	0.960*	0.169	8	0.870
Entiat	0.272	-3.057	18	0.007
Marsh	0.320	0.605	18	0.553
Sesech	0.903*	-2.059	18	0.054
Little Wenatchee	0.848*	-2.065	18	0.054

We also ran correlation and trend analyses on natural-log transformed spawner abundance, NORs, and productivity data. These analyses indicated that the Naches, Entiat, and Little Wenatchee abundance data correlated with the Chiwawa population data (Table 3). None of the abundance trends of the potential reference populations differed significantly from the Chiwawa population trend (Table 3; Figure 5). For NORs, all potential reference populations correlated with the Chiwawa population (Table 3). Only trends in NORs of the Entiat and Sesech differed significantly from the Chiwawa population (Table 2; Figure 6). For productivity, the Naches, Marsh, Sesech, and Little Wenatchee correlated with the Chiwawa population data (Table 3). Only the Entiat productivity trend differed significantly from the Chiwawa population trend (Table 3; Figure 7).

Table 3. Pearson correlation coefficients and t-test results comparing slopes of trends between potential reference populations and the Chiwawa spring Chinook population; d.f. = degrees of freedom and for correlation coefficients, an asterisk (*) indicates significance at $P < 0.05$. Analyses were conducted on natural-log transformed abundance and productivity data.

Reference populations	Pearson correlation coefficient	t-test on slopes		
		t-value	d.f.	P-value
LN Spawner Abundance Data				
Naches	0.642*	-1.323	8	0.222
Entiat	0.652*	0.412	18	0.685
Marsh	0.294	-1.324	18	0.202
Sesech	0.149	-1.431	18	0.170
Little Wenatchee	0.670*	1.325	18	0.202
LN Natural-Origin Recruits				
Naches	0.824*	-1.985	8	0.082
Entiat	0.886*	-2.563	18	0.019
Marsh	0.830*	-1.038	18	0.313
Sesech	0.730*	-2.664	18	0.016
Little Wenatchee	0.927*	-1.150	18	0.265
LN Productivity Data				

Reference populations	Pearson correlation coefficient	t-test on slopes		
		t-value	d.f.	P-value
Naches	0.944*	-0.042	8	0.968
Entiat	0.373	-3.043	18	0.007
Marsh	0.610*	0.428	18	0.674
Sesech	0.913*	-2.050	18	0.055
Little Wenatchee	0.862*	-1.811	18	0.087

In summary, based on correlation, trend, and graphic analyses, the Naches, Entiat, and Little Wenatchee populations appear to be reasonable reference populations for comparing spawner abundance data with Chiwawa data. For NORs, the Naches, Marsh, and Little Wenatchee appear to be reasonable reference populations. For productivity, the Naches, Marsh, Sesech, and Little Wenatchee populations appear to be reasonable reference populations for the Chiwawa population.

Minimal Detectable Differences (MDD)

Given a suite of potential reference populations, it is important to conduct power analyses to determine the minimum differences that can be detected when comparing the reference populations to the supplemented population. As a final exercise under Step 2, we examined potential reference populations for the smallest minimal detectable differences. Before conducting power analyses, several decisions needed to be made, including what statistical procedures will be used to analyze the data, the desired level of statistical power (probability of rejecting a false null hypothesis), the size of the type-I error (the probability of rejecting a true null hypothesis of no difference), and the number of samples (i.e., years) included in the analysis. In this case, the number of samples represents the number of treatment (supplementation) years. The number of pre-treatment years (1981-1992) was based on the number of years of quality data available for Chiwawa spring Chinook and potential reference populations.

We designed the study as a modified BACI (Before-After, Control-Impact) design, which includes replication before and after supplementation in both the treated (T) population and the reference (R) populations. A common approach used to analyze data from BACI designs includes analysis of difference scores (Stewart-Oaten et al. 1992; Smith et al. 1993). Differences are calculated between paired treatment and reference population scores (i.e., T-R). Another approach is to calculate ratios (treatment/reference; T/R) for paired treatment and reference population scores (Skalski and Robson 1992). Finally, differences in annual changes in paired treatment and reference population scores can be calculated (i.e., ΔT-ΔR) (Murdoch and Peven 2005; Hays et al. 2006).[17] These derived difference and ratio scores are then analyzed for a before-after treatment effect with a two-sample t-test, Aspin-Welch modification of the t-test, or a randomization test. For power analyses, we calculated minimal detectable differences assuming the use of an independent two-sample t-test with a type-I error rate of 0.05, power of 0.80 (beta or type-II error rate of 0.20), and sample sizes (treatment years) of 5, 10, 15, 20, 25, and 50 years.

---

[17] The difference of annual difference scores was estimated by first subtracting the population parameter (e.g., spawner abundance) in year 2 from year 1. This continues for all years in the data series for both treatment ($T_{t+1} - T_t$) and reference populations ($R_{t+1} - R_t$). We then calculated differences between paired treatment and reference annual difference scores [$(T_{t+1} - T_t) - (R_{t+1} - R_t) = \Delta T - \Delta R$].

The power analysis calculated the minimal detectable difference between mean difference or ratio scores before and during supplementation. We used existing data to calculate variances for the pre-supplementation and supplementation periods. Thus, variances were known and unequal. For both spawner abundance and NORs, the null hypothesis tested was that the mean difference or ratio before supplementation equaled the mean difference or ratio during supplementation. The alternative hypothesis was that the mean difference or ratio before supplementation was less than the mean difference during supplementation (one-tail test; Difference < 0). For productivity, the null hypothesis tested was that the mean difference or ratio before supplementation equaled the mean difference or ratio during supplementation. The alternative hypothesis was that the mean difference or ratio before supplementation was greater than the mean difference during supplementation (one-tail test; Difference > 0).

Based on spawner abundance data, power analysis indicated that the Sesech-Chiwawa pairing consistently produced the smallest detectable differences (Table 4). However, when the abundance data were transformed using natural logs, the Entiat-Chiwawa pairing produced the smallest detectable difference (Table 5). Minimal detectable differences, based on mean difference scores on untransformed data and a treatment period of 20 years, ranged from 334 to 394 adult spawners; transformed data ranged from 0.479 to 1.010. These analyses indicate that the Naches, Entiat, Sesech, and Little Wenatchee populations appear to be reasonable reference populations for comparing spawner abundance data with Chiwawa data. The Marsh Creek population produced some of the largest detectable differences and based on these analyses may not be a reasonable reference population.

Table 4. Minimal detectable differences between mean difference and ratio scores before and during supplementation. Analyses were conducted on spawner abundance data.

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
T-R	5	638	604	560	396	652
	10	464	448	444	354	481
	15	405	395	406	341	424
	20	376	368	387	334	394
	25	358	352	376	331	376
	50	322	319	354	323	340
T/R	5	0.600	2.084	39.251	1.569	5.498
	10	0.506	1.548	24.729	1.508	3.828
	15	0.478	1.367	19.646	1.490	3.256
	20	0.465	1.275	16.828	1.481	2.954
	25	0.458	1.219	14.974	1.475	2.765
	50	0.447	1.105	10.573	1.465	2.366
ΔT-ΔR	5	1,049	761	717	518	766
	10	750	542	539	411	547
	15	650	467	480	376	473
	20	598	429	450	359	434

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
	25	567	405	431	348	410
	50	506	355	395	329	361

Table 5. Minimal detectable differences between mean difference and ratio scores before and during supplementation. Analyses were conducted on natural-log transformed spawner abundance data.

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
T-R	5	0.975	0.871	2.061	0.828	1.013
	10	0.721	0.613	1.375	0.648	0.722
	15	0.637	0.525	1.138	0.588	0.623
	20	0.595	0.479	1.010	0.559	0.571
	25	0.569	0.450	0.928	0.541	0.539
	50	0.521	0.390	0.749	0.505	0.473
T/R	5	0.157	0.162	2.343	0.160	0.368
	10	0.116	0.115	1.474	0.125	0.247
	15	0.102	0.099	1.170	0.114	0.206
	20	0.095	0.090	1.001	0.108	0.183
	25	0.091	0.085	0.890	0.104	0.169
	50	0.082	0.075	0.625	0.098	0.138
ΔT-ΔR	5	1.261	1.288	3.076	1.160	1.467
	10	0.898	0.900	2.020	0.887	1.001
	15	0.776	0.768	1.653	0.797	0.840
	20	0.713	0.698	1.463	0.751	0.755
	25	0.675	0.655	1.325	0.724	0.701
	50	0.600	0.564	1.038	0.670	0.585

Based on NORs, power analysis indicated that the Entiat-Chiwawa, Marsh-Chiwawa, and Little Wenatchee-Chiwawa pairings produced the smallest detectable differences (Table 6). When NORs were transformed using natural logs, the Little Wenatchee-Chiwawa pairing produced the smallest detectable difference (Table 7). Minimal detectable differences, based on mean difference scores on untransformed data and a treatment period of 20 years, ranged from 483 to 640 NORs; transformed data ranged from 0.958 to 2.262. These analyses indicate that the Entiat, Marsh, and Little Wenatchee populations appear to be reasonable reference populations for comparing NORs with Chiwawa data.

Table 6. Minimal detectable differences between mean difference and ratio scores before and during supplementation. Analyses were conducted on natural-origin recruits.

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
T-R	5	1,139	541	573	630	546
	10	809	511	515	550	503
	15	698	502	498	526	489
	20	640	497	489	514	483
	25	604	494	484	507	479
	50	534	489	474	493	472
T/R	5	0.469	2.538	5.196	1.976	6.973
	10	0.451	2.183	4.183	1.894	5.118
	15	0.446	2.072	3.854	1.869	4.492
	20	0.445	2.017	3.691	1.857	4.170
	25	0.444	1.986	3.594	1.850	3.973
	50	0.443	1.924	3.405	1.836	3.572
ΔT-ΔR	5	1,639	500	519	609	531
	10	1,239	386	409	433	396
	15	1,109	348	374	372	351
	20	1,046	329	356	341	328
	25	1,009	318	346	321	314
	50	943	295	325	281	285

Table 7. Minimal detectable differences between mean difference and ratio scores before and during supplementation. Analyses were conducted on natural-log transformed natural-origin recruits.

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
T-R	5	2.380	1.646	1.967	2.247	1.174
	10	2.291	1.479	1.505	1.835	1.026
	15	2.270	1.428	1.351	1.702	0.980
	20	2.262	1.403	1.273	1.636	0.958
	25	2.258	1.389	1.227	1.597	0.945
	50	2.253	1.361	1.133	1.522	0.920
T/R	5	0.322	0.332	0.739	0.398	0.356
	10	0.301	0.289	0.581	0.334	0.322
	15	0.296	0.275	0.530	0.314	0.312
	20	0.294	0.269	0.504	0.305	0.307
	25	0.293	0.265	0.488	0.299	0.304

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
	50	0.291	0.258	0.458	0.288	0.298
ΔT-ΔR	5	2.858	2.400	2.355	3.283	2.109
	10	2.560	1.714	1.881	2.311	1.552
	15	2.485	1.481	1.728	1.979	1.365
	20	2.456	1.360	1.652	1.805	1.269
	25	2.443	1.285	1.607	1.697	1.210
	50	2.430	1.130	1.519	1.471	1.092

Using untransformed productivity data, power analysis indicated that the Little Wenatchee-Chiwawa pairing consistently produced the smallest detectable differences (Table 8). The Marsh-Chiwawa pairings produced the largest detectable differences. When we analyzed natural-log transformed productivity data, the Naches-Chiwawa and Little Wenatchee-Chiwawa pairings produced the smallest detectable differences (Table 9). Minimal detectable differences, based on mean difference scores on untransformed data and a treatment period of 20 years, ranged from 0.754 to 1.839; transformed data ranged from 0.277 to 0.477. These analyses indicate that the Naches, Entiat, Sesech, and Little Wenatchee populations appear to be reasonable reference populations for comparing productivity data with Chiwawa data. The Marsh Creek population produced some of the largest detectable differences and based on these analyses may not be a reasonable reference population.

Table 8. Minimal detectable differences between mean difference and ratio scores before and during supplementation. Analyses were conducted on productivity data.

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
T-R	5	2.181	1.382	2.033	3.517	1.192
	10	1.442	1.119	1.900	2.265	0.901
	15	1.186	1.033	1.859	1.828	0.804
	20	1.047	0.991	1.839	1.588	0.754
	25	0.959	0.966	1.828	1.432	0.724
	50	0.764	0.917	1.806	1.074	0.664
T/R	5	1.364	1.773	0.863	0.876	2.167
	10	1.095	1.359	0.831	0.687	1.587
	15	1.011	1.221	0.822	0.625	1.391
	20	0.971	1.152	0.817	0.594	1.290
	25	0.949	1.110	0.814	0.575	1.228
	50	0.910	1.027	0.908	0.538	1.102
ΔT-ΔR	5	3.298	1.864	3.211	4.420	1.942
	10	2.263	1.382	2.968	2.811	1.291

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
	15	1.909	1.220	2.894	2.248	1.066
	20	1.723	1.137	2.859	1.938	0.944
	25	1.606	1.087	2.839	1.735	0.866
	50	1.365	0.986	2.800	1.259	0.695

Table 9. Minimal detectable differences between mean difference and ratio scores before and during supplementation. Analyses were conducted on natural-log transformed productivity data.

Response variable	Treatment years	Minimal detectable differences by reference population				
		Naches	Entiat	Marsh	Sesech	Little Wenatchee
T-R	5	0.540	0.551	0.674	0.890	0.585
	10	0.367	0.452	0.542	0.590	0.413
	15	0.308	0.421	0.499	0.486	0.355
	20	0.277	0.405	0.477	0.430	0.324
	25	0.257	0.396	0.465	0.393	0.305
	50	0.215	0.378	0.440	0.314	0.265
T/R	5	0.915	1.286	0.743	0.697	1.685
	10	0.744	0.973	0.704	0.541	1.227
	15	0.691	0.868	0.692	0.489	1.072
	20	0.666	0.815	0.687	0.463	0.993
	25	0.652	0.783	0.683	0.447	0.943
	50	0.628	0.719	0.677	0.416	0.843
ΔT-ΔR	5	0.885	0.810	1.028	1.252	0.971
	10	0.631	0.609	0.822	0.809	0.640
	15	0.546	0.542	0.755	0.655	0.525
	20	0.502	0.508	0.722	0.570	0.463
	25	0.475	0.487	0.702	0.516	0.423
	50	0.423	0.446	0.664	0.391	0.333

Step 3: Quantitative Method for Ranking Selection Criteria

Not surprisingly, different selection criteria produced different results (Table 10). Determining whether a given population is or is not a suitable reference population based on selection criteria such as graphic analysis can be subjective. In addition, treating each selection criterion as equally important may not be appropriate. For example, using the information in Table 10, is it appropriate to select a reference population that has two or three "Yes" entries, or should only populations with four "Yes" entries be selected as suitable reference populations? This approach does not allow certain selection criteria to carry more weight in the overall selection process. That is, correlation may be more important than graphic analysis in the overall selection process. In order to reduce subjectivity, we developed a method of scoring and weighting each selection criterion. This method allows a more quantitative process for selecting suitable reference populations.

Table 10. Summary of results from graphic analysis, correlations, trend analysis, and power analysis (minimal detectable differences). "Yes" indicates that the population is a suitable reference population for the Chiwawa population; "No" indicates that it may not be a suitable reference population.

Potential reference populations	Graphic analysis	Correlation	Trends	Minimal detectable differences
Spawner Abundance				
Naches	Yes	Yes	Yes	Yes
Entiat	Yes	Yes	Yes	Yes
Marsh	No	No	Yes	No
Sesech	No	No	Yes	Yes
Little Wenatchee	Yes	Yes	Yes	Yes
Natural-Origin Recruits				
Naches	Yes	Yes	Yes	No
Entiat	No	Yes	No	Yes
Marsh	Yes	Yes	Yes	Yes
Sesech	No	Yes	No	No
Little Wenatchee	Yes	Yes	Yes	Yes
Productivity				
Naches	Yes	Yes	Yes	Yes
Entiat	No	No	No	Yes
Marsh	No	Yes	Yes	No
Sesech	Yes	Yes	Yes	Yes
Little Wenatchee	Yes	Yes	Yes	Yes

We developed scoring methods for each of the following five selection criteria:

(1) The proportion of natural-origin spawners (pNOS) in the reference population for the period before supplementation (pre-pNOS);

(2) pNOS in the reference population for the period following supplementation (post-pNOS);

(3) The correlation between the reference and supplemented populations before supplementation;

(4) The relative difference in slopes between the reference and supplemented populations before supplementation; and

(5) The coefficient of variation (CV) of the ratio of supplemented to reference populations before the period of supplementation.

Each selection criteria was scored from 0 to 1, with 0 being the worst possible score and 1 being the best.

The pre- and post-pNOS values were calculated as the average pNOS values before and after supplementation, respectively. Because pNOS values range from 0-1, we did not need to rescale these values. When using reference populations to evaluate the effects of supplementation programs, it is important that the reference populations maintain high values of pNOS throughout the life of the monitoring program. Therefore, we heavily weighted the mean pNOS scores. We assigned weights of 30 and 40 to the mean pre- and post-pNOS scores, respectively. The relatively larger weight for the post-supplementation period is to reduce the likelihood of retaining a reference population that becomes influenced by hatchery fish during the supplementation period.

We assessed the association between the reference and supplemented populations during the pre-supplementation period by calculating the Pearson correlation coefficient, which ranges from -1 to 1. To scale the coefficient between 0 and 1, we took the absolute value of the coefficient. Thus, a coefficient of -0.92 would be reported as 0.92. For our analyses, we were not concerned with the direction of the relationship, only the strength of the relationship. The correlation coefficient was given a weight of 12.5.

As noted earlier, we used least squares to fit a linear trend to each of the reference populations and the supplemented population during the pre-supplementation period. Using the slope estimates for each trend line, we calculated the relative difference in slopes as the slope of the supplemented population minus the slope of the reference population, divided by the slope of the reference population. To scale this value between 0 and 1, we used absolute values, and depending on the direction of the slopes, we subtracted the relative difference from 1. The latter was needed to make sure a larger relative difference value indicated a small difference in slopes between the supplemented and reference populations. The relative difference score was given a weight of 7.5.

Finally, as a means to score effect size, we calculated the CV of the ratio of supplemented to reference population parameters (i.e., T/R). The CV was calculated as the standard deviation of the ratios divided by the absolute value of the mean ratios. The CV was subtracted from 1. This scaled the value from 0 to 1 with larger values representing the best condition. The CV was given a weight of 10, which is greater than the weight for trend, but less than the weight for correlation.

The total score for a reference population was calculated by multiplying the estimated value, which ranged from 0 to 1, by its weight. The sum of the five weighted values provided a total score, which ranged from 0 to 100. Based on several simulations, we set the cut-off score at 81. That is, if the total score for a given reference population equaled or exceeded 81, the population was included as a suitable reference population. If the total score fell below 81, the population was not considered a suitable reference. Based on the distribution of all scores possible, a score of 81 or greater represented only 3% of the total distribution. Thus, a cut-off of 81 is quite conservative.

Under Step 3, we used this method to select the final suite of suitable reference populations. Table 11 shows results from scoring each of the reference populations using the quantitative method. Using the cut-off criterion of 81, only the Naches, Marsh, and Sesech populations would be considered suitable reference populations for the Chiwawa supplementation program. Both the Entiat and Little Wenatchee populations failed to meet the minimum score, largely because of the influence of hatchery fish within those populations (i.e., relatively low pNOS values).

Table 11. Results from scoring potential reference populations using the selection criteria (pNOS, correlation, trend, and effect size). Populations with scores less than 81 were considered unsuitable as reference populations. Populations with scores equal to or greater than 81 were considered suitable references. These results were based on natural-log transformed data.

Potential reference populations	**Population metric**		
	Abundance	**NORs**	**Productivity**
Naches	85	88	91
Entiat	23	21	16
Marsh	79	91	87
Sesech	84	85	88
Little Wenatchee	51	53	49

An important benefit from scoring the different selection criteria is that the total scores can be used to weight the outcome of differing statistical results. For example, analyses may show that when three suitable reference populations are compared to the supplemented population, two of the reference populations may indicate a significant treatment effect, while the third indicates no effect. Under this scenario it is not clear if the supplementation program has or has not affected the abundance or productivity of the supplemented population. If, however, the two reference populations that produced a significant result had higher total scores than the reference population that did not indicate a significant result, one can place more weight on the results from populations with higher total scores.

Conclusions

The purpose of this exercise was to develop a method for selecting suitable reference populations that could be used to assess the effects of supplementation programs on spawner abundance, NORs, and productivity. The selection process included a three-step process (Figure 8). Step 1 identified populations with similar life-history characteristics, few or no hatchery spawners, a long time series of accurate abundance and productivity estimates, and similar freshwater habitat impairments and out-of-basin effects. Populations that met these criteria were then examined for their graphical and statistical relationship with the supplemented population (Step 2). The statistical analysis under Step 2 were converted to a quantitative model (Step 3) that was used to generate a weighted score for pNOS, correlation, trends, and effect sizes for each potential reference population. Reference populations with total scores of 81 or greater were selected as suitable reference populations.

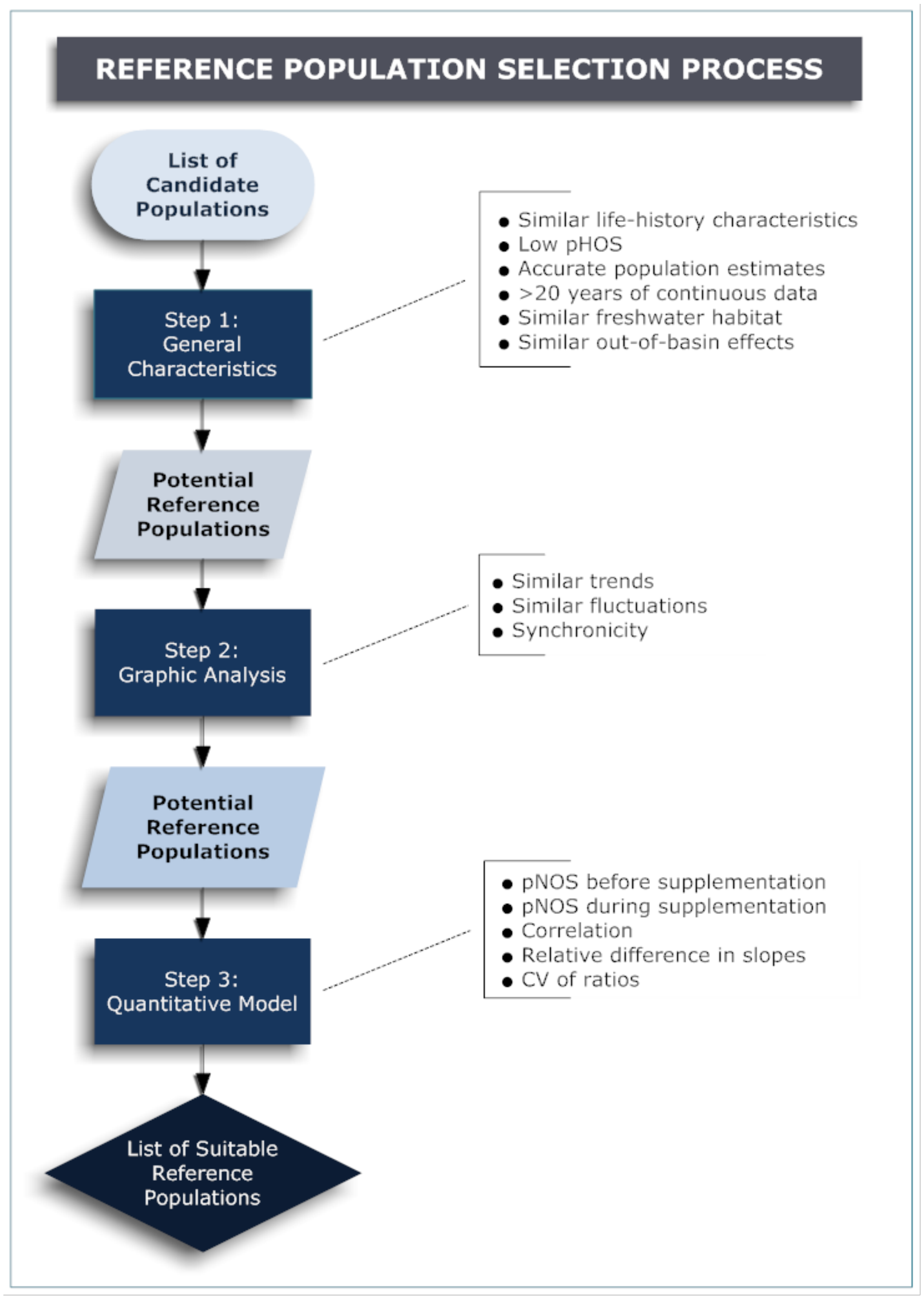


Figure 8. Three-step process for selecting suitable reference populations for supplemented populations.

We used this approach to select suitable reference populations for analyzing the effects of the Chiwawa spring Chinook supplementation program on fish abundance and productivity. The method indicated that the Naches, Marsh, and Sesech populations would serve as suitable reference populations for the Chiwawa spring Chinook supplementation program. Both the Entiat and Little Wenatchee populations failed to meet the minimum score, largely because of the influence of hatchery fish within those populations (i.e., relatively low pNOS values). However, because the presence of hatchery spring Chinook within those populations should decrease, they may serve as unique reference populations in which the comparisons change from all populations receiving hatchery fish to only the Chiwawa population receiving hatchery fish. Therefore, we will continue to include both the Little Wenatchee and Entiat populations in future analyses.

An important assumption in the use of reference populations is that the supplemented and reference populations that tracked each other before supplementation would continue to track each other in the absence of supplementation. Given that the reference populations did not match the Chiwawa population on all criteria examined (Table 1) and some reference populations tracked the Chiwawa population more poorly than others (Figures 2-7; Tables 2-4), there may be some uncertainty as to whether differences observed between the Chiwawa and reference populations during the supplementation period are associated with the hatchery program, "nuisance" factors[18], or a combination of both. In addition, we have no ability to regulate or control activities in reference areas. Any large-scale change (man-made or natural) in reference areas could affect our ability to assess the effectiveness of the supplementation program.

Because we have no ability to maintain reference areas for long periods of time and may not be able to control all activities even within the supplemented populations, we propose the use of a "causal-comparative" approach to strengthen the certainty of our inferences (Pearsons and Temple 2010). The causal-comparative approach relies on correlative data to try and make a case for causal inference.[19] Correlation is used to rule out alternative hypotheses (note that we make our case as much if not more by disproving plausible alternatives as we do by showing that the data are consistent with a hypothesis). For example, large scale land-use activities or natural events can affect stream flows, fine sediment recruitment, and water temperatures. Changes in these factors can affect the freshwater survival and productivity of fish independently of supplementation programs. If changes in habitat, migratory, and ocean conditions do not affect reference and treatment populations similarly, inferences associated with supplementation programs may be confounded. By measuring and tracking these extraneous factors within reference and treatment areas, we can assess the effects of these state variables on population conditions independent of the supplementation programs. This allows us to more effectively assess the influence of supplementation programs on populations.

To that end, we recommend that the following state variables be measured and tracked within the Chiwawa Basin and each of the reference areas: mean annual precipitation, total and riparian forest cover, road density, impervious surface, and alluvium. These variables can be used to describe differences in water temperatures at different life stages (pre-spawning, egg incubation, and summer rearing) and substrate characteristics, including fine sediments and embeddedness (Jorgensen et al. 2009). They can be used to assess possible changes in spawner abundance, NORs, and productivity that are independent of supplementation.

---

[18] A "nuisance" factor is any factor that is outside the control of the experimenter and can affect the response variable (spawner abundance or productivity). In this case, nuisance factors may include differences in freshwater habitat trends and conditions, out-of-basin effects (e.g., migration and ocean survival), and hatchery strays that affect the Chiwawa and reference populations differently.

[19] It is important to point out that correlation does not demonstrate cause-and-effect. It only suggests a relationship between variables. Thus, inferences based on correlation lack the certainty that is associated with a design-based approach.

Analyses with Reference Populations

Once suitable reference populations are selected, methods for analyzing the supplemented and reference populations need to be identified. What follows is a description of different analyses that can be used to assess the effects of supplementation programs on spawner abundance, NORs, and productivity using reference populations. Later in this report we describe methods for assessing supplementation effects when reference populations are not available.

We used some of the reference populations selected for the Chiwawa program to illustrate the different methods for evaluating the effects of the supplementation program on spawner abundance, NORs, and productivity. For abundance, we selected the Naches, Entiat, Little Wenatchee, and Sesech populations as suitable references for the Chiwawa population. For NORs, we selected the Naches, Entiat, Marsh, and Little Wenatchee populations as suitable references. For productivity, we selected the Naches, Sesech, Little Wenatchee, and Marsh Creek as suitable references for the Chiwawa. As noted earlier, we included the Little Wenatchee and Entiat populations, even though they did not meet all the criteria for suitable reference populations.

Analysis of Trends

As a first step, we used trend analyses to assess the effects of the Chiwawa supplementation program on spring Chinook spawner abundance, NORs, and productivity. Here, we compared the slopes of the trends between each treatment/reference pair before and during supplementation using t-tests. If the hatchery program is successfully supplementing the natural spring Chinook population, trends in spawner abundance and NORs should deviate significantly (i.e., the slope of the supplemented population should be greater than the slopes of the reference populations during the supplementation period). For productivity, the slope of the supplemented population, relative to the reference population, should increase or remain the same.

Trend analysis indicated that the relationship of slopes of spawner abundance between the Chiwawa and reference populations did not change significantly after the initiation of supplementation (Figure 9; Table 12). This was true for both transformed and untransformed abundance data. Before supplementation, spawner abundances trended down in both the Chiwawa and reference populations (Figure 9). During the period of supplementation, abundances in both the Chiwawa and reference populations trended upward. Interestingly, in nearly all treatment/reference comparisons, the Pearson correlation coefficient was greater in the supplementation period than in the pre-supplementation period (Table 12). This was most evident in the transformed abundance data (Figure 9).

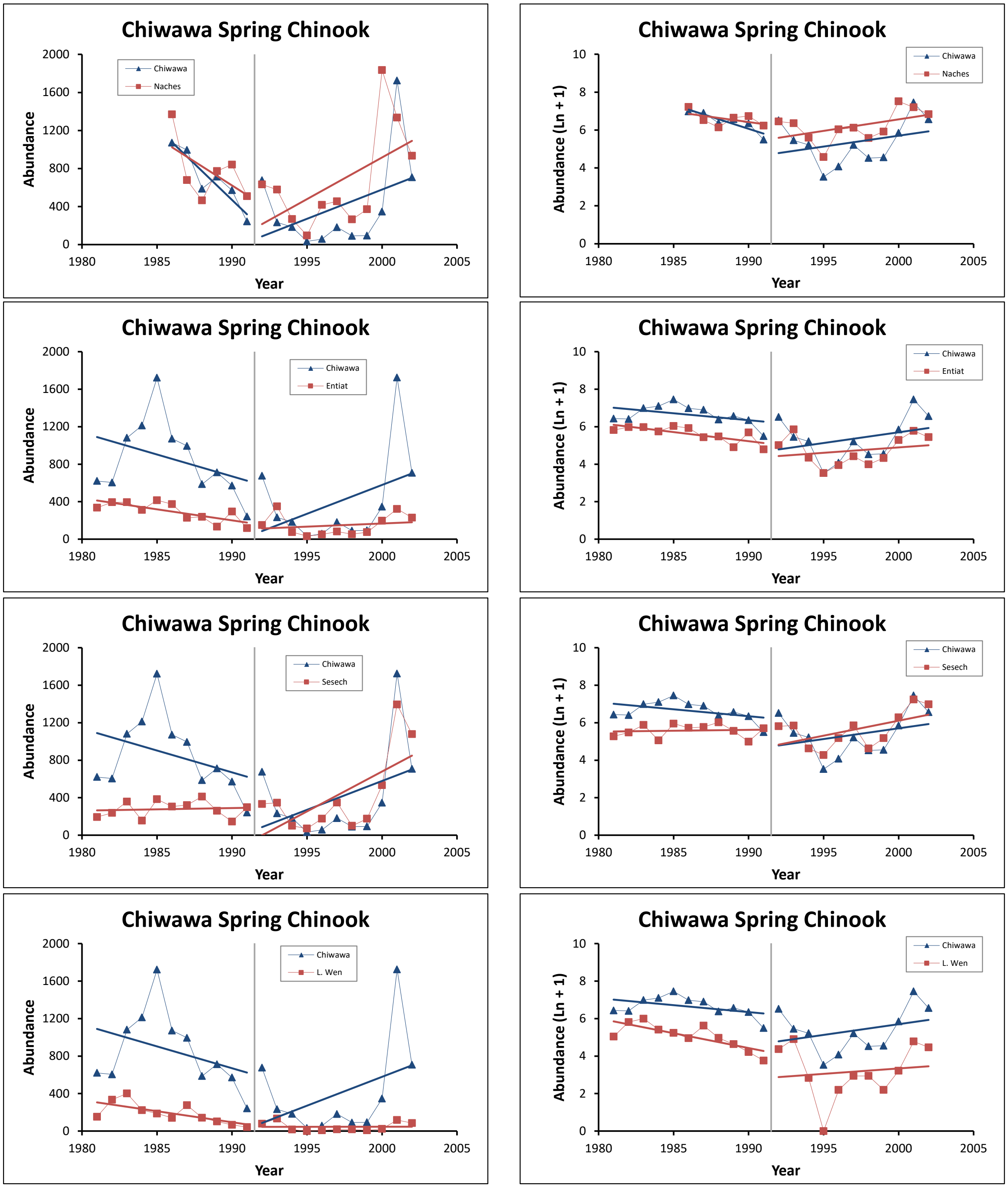


Figure 9. Trends in spring Chinook spawner abundance in the Chiwawa and reference populations. The vertical lines in the figures separate the pre- and post-supplementation periods. Figures on the left include untransformed spawner abundance data; those on the right include natural-log transformed data.

Table 12. Pearson correlation coefficients and t-test results comparing slopes of spawner abundance trends between reference populations and the Chiwawa spring Chinook population before and during the supplementation periods; for correlation coefficients, an asterisk (*) indicates significance at P < 0.05. Analyses include both untransformed and natural-log transformed spawner abundance data.

Reference population	Pearson correlation coefficient		Test on slopes			
			t-value		P-value	
	Before	During	Before	During	Before	During
Spawner Abundance						
Naches	0.684*	0.595	-0.659	-0.414	0.528	0.684
Entiat	0.598*	0.672*	-0.596	1.162	0.559	0.260
Sesech	0.274	0.904*	-1.265	-0.418	0.222	0.681
Little Wenatchee	0.399	0.685*	-0.591	1.330	0.562	0.200
LN Spawner Abundance						
Naches	0.642*	0.813*	-1.323	-0.047	0.222	0.963
Entiat	0.652*	0.860*	0.412	0.422	0.685	0.678
Sesech	0.149	0.878*	-1.431	-0.333	0.170	0.743
Little Wenatchee	0.670*	0.861*	1.325	0.316	0.202	0.756

Trend analysis indicated that the relationship of slopes of NORs between the Chiwawa and reference populations did not change significantly after the initiation of supplementation (Figure 10; Table 13). Before supplementation, Chiwawa NORs trended downward more strongly than the reference populations (Figure 10). However, during the supplementation period, both the Chiwawa and reference population NORs trended upward in parallel. In nearly all treatment/reference comparisons, the Pearson correlation coefficient was greater in the pre-supplementation period than in the supplementation period (Table 13).

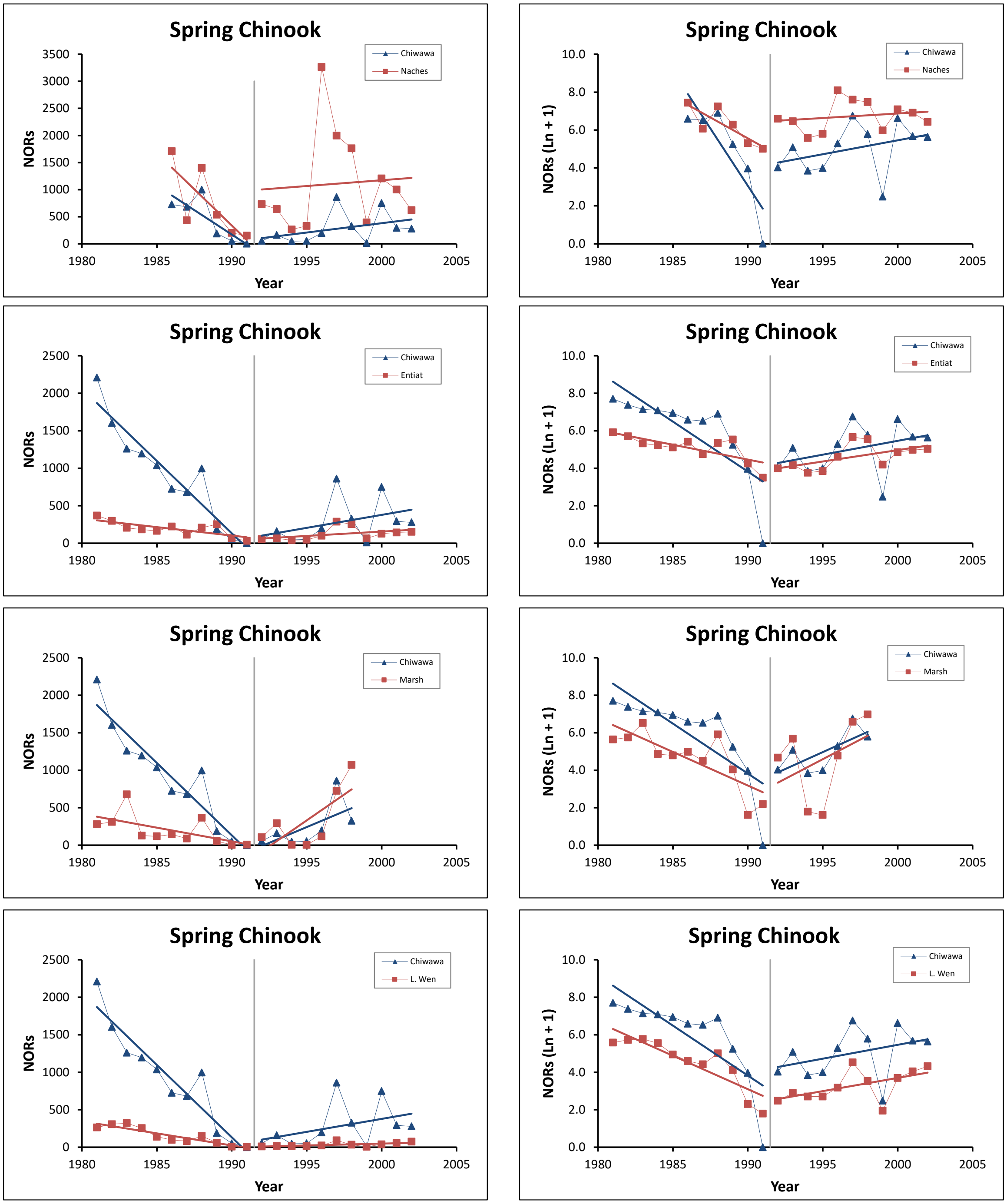


Figure 10. Trends in spring Chinook natural-origin recruits (NORs) in the Chiwawa and reference populations. The vertical lines in the figures separate the pre- and post-supplementation periods. Figures on the left include untransformed NORs; those on the right include natural-log transformed data.

Table 13. Pearson correlation coefficients and t-test results comparing slopes of natural-origin recruits trends between reference populations and the Chiwawa spring Chinook population before and during the supplementation periods; for correlation coefficients, an asterisk (*) indicates significance at $P < 0.05$. Analyses include both untransformed and natural-log transformed natural-origin recruits.

Reference population	Pearson correlation coefficient		Test on slopes			
			t-value		P-value	
	Before	During	Before	During	Before	During
Natural-Origin Recruits						
Naches	0.803*	0.432	0.666	0.140	0.524	0.890
Entiat	0.795*	0.754*	-7.495	0.847	0.000	0.408
Marsh	0.605*	0.677*	-5.786	-0.718	0.000	0.489
Little Wenatchee	0.880*	0.758*	-7.206	1.128	0.000	0.274
LN Natural-Origin Recruits						
Naches	0.824*	0.710*	-1.985	0.693	0.082	0.497
Entiat	0.886*	0.796*	-2.563	0.202	0.019	0.842
Marsh	0.830*	0.835*	-1.038	-0.134	0.313	0.896
Little Wenatchee	0.927*	0.898*	-1.150	0.046	0.265	0.964

As with NORs and spawner abundance data, trend analysis indicated that the relationship of slopes of productivity (recruits/spawner) between the Chiwawa and reference populations did not change significantly after the initiation of supplementation (Figure 11; Table 14). This was true for both transformed and untransformed productivity data. Before supplementation, productivities trended down in both the Chiwawa and reference populations (Figure 11). During the period of supplementation, productivities fluctuated widely in both the Chiwawa and reference populations. Nevertheless, during the supplementation period, productivities generally increased in both the reference and Chiwawa populations. Unlike with spawner abundance, the Pearson correlation coefficients resulting from analysis of productivity data were generally higher in the pre-supplementation period than during the supplementation period (Table 14).

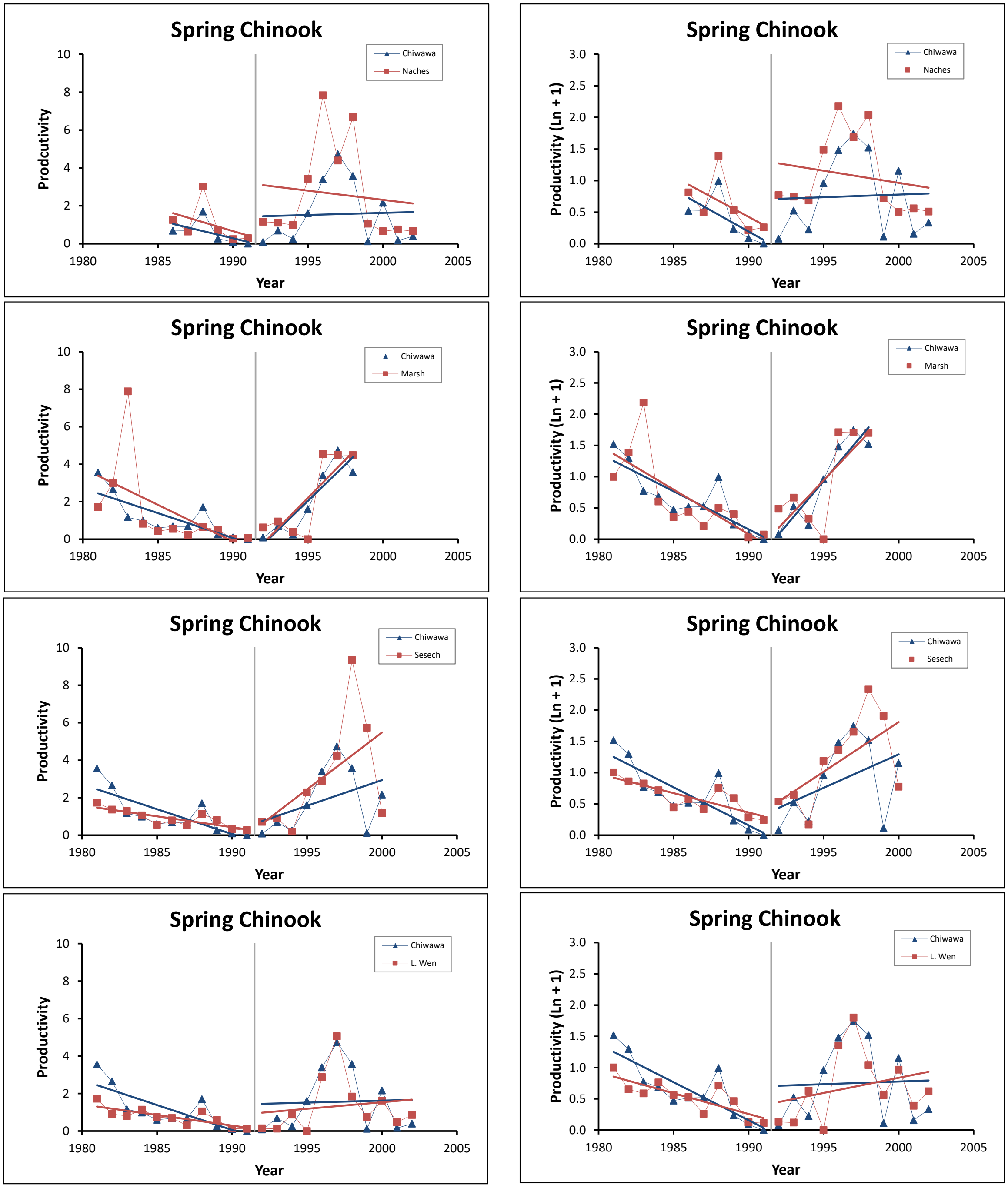


Figure 11. Trends in spring Chinook productivity (recruits/spawner) in the Chiwawa (supplemented) and reference populations. The vertical lines in the figures separate the pre- and post-supplementation periods. Figures on the left include untransformed productivity data; those on the right include natural-log transformed data.

Table 14. Pearson correlation coefficients and t-test results comparing slopes of productivity (recruits/spawner) trends between reference populations and the Chiwawa spring Chinook population before and during the supplementation periods; for correlation coefficients, an asterisk (*) indicates significance at P < 0.05. Analyses include both untransformed and natural-log transformed productivity data.

Reference population	Pearson correlation coefficient		Test on slopes			
			t-value		P-value	
	Before	During	Before	During	Before	During
Productivity						
Naches	0.960*	0.802*	0.169	0.387	0.870	0.703
Marsh	0.320	0.910*	0.605	-0.132	0.553	0.898
Sesech	0.903*	0.491	-2.059	-0.837	0.054	0.417
Little Wenatchee	0.848*	0.864*	-2.065	-0.213	0.054	0.834
LN Productivity						
Naches	0.944*	0.805*	-0.042	0.526	0.968	0.605
Marsh	0.610*	0.804*	0.428	0.281	0.674	0.784
Sesech	0.913*	0.531	-2.050	-0.463	0.055	0.651
Little Wenatchee	0.862*	0.751*	-1.811	-0.480	0.087	0.637

Using trend analysis, we found no evidence that the supplementation program has significantly increased the spawner abundance and NORs of spring Chinook in the Chiwawa Basin. Even though we documented an increasing trend in spawner abundance and NORs during the supplementation period, a similar increase in spawner abundance and NORs was observed in the reference populations. In addition, we found no evidence that the supplementation program has increased the productivity of spring Chinook in the Chiwawa Basin. Importantly, the productivity of spring Chinook in the Chiwawa Basin did not trend downward during the supplementation period. Thus, based on trend analysis, it appears that the supplementation program has not increased or decreased the abundance and productivity of spring Chinook in the Chiwawa Basin.

We note that this exercise only tests the slopes of the trend lines. It does not test for differences in elevations of the trend lines. A supplementation program could increase spawner abundance, NORs, and productivity of the target population without changing the slopes of the trend lines. That is, supplementation could cause the elevation of the trend line to be greater during the supplementation period than during the pre-supplementation period. In the next section we evaluate elevation differences by testing mean differences before and after supplementation.

Analysis of Mean Differences, Ratios, and Rates

For assessing mean differences between supplemented and reference populations, we derived three different response variables using transformed and untransformed spawner abundance, NORs, and productivity data. The first included difference scores, which were calculated as the difference between paired treatment and reference data (T-R). The second included ratios, which were calculated as the ratio of paired treatment and reference data (T/R). Finally, we calculated the differences in annual changes in paired treatment and reference population data (ΔT-ΔR; see footnote #2).

If the hatchery program is successfully supplementing the natural spring Chinook population, the mean difference or ratio score of paired spawner abundance data and NORs during the supplementation period should be greater than the pre-supplementation period. For productivity, the mean difference or ratio score during the supplementation period should be equal to or higher than the pre-supplementation period. We tested the following statistical hypotheses.

Spawner Abundance and NORs:

> Ho: Mean Difference (or Ratio) before supplementation ≥ Mean Difference (or Ratio) during supplementation.
>
> Ha: Mean Difference (or Ratio) before supplementation < Mean Difference (or Ratio) during supplementation (i.e., $\mu_{pre} - \mu_{post} < 0$).

Productivity (Recruits/Spawner):

> Ho: Mean Difference (or Ratio) before supplementation ≤ Mean Difference (or Ratio) during supplementation.
>
> Ha: Mean Difference (or Ratio) before supplementation > Mean Difference (or Ratio) during supplementation (i.e., $\mu_{pre} - \mu_{post} > 0$).[20]

For each set of response variables, we tested before/after supplementation effects using a one-tailed Aspin-Welch unequal-variance test. We used the Aspin-Welch unequal-variance test instead of Student's t-test, because in nearly every case, the variances of response variables in the pre-treatment and supplementation periods were unequal.[21] This was true even for natural-log transformed variables. We used the modified Levene equal-variance test to assess the equality of variance. In some cases, the distributions of response variables were not normal (based on the Omnibus Normality test and examination of histograms, normal probability plots, and box plots). Therefore, we also used a randomization test, based on 10,000 Monte Carlo simulations, to assess differences in response variables before and during supplementation. The randomization procedure only allowed the testing of two-tailed hypotheses. Therefore, we generated 95% confidence intervals on the mean difference ($\mu_{pre} - \mu_{post}$) using bootstrapping methods to determine the direction of the difference. We generated 5,000 bootstrap samples to calculate confidence intervals.

All these statistical methods assume that the samples of derived difference or ratio scores from the pre-supplementation and supplementation periods were independent. However, BACI designs, like time-series trend analysis, are repeated-measures designs and therefore are susceptible to temporal correlations in the data. This means that the two samples of difference or ratio scores may not be independent. Under this scenario, ARIMA models can be used to describe the correlation structure in temporal data (Gotelli and Ellison 2004). ARIMA models can be fit individually to the reference and supplemented time series data, or to a derived data series created by taking the ratio or difference of the supplemented/reference data at each time step. ARIMA models, however, require a long time series (N > 40) and therefore we could not use them to model

---

[20] Because of the logic of null hypothesis testing, the rejection of the null hypothesis of no difference in productivity would mean that the supplementation program has reduced the productivity of the target population (here rejection of the null indicates "harm"). Notice that the rejection of the null hypothesis of no difference in spawner abundance means that the supplementation program has improved the spawner abundance in the target population (here rejection of the null indicates "benefit").

[21] In cases in which the variances were equal, both the Aspin-Welch test and Student's t-test gave the same result.

the spring Chinook data. Thus, we acknowledge that our analyses may be confounded if the samples are not independent.

Difference Scores (T-R)

Analysis of supplementation effects on spawner abundance using difference scores indicated that supplementation did not significantly increase spawning abundance in the Chiwawa Basin (Table 15; Figure 12). Only the Little Wenatchee-Chiwawa pairing using transformed abundance data indicated a significant increase in spawning abundance following supplementation. The randomization test indicated significant differences in several of the treatment-reference pairs; however, the bootstrap CIs indicated that those differences were in the wrong direction (i.e., CIs > 0). That is, compared to the reference populations, spawner abundance decreased in the Chiwawa Basin during the supplementation period (Figure 12).

Table 15. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed spawner abundance data. Tests determined if the mean difference scores during the supplementation period were greater than mean difference scores during the pre-supplementation period.

Reference population	**Aspin-Welch unequal-variance test**			**Randomization test P-value**	**Bootstrap 95% CI**
	t-value	**P-value**	**Effect size**		
Spawner Abundance					
Naches	1.066	0.848	184	0.322	-162 – 472
Entiat	1.872	0.962	316	0.078	17 – 633
Sesech	4.502	0.999	607	0.000	349 – 851
Little Wenatchee	1.773	0.954	321	0.093	0 – 690
LN Spawner Abundance					
Naches	2.603	0.990	0.701	0.026	0.210 – 1.214
Entiat	1.701	0.946	0.388	0.108	-0.033 – 0.811
Sesech	5.394	0.999	1.327	0.000	0.891 – 1.805
Little Wenatchee	-2.259	0.018	0.609	0.034	-1.125 – -0.097

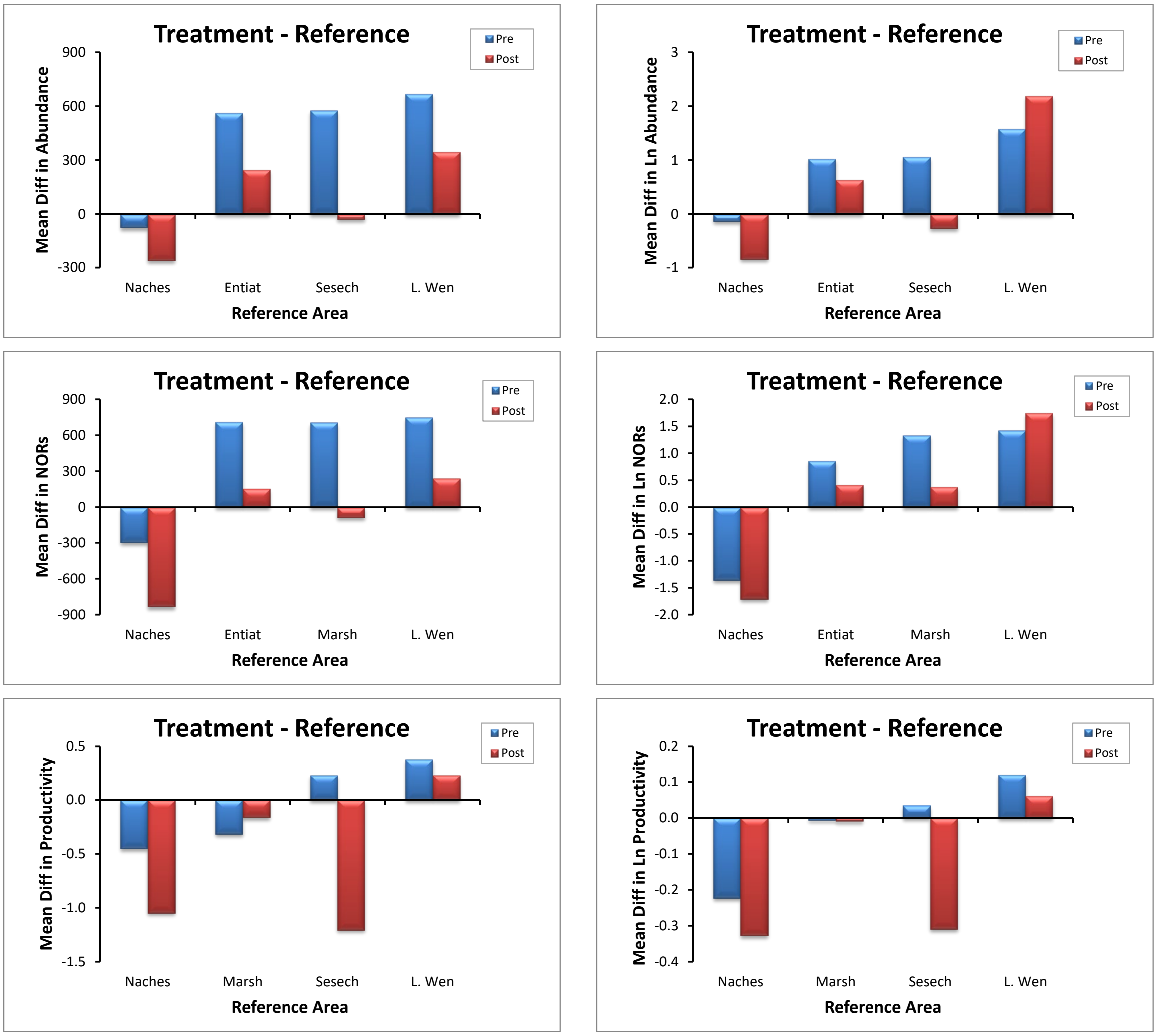


Figure 12. Mean difference (Treatment – Reference) scores of untransformed (figures on the left) and transformed (figures on the right) spawner abundance, natural-origin recruits (NORs), and productivity data before (pre) and after (post) spring Chinook supplementation in the Chiwawa Basin. Positive effects of supplementation on spawner abundance and NORs are indicated when the post-supplementation (red) bars are greater than their corresponding pre-supplementation (blue) bars. Negative effects of supplementation on productivity are indicated when the pre-supplementation (blue) bars are greater than their corresponding post-supplementation (red) bars.

Analysis of supplementation effects on NORs using difference scores indicated that supplementation did not significantly increase NORs in the Chiwawa Basin (Table 16; Figure 12). The randomization test indicated significant differences in several of the treatment-reference pairs; however, the bootstrap CIs indicated that those differences were in the wrong direction. That is, compared to the reference populations, NORs decreased in the Chiwawa Basin during the supplementation period (Figure 12).

Table 16. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed natural-origin recruits. Tests determined if the mean difference scores during the supplementation period were greater than mean difference scores during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Natural-Origin Recruits					
Naches	1.787	0.953	537	0.081	-60 – 1039
Entiat	2.879	0.993	558	0.007	201 – 916
Marsh	3.817	0.999	795	0.001	381 – 1153
Little Wenatchee	2.668	0.991	510	0.013	145 – 863
LN Natural-Origin Recruits					
Naches	0.430	0.659	0.354	0.686	-0.948 – 1.975
Entiat	0.788	0.779	0.445	0.465	-0.504 – 1.583
Marsh	1.45	0.916	0.953	0.168	-0.169 – 2.243
Little Wenatchee	-0.813	0.214	-0.319	0.506	-0.948 – 0.484

Analysis of supplementation effects on productivity (adult recruits/spawner) using difference scores indicated that supplementation did not significantly decrease productivity in the Chiwawa Basin (Table 17; Figure 12). All tests, regardless of treatment-reference pairs, indicated that productivity did not change significantly during the supplementation period. These tests indicate that supplementation has not negatively affected the productivity of spring Chinook salmon in the Chiwawa Basin.

Table 17. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed productivity data. Tests determined if the mean difference scores during the supplementation period were less than mean difference scores during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Productivity					
Naches	1.134	0.139	0.594	0.296	-0.427 – 1.540
Marsh	-0.203	0.579	0.152	0.932	-0.304 – 1.381
Sesech	1.607	0.071	1.435	0.151	-0.403 – 2.917
Little Wenatchee	0.431	0.335	0.147	0.665	-0.498 – 0.762
LN Productivity					
Naches	0.770	0.227	0.104	0.480	-0.125 – 0.378
Marsh	0.012	0.495	0.003	0.992	-0.375 – 0.493
Sesech	1.463	0.087	0.343	0.161	-0.135 – 0.732
Little Wenatchee	0.390	0.351	0.060	0.701	-0.229 – 0.347

Ratio Scores (T/R)

As with difference scores, analysis of supplementation effects on spawner abundance using ratios indicated that supplementation did not significantly increase spawning abundance in the Chiwawa Basin (Table 18; Figure 13). Only the Little Wenatchee-Chiwawa pairing indicated a significant increase in spawning abundance following supplementation. Analysis with both transformed and untransformed Little Wenatchee-Chiwawa data indicated a significant effect. In contrast, only difference scores derived from transformed data indicated a significant effect. The randomization test indicated significant differences in several of the treatment-reference pairs; however, the bootstrap CIs indicated that those differences were in the wrong direction. That is, compared to the reference populations, spawner abundance decreased in the Chiwawa Basin during the supplementation period (Figure 13).

Table 18. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed spawner abundance data. Tests determined if the mean ratios during the supplementation period were greater than mean ratios during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Spawner Abundance					
Naches	2.110	0.970	0.398	0.065	0.056 – 0.737
Entiat	1.254	0.888	0.731	0.223	-0.365 – 1.834
Sesech	4.251	0.999	2.428	0.000	1.278 – 3.435
Little Wenatchee	-2.649	0.009	3.897	0.018	-6.579 – -1.202
LN Spawner Abundance					
Naches	2.783	0.993	0.120	0.021	0.045 – 0.199
Entiat	1.273	0.890	0.055	0.220	-0.026 – 0.135
Sesech	5.143	0.999	0.244	0.000	0.160 – 0.335
Little Wenatchee	-3.462	0.002	0.327	0.003	-0.516 – -0.154

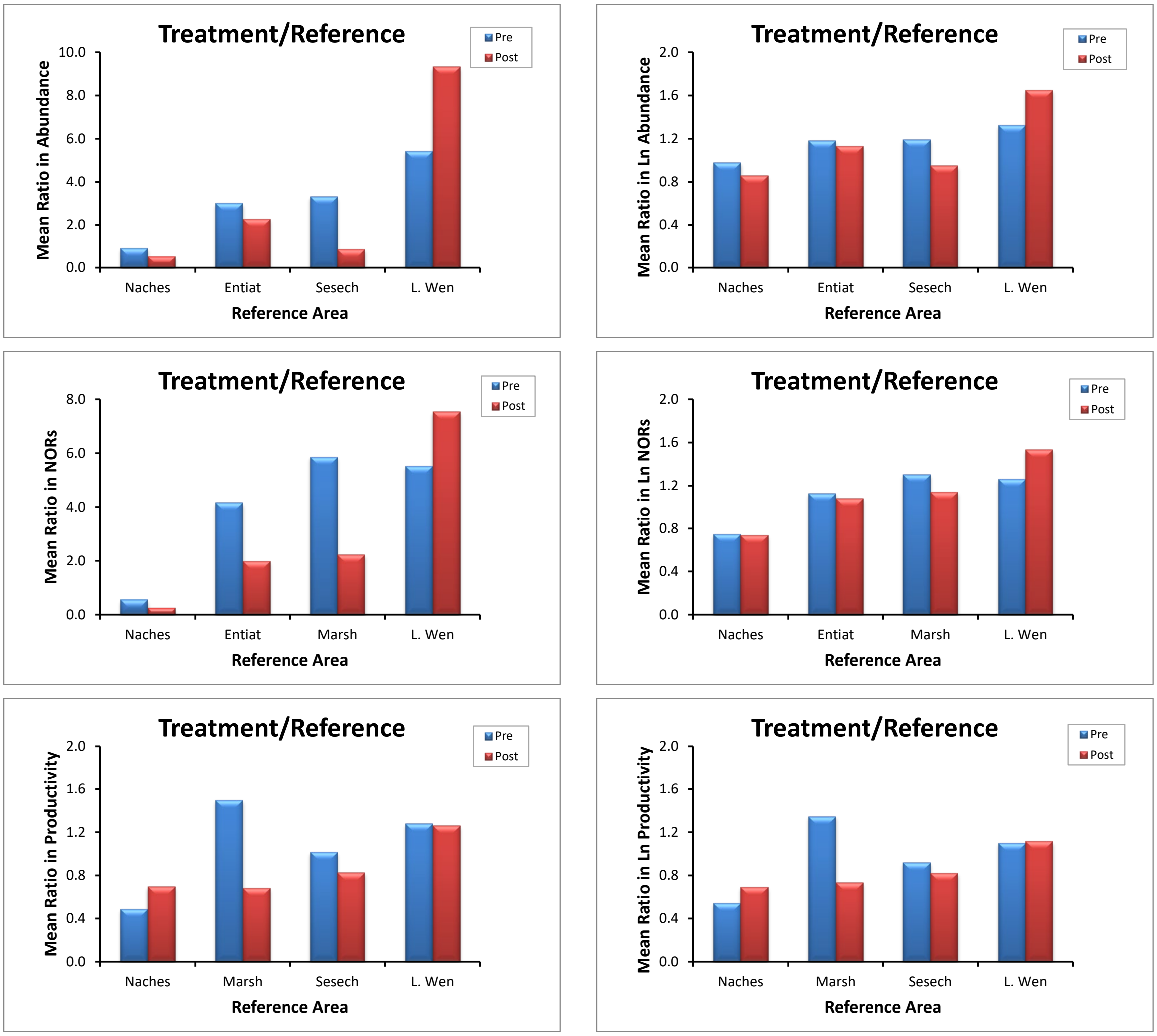


Figure 13. Mean ratios (Treatment/Reference) scores of untransformed (figures on the left) and transformed (figures on the right) spawner abundance, natural-origin recruits (NORs), and productivity data before (pre) and after (post) spring Chinook supplementation in the Chiwawa Basin. Positive effects of supplementation on spawner abundance and NORs are indicated when the post-supplementation (red) bars are greater than their corresponding pre-supplementation (blue) bars. Negative effects of supplementation on productivity are indicated when the pre-supplementation (blue) bars are greater than their corresponding post-supplementation (red) bars.

Analysis of supplementation effects on NORs using ratios indicated that supplementation did not significantly increase NORs in the Chiwawa Basin (Table 19; Figure 13). Only the Little Wenatchee-Chiwawa pairing indicated a significant increase in transformed NORs following supplementation. The randomization test indicated significant differences in several of the treatment-reference pairs; however, the bootstrap CIs indicated that those differences were in the wrong direction. That is, compared to the reference populations, NORs decreased in the Chiwawa Basin during the supplementation period (Figure 13).

Table 19. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed natural-origin recruits. Tests determined if the mean ratios during the supplementation period were greater than mean ratios during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Natural-Origin Recruits					
Naches	1.318	0.881	0.306	0.219	-0.157 – 0.670
Entiat	2.447	0.987	2.172	0.028	0.593 – 3.871
Marsh	2.001	0.965	3.638	0.075	0.532 – 7.201
Little Wenatchee	-1.148	0.136	2.020	0.284	-5.055 – 1.516
LN Natural-Origin Recruits					
Naches	0.057	0.522	0.009	0.967	-0.230 – 0.351
Entiat	0.359	0.638	0.049	0.759	-0.173 – 0.336
Marsh	0.603	0.721	0.161	0.579	-0.272 – 0.681
Little Wenatchee	-1.914	0.038	0.277	0.027	-0.504 – 0.031

Analysis of supplementation effects on productivity (adult recruits/spawner) using ratios indicated that supplementation did not significantly decrease productivity in the Chiwawa Basin (Table 20; Figure 13). Although the Aspin-Welch test indicated a significant effect when comparing the Chiwawa to the Marsh Creek population, both the randomization test and the bootstrap CI did not indicate a significant effect. These tests indicate that supplementation has probably not negatively affected the productivity of spring Chinook salmon in the Chiwawa Basin.

Table 20. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed productivity data. Tests determined if the mean ratios during the supplementation period were less than mean ratios during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Productivity					
Naches	-0.677	0.745	0.209	0.688	-0.700 – 0.425
Marsh	2.236	0.022	0.814	0.054	0.112 – 1.459
Sesech	0.677	0.253	0.191	0.515	-0.356 – 0.718
Little Wenatchee	0.033	0.487	0.018	0.979	-0.879 – 1.162
LN Productivity					
Naches	-0.639	0.734	0.148	0.616	-0.548 – 0.316
Marsh	1.952	0.036	0.613	0.081	-0.003 – 1.170
Sesech	0.447	0.330	0.098	0.663	-0.301 – 0.515
Little Wenatchee	-0.034	0.513	0.015	0.982	-0.692 – 0.861

Difference of Annual Difference Scores (ΔT-ΔR)

Analysis of supplementation effects on spawner abundance using difference scores of annual changes indicated that supplementation did not significantly increase spawning abundance in the Chiwawa Basin (Table 21; Figure 14). None of the statistical analyses detected a significant increase in annual change in the Chiwawa Basin relative to the reference populations.

Table 21. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed spawner abundance data. Tests determined if mean difference scores of annual change during the supplementation period were greater than mean difference scores of annual change during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Spawner Abundance					
Naches	0.009	0.503	2	0.995	-502 – 539
Entiat	-0.239	0.407	48	0.826	-414 – 327
Sesech	-0.126	0.451	20	0.902	-311 – 266
Little Wenatchee	-0.318	0.377	65	0.761	-452 – 311
LN Spawner Abundance					
Naches	-0.425	0.339	0.142	0.698	-0.744 – 0.466
Entiat	-0.084	0.467	0.028	0.933	-0.681 – 0.593
Sesech	-0.349	0.366	0.117	0.740	-0.741 – 0.515
Little Wenatchee	0.001	0.500	0.000	0.999	-0.663 – 0.687

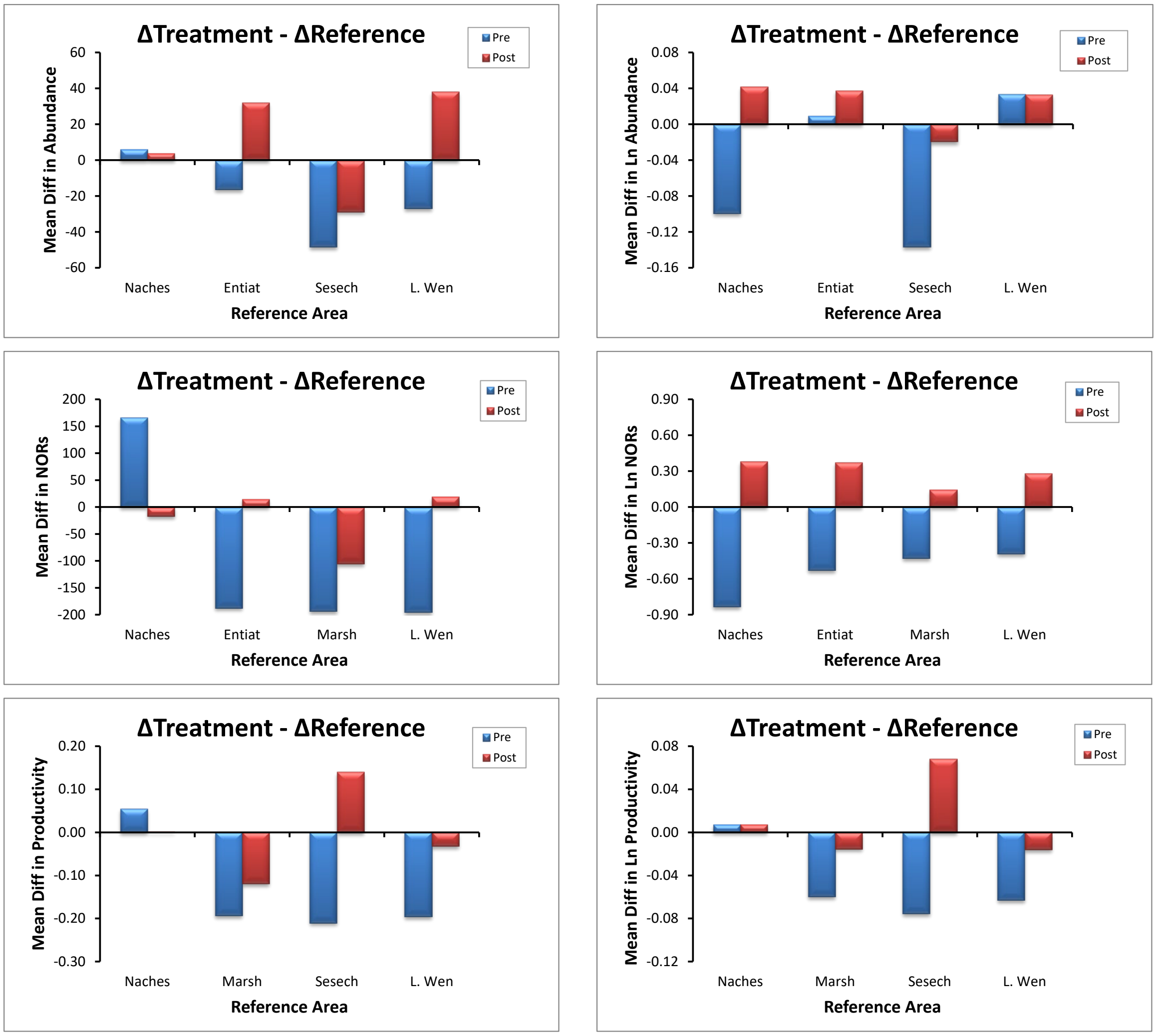


Figure 14. Mean difference scores of annual changes (ΔTreatment – ΔReference) of untransformed (figures on the left) and transformed (figures on the right) spawner abundance and productivity data before (pre) and after (post) spring Chinook supplementation in the Chiwawa Basin.

Analysis of supplementation effects on NORs using difference scores of annual changes indicated that supplementation did not significantly increase NORs in the Chiwawa Basin (Table 22; Figure 14). None of the statistical analyses detected a significant increase in annual change in the Chiwawa Basin relative to the reference populations.

Table 22. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed natural-origin recruits. Tests determined if mean difference scores of annual change during the supplementation period were greater than mean difference scores of annual change during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Natural-Origin Recruits					
Naches	0.399	0.652	184	0.741	-699 – 989
Entiat	-1.381	0.092	202	0.194	-471 – 86
Marsh	-0.505	0.311	88	0.624	-425 – 206
Little Wenatchee	-1.437	0.084	214	0.179	-481 – 64
LN Natural-Origin Recruits					
Naches	-1.301	0.118	1.214	0.224	-2.783 – 0.531
Entiat	-1.408	0.088	0.901	0.188	-1.977 – 0.387
Marsh	-0.712	0.244	0.570	0.517	-1.952 – 0.975
Little Wenatchee	-1.154	0.132	0.674	0.274	-1.706 – 0.497

Analysis of supplementation effects on productivity (adult recruits/spawner) using difference scores of annual changes indicated that supplementation did not significantly decrease productivity in the Chiwawa Basin (Table 23; Figure 14). All tests, regardless of treatment-reference pairs, indicated that productivity did not change significantly during the supplementation period.

Table 23. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed productivity data. Tests determined if the mean difference scores of annual change during the supplementation period were less than mean difference scores of annual change during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Productivity					
Naches	0.002	0.475	0.054	0.952	-1.464 – 1.583
Marsh	-0.063	0.525	0.074	0.948	-2.395 – 2.031
Sesech	-0.317	0.621	0.350	0.628	-2.387 – 1.695
Little Wenatchee	-0.347	0.633	0.163	0.728	-1.023 – 0.725
LN Productivity					
Naches	0.000	0.500	0.000	0.999	-0.408 – 0.445
Marsh	-0.126	0.549	0.044	0.904	-0.715 – 0.595
Sesech	-0.449	0.668	0.144	0.727	-0.685 – 0.509
Little Wenatchee	-0.200	0.578	0.047	0.842	-0.466 – 0.391

We believe results from analysis of mean differences of annual change (ΔT-ΔR) in spawning abundance, NORs, and productivity are difficult to interpret and may be insensitive to treatment effects. A simpler analysis, which is also easier to interpret, is the use of trend analysis. Therefore, we recommend that analyses using differences of annual change be replaced with trend analysis.

Corrections for Density Dependence and Carrying Capacity

The analyses described above assume that the density of spawners or recruits does not affect the survival and productivity of fish. However, it is well known that the density of fish can affect the number of recruits as well as the productivity of the population. This occurs through the relationship between density and mortality. Mortality of fish can be generally classified as density independent and density dependent. In general, when densities are low, the mortality is density independent, but as densities increase, the amount of density-dependent mortality increases. Monitoring programs can make use of this information to derive density-corrected estimates of productivity. In this section, we describe two different methods for deriving density-corrected estimates of productivity.

The first method controlled the effects of density on productivity (adult recruits/spawner; R/S) by partitioning observed productivities into density-independent and density-dependent productivity. When abundance is below the minimum number of spawners (S) needed to produce the maximum number of recruits (K_{sp}), the observed productivity is used in statistical tests. However, when the abundance is equal to or above K_{sp}, the modeled value of productivity (R/K_{sp}) is used in statistical tests.

$$\boldsymbol{Adj\ R/S = \begin{cases} R/S, & if\ S < K_{sp} \\ R/K_{sp}, & if\ S \geq K_{sp} \end{cases}}$$

The density-independent and density-dependent productivities were then combined in a single test.

The second method was based on one of the goals of supplementation, which is to fill the capacity of the environment with fish. This method corrects for differences in carrying capacities between the supplemented and reference populations. We did this by calculating the percent saturation of NORs. That is, we calculated the fraction of the habitat (τ) that was filled with NORs by dividing the observed NOR by the modeled maximum number of NORs (K_R) that the habitat could support.

$$\boldsymbol{\tau = \frac{NOR_{obs}}{K_R}}$$

Note that 1-τ represents the unused portion of the carrying capacity and is the term that is multiplied by the exponential growth equation to derive the logistic growth equation. We included τ in the statistical analyses.

These two methods require the estimation of carrying capacity (K_R) and the spawning abundance that produces the maximum number of recruits (K_{sp}). We estimated these parameters for both reference populations and the supplemented population using Ricker, Beverton-Holt, and smooth hockey stick stock-recruitment models. We used only spawner abundance as a predictor of subsequent brood recruitment. We made the following assumptions in proceeding with the analysis:

- Density-dependent mortality—For some time period before recruitment, the brood instantaneous mortality rate is proportional to the number of parent spawners (Ricker 1954).
- Lognormal variation—At any particular spawning stock size, the variation in recruitment is log-normally distributed about its average, and acts multiplicatively (Quinn and Deriso 1999).
- Measurement error—Error in spawning stock size estimates (measurement error) is small relative to the range of spawning stock sizes observed (Hilborn and Walters 1992). Variation in realized recruitment at any particular spawning stock size (process error) dominates recruitment measurement error.
- Stationarity—The average stock-recruitment relationship is constant over time (Hilborn and Walters 1992). That is, environmental conditions randomly affect survival independent of stock size or time.

In general, the methods we used to fit the models to the data followed those outlined in Hilborn and Walters (1992) and Froese (2008). The Ricker model, which assumes that the number of recruits increases to a maximum and then declines as the number of spawners increases, takes the form:

$$\boldsymbol{E(R) = \alpha S e^{-\beta S}}$$

where E(R) is the expected recruitment, S is spawner abundance, α is the number of recruits per spawner at low spawning levels, and β describes how quickly the recruits per spawner drop as the number of spawners increases. We estimated K_R as:

$$\boldsymbol{K_R = \left(\frac{\alpha}{\beta}\right) e^{-1}}$$

and K_{sp} as:

$$\boldsymbol{K_{sp} = \frac{1}{\beta}}$$

The Beverton-Holt model assumes that the number of recruits increases constantly toward an asymptote as the number of spawners increases. After the asymptote is reached, the number of recruits neither increases nor decreases. The asymptote represents the maximum number of recruits the system can support (i.e., carrying capacity for the system; K_R). The Beverton-Holt curve takes the form:

$$\boldsymbol{E(R) = \frac{(\alpha S)}{(\beta + S)}}$$

where E(R) and S are as above, α is the maximum number of recruits produced (K_R), and β is the number of spawners needed to produce (on average) recruits equal to one-half the maximum number of recruits. Because $K_{sp} = \infty$ in the Beverton-Holt model, we estimated K_{sp} as the number of spawners needed to produce 0.99(K_R).

Like the Beverton-Holt model, the smooth hockey stick model assumes that the number of recruits increases toward an asymptote (carrying capacity; K_R) as the number of spawners increases. After the carrying capacity is reached, the number of recruits neither increases nor decreases. The carrying capacity represents the maximum number of recruits the system can support. This curve takes the form (Froese 2008):

$$\boldsymbol{E(R) = R_{\infty}\left(1 - e^{-\left(\frac{\alpha}{R_{\infty}}\right)S}\right)}$$

where E(R) and S are as above, α is the slope at the origin of the spawner-recruitment curve, and R_{∞} is the carrying capacity of recruits (note that $R_{\infty} = K_R$). As with the Beverton-Holt model, we estimated K_{sp} as the number of spawners needed to produce $0.99(K_R)$.

We used non-linear regression to fit the three models to spawner-recruitment data. Before fitting the models, we transformed recruitment data using natural logs. We estimated bias and uncertainty measures (95% CI) for the model parameters using bootstrap procedures, which assumed that the {R, S} sample represented or approximated the population. The number of bootstrap samples was 3,000. We computed and stored the non-linear regression results for each bootstrap sample. We then calculated the bootstrap 95% CI by arranging the 3,000 bootstrap parameter values in sorted order and selected the 2.5 and 97.5 percentiles from the list.

We used Akaike's Information Criterion for small sample size (AIC_c) to determine which model(s) best explained the relationship between spawners and recruitment in the supplemented and reference populations. AIC_c was estimated as:

$$\boldsymbol{AIC_c = -2log\left(£(\theta|data)\right) + 2K + \left(\frac{2K(K+1)}{n-K-1}\right)}$$

where $log(£(\theta|data))$ is the maximum likelihood estimate, K is the number of estimable parameters (structural parameters plus the residual variance parameter), and n is the sample size (Burnham and Anderson 2002). We used least-squares methods to estimate $log(£(\theta|data))$, which was calculated as $log(\sigma^2)$, where σ^2 = residual sum of squares divided by the sample size ($\sigma^2 = RSS/n$). AIC_c assessed model fit in relation to model complexity (number of parameters). The model with the smallest AIC_c value represented the "best approximating" model within the model set. Remaining models were ranked relative to the best model using AIC_c difference scores (ΔAIC_c), Akaike weights (w_i), and evidence ratios. Models with ΔAIC_c values less than 2 indicated that there is substantial support for these models as being the best-fitting models within the set (Burnham and Anderson 2002). Models with values greater than 2 had less support. Akaike weights are probabilities estimating the strength of the evidence supporting a particular model as being the best model within the model set. Models with small w_i values are less plausible as competing models (Burnham and Anderson 2002). If no single model could be specified as the best model, a "best subset" of competing models was identified using (1) AIC_c differences to indicate the level of empirical support each model had as being the best model, (2) evidence ratios based on Akaike weights to indicate the relative probability that any model is the best model, and (3) coefficients of determination (R^2) assessing the explanatory power of each model.

Stock-Recruitment Analysis

We successfully fit stock-recruitment models to the Chiwawa and reference population data. The span of spawner data for the Chiwawa and reference populations was greater than 14 times the minimum observed spawners, which should provide sufficient contrast for estimation of model parameters. In addition, the span of recruitment data was greater than 12 times the minimum observed recruitment, again providing sufficient contrast for estimation of parameters. The relationship between natural log R/S and spawners indicated that some of the highest productivities occurred at the lower spawner levels and the lowest productivities generally occurred at the highest spawner levels (Figure 15). This is consistent with the assumption of density-dependent mortality.

Although model fits were generally poor, explaining less than 40% of the residual variation in natural-log recruitment data, we were able to estimate average maximum recruitment levels (K_R) and the spawning levels needed to produce maximum recruitment (K_{sp}) (Table 24; Figure 15). For all populations examined, Akaike information criterion was unable to identify a best approximating model (i.e., ΔAIC_c values were less than 2, indicating support for all three models). However, evaluation of 95% CIs and the asymptotic correlation coefficients indicated that the smooth hockey stick model may be the best approximating model for each population. Therefore, we used estimates of K_R and K_{sp} derived from the smooth hockey stick model to correct for density dependence and different carrying capacities in treatment-reference comparisons.

As part of the regression diagnostics, we examined the dependence of the model residuals on time and found a significant ($P < 0.05$), positive, one-year-lag autocorrelation for the Entiat (0.562), Marsh (0.551), Sesech (0.564), and Little Wenatchee (0.629) populations. For the purposes of our work here, we did not attempt to correct for this one-year-lag correlation in the residuals. Future analyses will explore the use of autoregressive models (e.g., AR1; Noakes et al. 1987) to correct for autocorrelation.

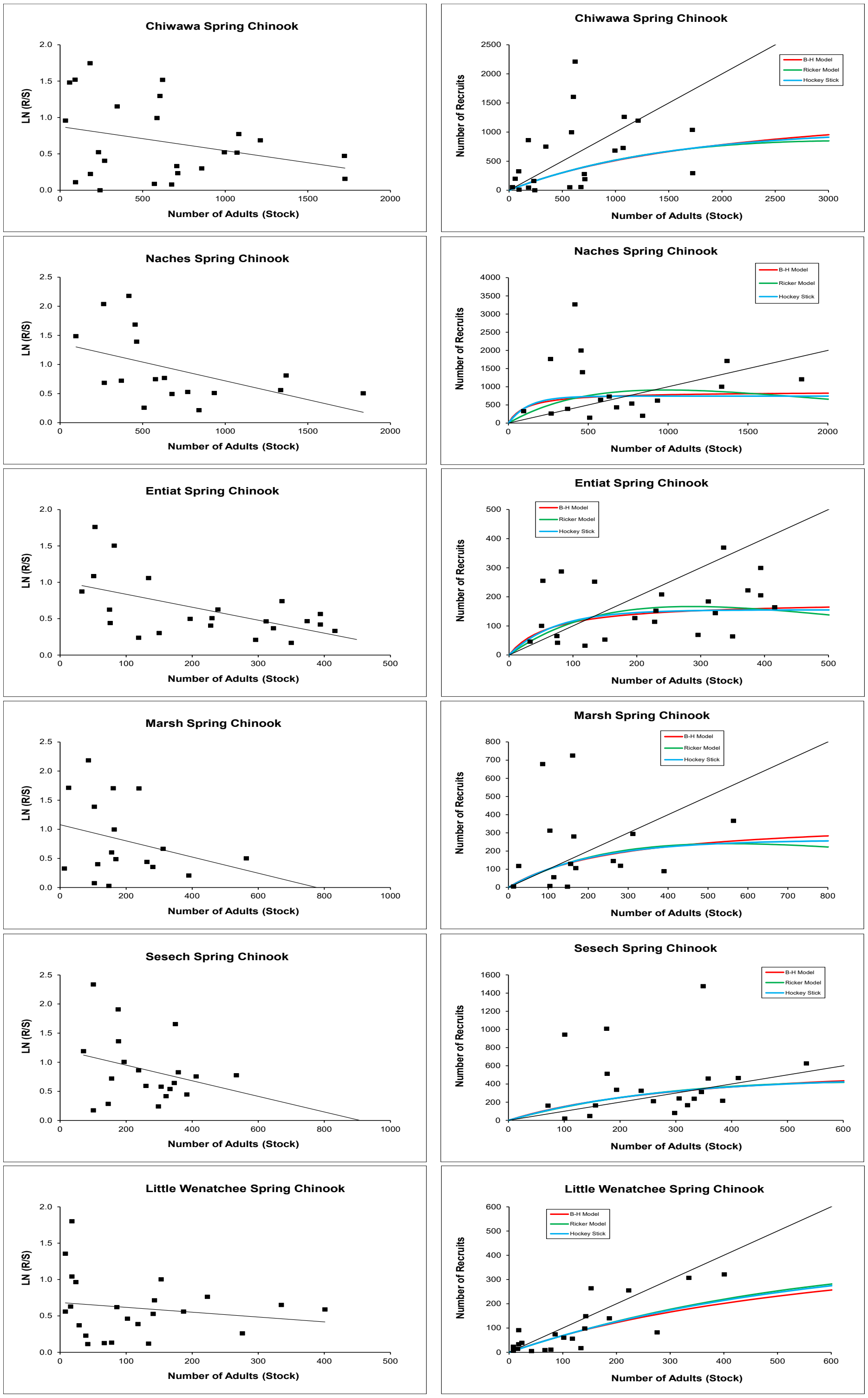


Figure 15. Relationships between natural log recruits/spawner (LN R/S) and spawners (Stock) in the Chiwawa and reference populations (figures on the left) and relationships between numbers of untransformed recruits and spawners in the Chiwawa and reference populations (figures on the right). Figures on the right also show the fit of the Ricker, Beverton-Holt, and the smooth hockey stick models to the data (black straight line represents R=S).

Table 24. Results from fitting Ricker, Beverton-Holt, and smooth hockey stick models to stock-recruitment data from the Chiwawa and reference populations. 95% CI on parameter estimates are based on 3,000 bootstrap trials; Corr coef = asymptotic correlation of the parameter estimates; K_R = maximum natural origin recruits (recruits at carrying capacity); K_{sp} = number of spawners needed to produce K_R; AICc = Akaike's Information Criterion for small sample size; Adj R^2 = coefficient of determination that is adjusted for the number of parameters in the model.

Model	Parameter	Parameter value	Bootstrap 95% CI	Corr coef	K_R	K_{sp}	AICc	Adj R^2
Chiwawa Population								
Ricker	α	0.7048	-0.6197 1.1055	0.791	852	3,285	-47.949	0.125
	β	0.000304	-0.000668 0.000609					
Beverton-Holt	α	1687.4	-65654539 3062.1	0.989	1,687	43,760	-47.962	0.125
	β	2308.5	-99999538 4526.1					
Smooth hockey stick	α	6.956	-41.313 8.2270	-0.708	1,049	6,847	-47.949	0.125
	β	0.7118	-2.397 1.122					
Naches Population								
Ricker	α	2.5223	-2.0003 3.9672	0.844	912	983	-45.063	-0.143
	β	0.001018	-0.000752 0.001717					
Beverton-Holt	α	869.4	97.4 1641.4	0.858	869	11,455	-46.801	-0.097
	β	111.8	-346.2 569.8					
Smooth hockey stick	α	6.612	5.9223 7.006	-0.399	744	565	-46.831	-0.095
	β	6.013	-89.071 12.026					
Entiat Population								
Ricker	α	1.5843	0.1609 2.4178	0.867	167	286	-68.365	-0.049
	β	0.003496	0.001141 0.005906					
Beverton-Holt	α	186.1	67.9 304.3	0.880	186	1,277	-69.895	0.029
	β	65.0	-59.1 189.2					
Smooth hockey stick	α	5.045	4.381 5.378	-0.450	155	344	-69.379	0.003
	β	2.180	-89.369 3.704					
Marsh Creek Population								
Ricker	α	1.1852	-1.8268 1.9269	0.823	241	552	-32.237	0.218

Model	Parameter	Parameter value	Bootstrap 95% CI	Corr coef	K_R	K_{sp}	AICc	Adj R^2
	β	0.001810	-0.003063 0.003625					
Beverton-Holt	α	383.3	-85109314 665.4	0.970	383	5,310	-32.291	0.234
	β	282.4	-99999944 564.9					
Smooth hockey stick	α	5.565	-22.631 6.584	-0.694	261	984	-32.264	0.227
	β	1.265	-108.574 2.531					
Sesech Population								
Ricker	α	1.6835	-2.9253 2.5951	0.912	421	680	-54.589	-0.005
	β	0.001470	-0.002951 0.002941					
Beverton-Holt	α	689.9	-986.8 2366.7	0.981	690	6,591	-54.678	0.000
	β	351.7	-1059.0 1762.5					
Smooth hockey stick	α	6.1528	-22.851 6.815	-0.821	470	1,185	-54.633	-0.002
	β	0.8000	-119.370 2.909					
Little Wenatchee Population								
Ricker	α	0.7447	0.0828 1.0280	0.735	356	1,298	-66.978	0.357
	β	0.000770	-0.003052 0.001541					
Beverton-Holt	α	564.7	-74423355 1067.6	0.994	565	13,400	-67.055	0.358
	β	719.7	-99999856 1413.4					
Smooth hockey stick	α	6.0181	-49.5620 8.1122	-0.683	411	2,544	-67.000	0.357
	β	0.7550	-0.9539 1.0452					

Method 1: Productivity Data Adjusted for Density Dependence

Analysis of supplementation effects on productivity (adult recruits/spawner adjusted for density-dependent effects based on the smooth hockey stick model) using difference scores indicated that supplementation did not significantly decrease productivity in the Chiwawa Basin (Table 25; Figure 16). All tests, regardless of treatment-reference pairs, indicated that productivity did not change significantly during the supplementation period, even though productivity did decrease during the supplementation period (Figure 16). These results are consistent with those based on unadjusted productivity data (Table 17). This is because most abundance estimates were below the level of assumed density dependence.

Table 25. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed productivity data corrected for density dependence. Tests determined if the mean difference scores during the supplementation period were greater than mean difference scores during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Productivity					
Naches	0.904	0.190	0.496	0.412	-0.511 – 1.497
Marsh	-0.203	0.579	0.152	0.927	-1.298 – 1.372
Sesech	1.607	0.071	1.435	0.146	-0.359 – 2.911
Little Wenatchee	0.431	0.335	0.147	0.668	-0.487 – 0.781
LN Productivity					
Naches	0.570	0.290	0.083	0.568	-0.168 – 0.362
Marsh	0.012	0.495	0.003	0.991	-0.373 – 0.480
Sesech	1.463	0.087	0.343	0.171	-0.125 – 0.732
Little Wenatchee	0.390	0.351	0.060	0.709	-0.218 – 0.365

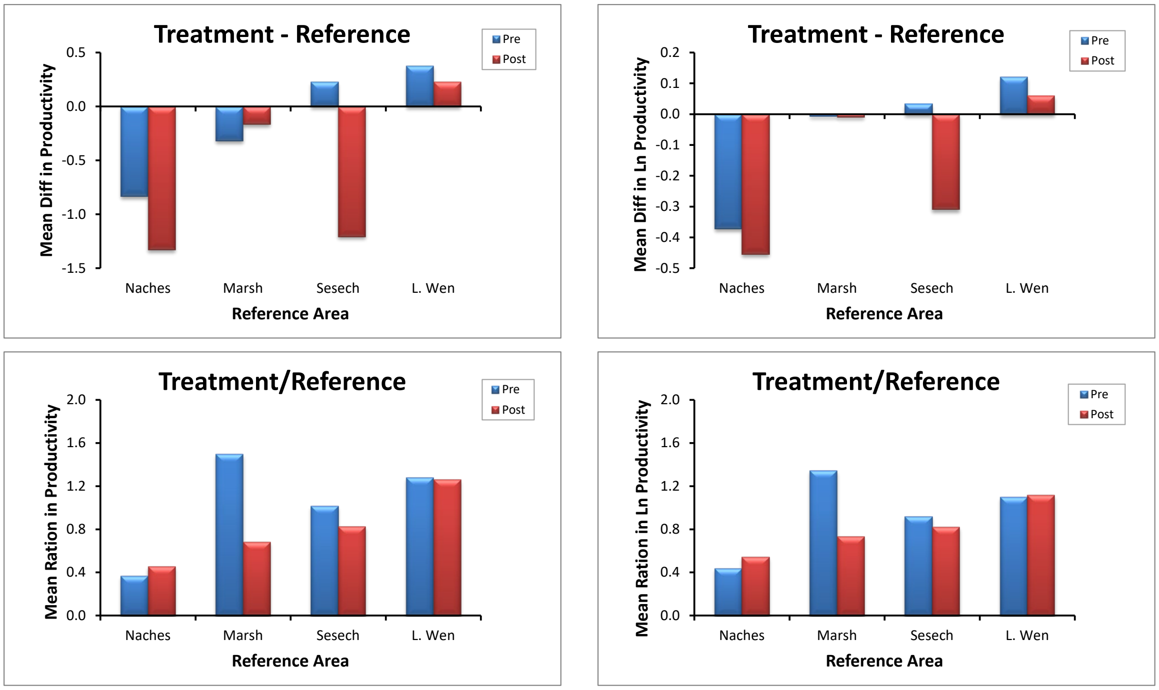


Figure 16. Mean differences (Treatment – Reference; figures on the top) and mean ratios (Treatment/Reference; figures on the bottom) of transformed and untransformed productivity data (adjusted for density dependence) before (pre) and after (post) spring Chinook supplementation in the Chiwawa Basin. Negative effects of supplementation on productivity are indicated when the pre-supplementation (blue) bars are greater than their corresponding post-supplementation (red) bars.

Analysis of supplementation effects on productivity (adult recruits/spawner adjusted for density-dependent effects) using ratios indicated that supplementation did not significantly decrease productivity in the Chiwawa Basin (Table 26; Figure 16). The Aspin-Welch test and the 95% CIs did indicate a significant effect when comparing the Chiwawa to the Marsh Creek population. These results are consistent with those using unadjusted productivity data (Table 20). Again, this is because most abundance estimates were below the level of assumed density dependence.

Table 26. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on transformed and untransformed productivity data corrected for density dependence. Tests determined if the mean ratios during the supplementation period were less than mean ratios during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Productivity					
Naches	-0.529	0.696	0.087	0.597	-0.394 – 0.214
Marsh	2.236	0.022	0.814	0.056	0.140 – 1.470
Sesech	0.677	0.253	0.191	0.496	-0.343 – 0.727
Little Wenatchee	0.033	0.487	0.018	0.978	-0.902 – 1.181
LN Productivity					
Naches	-0.621	0.726	0.104	0.536	-0.406 – 0.191

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Marsh	1.952	0.036	0.613	0.076	0.005 – 1.163
Sesech	0.447	0.330	0.098	0.649	-0.312 – 0.498
Little Wenatchee	-0.034	0.513	0.015	0.980	-0.697 – 0.852

Our analyses assume that there is a spawner abundance (K_{sp}) at which density-independent effects end and density-dependent effects begin. In reality, density-dependent effects occur at low spawning abundance and intensify as spawning abundance increases (evident in the changing slope of the three stock-recruitment curves used in our analyses). We did not account for these increasing density-dependent effects at spawner abundances less than K_{sp}. If we accounted for the increasing effects of density dependence at spawning abundances less than K_{sp}, the analysis with and without productivity adjustments may give different results.

Method 2: Fraction of Carrying Capacity Filled with NORs

We analyzed the effects of supplementation on filling the capacity of the habitat with natural-origin recruits. The smooth hockey stick model derived the carrying capacity (K_R) estimates for the Chiwawa and reference populations. The fraction of the carrying capacity filled with Chinook recruits before and during supplementation for the Chiwawa and reference populations is provided in Table 27. These data indicate that for the Chiwawa population, the mean fraction of the K_R filled with fish decreased significantly from the pre-supplementation period through the supplementation period (Table 27). Likewise, the Entiat and Little Wenatchee populations showed a significant decline in the mean fraction of K_R filled with adult recruits. In contrast, the mean fraction of K_R in the Naches and Marsh Creek populations increased during the same period (Table 27).[22] Interestingly, the fraction of K_R filled with adult recruits for all populations trended downward during the pre-supplementation period (Figure 17). During the supplementation period, however, the fraction of K_R filled with adult recruits trended upward for all populations. These results suggest that agents of mortality outside the Chiwawa and reference populations were reducing recruitment to the populations.

---

[22] Although we do not show the results here, statistical analysis of the mean fraction of carrying capacity filled by adult recruits using natural-log transformed data produced the same result as using untransformed data. This was true for all populations.

Table 27. Fraction of the carrying capacity that was filled with Chinook salmon adult recruits in the Chiwawa and reference populations before (pre) and during (post) supplementation in Chiwawa Basin. The smooth hockey stick model estimated carrying capacity for each population. Statistical results from comparing the pre and post mean scores using the Aspin-Welch unequal-variance test are provided at the bottom of the table.

Supplementation period	**Chiwawa**	**Reference populations**			
		Naches	**Entiat**	**Marsh**	**L. Wenatchee**
Pre-supplementation period (1981-1992)	2.11		2.38	1.07	0.64
	1.53		1.93	1.20	0.75
	1.20		1.32	2.60	0.78
	1.14		1.19	0.49	0.62
	0.99		1.06	0.46	0.34
	0.70	2.30	1.43	0.56	0.24
	0.65	0.58	0.74	0.34	0.20
	0.95	1.88	1.34	1.40	0.36
	0.18	0.72	1.63	0.22	0.15
	0.05	0.27	0.45	0.02	0.02
	0.00	0.20	0.21	0.03	0.01
Pre-Mean:	0.86	0.99	1.24	0.76	0.37
Pre-Range:	0.00 – 2.11	0.20 – 2.30	0.21 – 2.38	0.02 – 2.60	0.01 – 0.78
Post-supplementation period (1992-2002)	0.05	0.98	0.34	0.41	0.03
	0.15	0.86	0.41	1.13	0.04
	0.04	0.35	0.27	0.02	0.03
	0.05	0.44	0.30	0.02	0.03
	0.19	4.39	0.65	0.45	0.06
	0.82	2.68	1.85	2.78	0.22
	0.31	2.37	1.65	4.10	0.08
	0.01	0.53	0.42		0.02
	0.71	1.62	0.82		0.10
	0.28	1.35	0.93		0.14
	0.27	0.83	0.98		0.18
Post-Mean:	0.26	1.49	0.78	1.27	0.08
Post-Range:	0.04 – 0.82	0.35 – 4.39	0.30 – 1.85	0.02 – 4.10	0.02 – 0.22
One-sided Aspin-Welch t-test of pre and post means	t = 2.846; P = 0.007	t = -0.967; P = 0.825	t = 1.833; P = 0.041	t = -0.799; P = 0.776	t = 3.321; P = 0.003

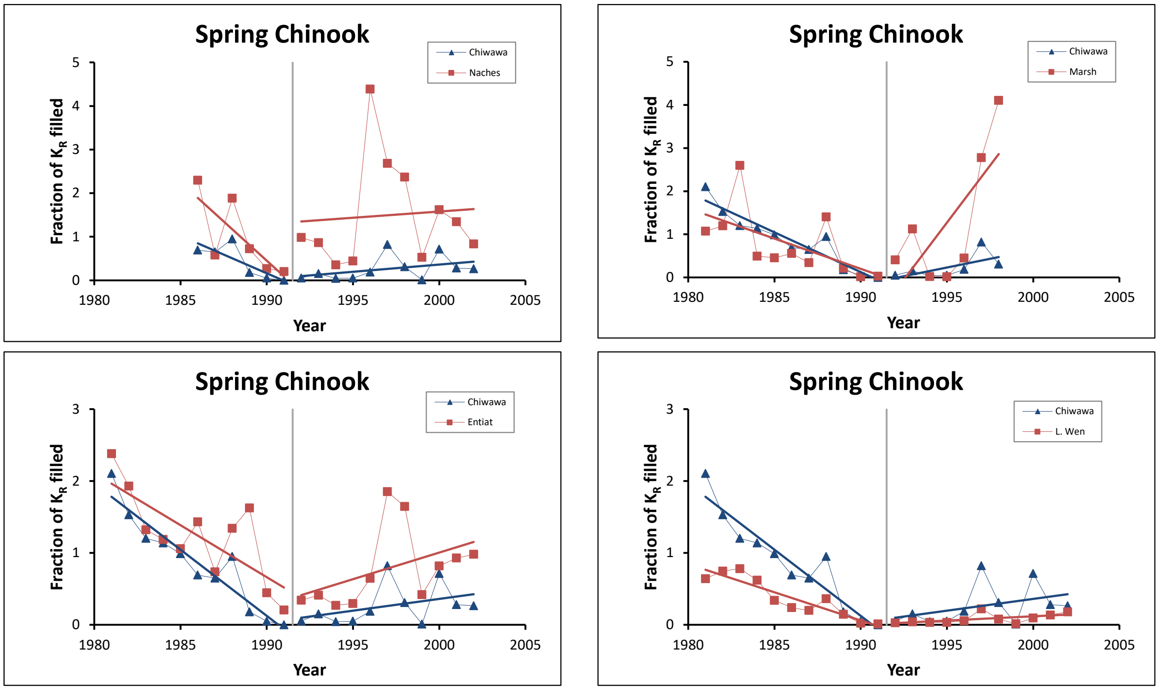


Figure 17. Trends in the fraction of the carrying capacity that was filled with Chinook salmon adult recruits in the Chiwawa and reference populations before (pre) and during (post) supplementation in Chiwawa Basin. The vertical lines in the figures separate the pre- and post-supplementation periods. The smooth hockey stick model estimated carrying capacity for each population.

We then compared the mean difference scores and ratios between the Chiwawa and reference populations before and during supplementation using data representing the fraction of K_R filled with adult recruits. In most of the Chiwawa-reference population comparisons, the absolute value of the mean difference between the fraction of K_R filled with recruits was greater in the supplementation period than during the pre-supplementation period; two of the four pairings were significant (Table 28; Figure 18). Analysis of difference scores using natural-log transformed data indicated that three of the four pairings were significant (Table 28).

Results from analyses using ratios were similar to results using difference scores. Mean ratio scores were generally smaller during the supplementation period than during the pre-supplementation period (Figure 18). This indicated that the mean fraction of K_R filled by adult recruits in most reference populations was greater during the supplementation period than during the pre-supplementation period (i.e., the denominator in the ratio increased between the pre- and post-supplementation periods). In contrast, the fraction of K_R filled by adult recruits in the Chiwawa decreased from the pre- to post-supplementation period (i.e., the numerator in the ratio decreased between the pre- and post-supplementation periods). Thus, unlike the Chiwawa population, the capacity of most reference populations was becoming more saturated during the period when the Chiwawa was being supplemented. Statistical analysis with mean ratios indicated that two of the four pairings were significant (Table 29).

Analyses comparing the Little Wenatchee with the Chiwawa indicate that adult recruits to the Little Wenatchee have been well below its carrying capacity. During the pre-supplementation period, the capacity of the Little Wenatchee was on average 37% saturated with adult recruits. During the supplementation period, the capacity of the Little Wenatchee declined to 8% saturation with adult recruits (a 22% decline). The Chiwawa, during the pre-supplementation period, was on average 86% saturated. During the supplementation period, percent saturation in the Chiwawa decreased to 26% (a 30% decrease). During the same time periods, the capacity of the Entiat population, which until recently has been supplemented, declined from 124% to 78% saturation (a 63% decline).

Table 28. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on the fraction of the habitat capacity (K_R) that is filled with natural origin recruits. Analyses include both transformed and untransformed data. Tests determined if the mean difference scores during the supplementation period were greater than mean difference scores during the pre-supplementation period.

Reference population	**Aspin-Welch unequal-variance test**			**Randomization test P-value**	**Bootstrap 95% CI**
	t-value	**P-value**	**Effect size**		
Fraction of Capacity Filled					
Naches	1.550	0.071	0.657	0.145	-0.173 – 1.378
Entiat	0.835	0.207	0.141	0.422	-0.167 – 0.475
Marsh	2.026	0.040	1.141	0.055	0.064 – 2.054
Little Wenatchee	2.166	0.023	0.310	0.031	0.035 – 0.569
LN Fraction of Capacity Filled					
Naches	2.123	0.026	0.311	0.039	0.031 – 0.575
Entiat	1.405	0.087	0.122	0.176	-0.034 – 0.289
Marsh	2.547	0.017	0.519	0.017	0.125 – 0.864
Little Wenatchee	1.744	0.049	0.130	0.100	-0.004 – 0.273

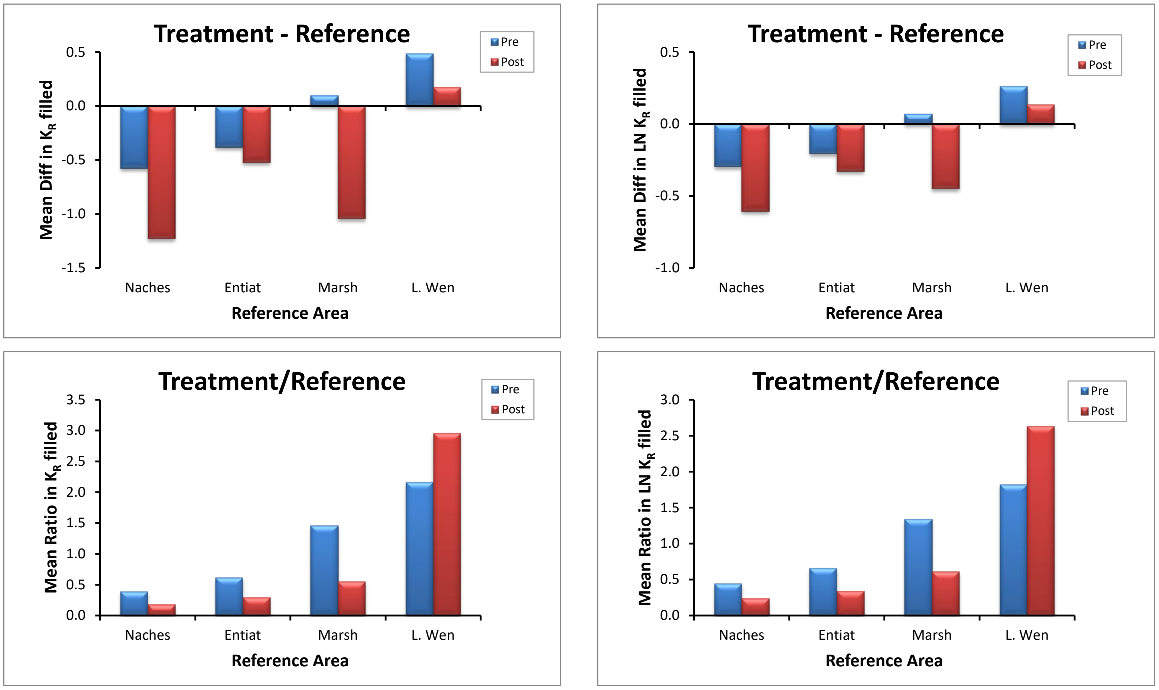


Figure 18. Mean differences (Treatment – Reference; figures on the top) and mean ratios (Treatment/Reference; figures on the bottom) of transformed and untransformed fractions of carrying capacity filled with adult recruits before (pre) and after (post) spring Chinook supplementation in the Chiwawa Basin.

Table 29. Results of the Aspin-Welch unequal-variance test, randomization test (based on 10,000 Monte Carlo samples), and 95% CI (based on 5,000 bootstrap samples) on the fraction of the habitat capacity (K_R) that is filled with natural origin recruits. Analyses include both transformed and untransformed data. Tests determined if the mean ratios during the supplementation period were less than mean ratios during the pre-supplementation period.

Reference population	Aspin-Welch unequal-variance test			Randomization test P-value	Bootstrap 95% CI
	t-value	P-value	Effect size		
Fraction of Capacity Filled					
Naches	1.317	0.119	0.217	0.219	-0.103 – 0.482
Entiat	2.449	0.013	0.321	0.028	0.085 – 0.577
Marsh	2.001	0.035	0.905	0.070	0.138 – 1.788
Little Wenatchee	-1.148	0.864	0.791	0.278	-1.979 – 0.578
LN Fraction of Capacity Filled					
Naches	1.257	0.127	0.207	0.249	-0.099 – 0.484
Entiat	2.346	0.016	0.313	0.031	0.072 – 0.583
Marsh	1.737	0.056	0.729	0.111	0.028 – 1.531
Little Wenatchee	-1.525	0.924	0.815	0.142	-1.751 – 0.195

Comparing Stock-Recruitment Curves

As a final set of treatment and reference population comparisons, we compared the stock-recruitment curves of the Chiwawa population (using {R, S} data only from the supplementation period) to the reference populations (using all available {R, S} data). Specifically, we tested whether the regression parameters were equal between the Chiwawa population and the reference populations, and whether the fitted curves coincided between populations. Earlier in this report we described the data, methods, and results of fitting the Ricker, Beverton-Holt, and smooth hockey stick curves to the data. Because AIC_c was unable to identify a best approximating model, here we included all three models in our analyses. We tested the following hypotheses.

Parameter equivalence:

Ho: Stock-recruitment parameters (α and β) of the Chiwawa population = Stock-recruitment parameters of the reference populations.

Ha: Stock-recruitment parameters (α and β) of the Chiwawa population $\neq$ Stock-recruitment parameters of the reference populations.

Curve equivalence:

Ho: Modeled stock-recruitment curves of the Chiwawa population = Modeled stock-recruitment curves of the reference populations.

Ha: Modeled stock-recruitment curves of the Chiwawa population $\neq$ Modeled stock-recruitment curves of the reference populations.

We used two-sided randomization tests to test the null hypotheses of equal model parameters and that fitted curves coincided. Because the total number of permutations was in the millions, we used a Monte Carlo approach to randomly select 10,000 permutations. The test statistic for comparing the model parameters was formed by summing the difference between the population parameter estimates for each pair of populations. The test statistic for comparing the whole curve was formed by summing the difference between the estimated predicted values for each pair of populations at 500 equally spaced points along the curve.

Ricker Relationships

Ricker curves differed significantly between the Chiwawa and reference populations (Figure 19; Table 30). Interestingly, however, the parameters in the Ricker model did not differ significantly among most populations (Table 30). Only the β parameter differed significantly between the Chiwawa and Entiat populations.

In the Ricker model, the α parameter represents intrinsic productivity (i.e., recruits per spawner at low spawner densities). In this analysis, there was not enough evidence in the stock-recruitment data to reject the hypothesis of inequality in intrinsic productivity. Thus, this test was unable to demonstrate that supplementation, based on the Ricker curve, affected productivity in the Chiwawa population.

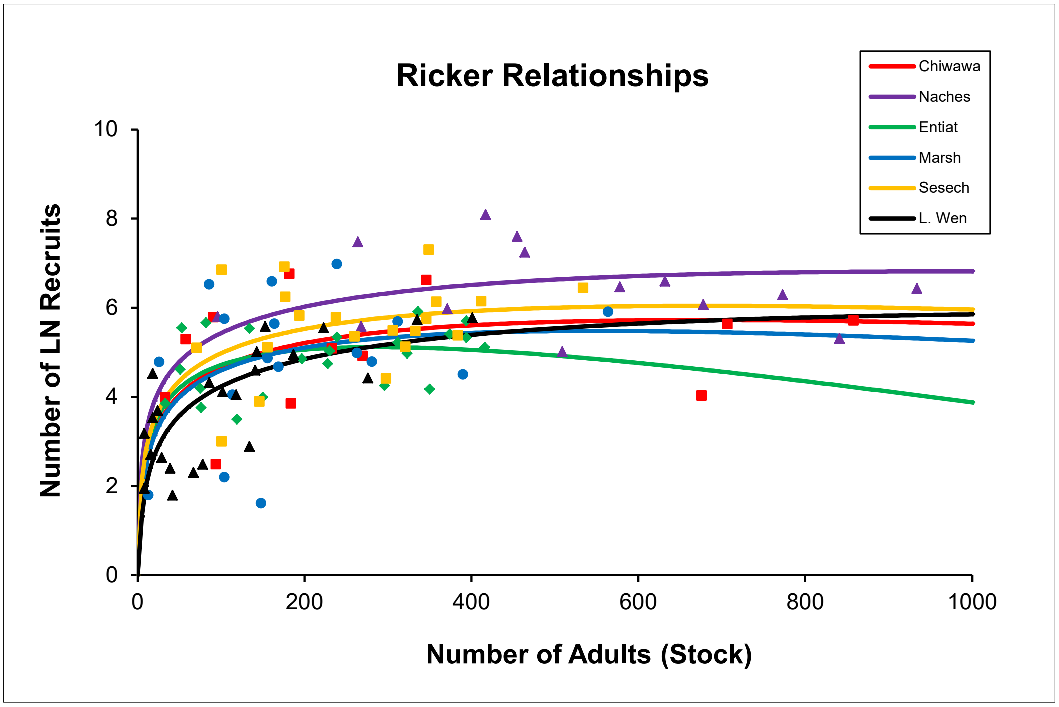


Figure 19. Scatter plot of the number of spawners and natural log adult recruits and fitted Ricker curves to the Chiwawa (supplemented population) and reference (un-supplemented) populations.

Table 30. Randomization test results comparing the equality of Ricker curves and equality of parameter values (α and β). Randomization tests were based on 10,000 Monte Carlo samples. Equality or curves was based on 500 points along the x-axis (spawner abundance axis).

Curves tested	Curve inequality randomization P-value	Parameter inequality		
		Model Parameter		Randomization P-value
		Chiwawa	Reference	
Chiwawa v. Naches	0.008	α = 1.2247	α = 2.5267	0.236
		β = 0.0015	β = 0.0010	0.600
Chiwawa v. Entiat	0.004	α = 1.2247	α = 1.5836	0.978
		β = 0.0015	β = 0.0035	0.025
Chiwawa v. Marsh	0.034	α = 1.2247	α = 1.1855	0.997
		β = 0.0015	β = 0.0018	0.688
Chiwawa v. Sesech	0.036	α = 1.2247	α = 1.6818	0.972
		β = 0.0015	β = 0.0015	0.997
Chiwawa v. L. Wenatchee	0.034	α = 1.2247	α = 0.7439	0.969
		β = 0.0015	β = 0.0008	0.203

Beverton-Holt Relationships

Beverton-Holt curves differed significantly only between the Chiwawa and Naches populations (Figure 20; Table 31). There was no significant difference in curves between the Chiwawa and the other reference populations. The parameters in the Beverton-Holt model did not differ significantly among any of the populations (Table 31). This was true even for the Chiwawa and Naches populations.

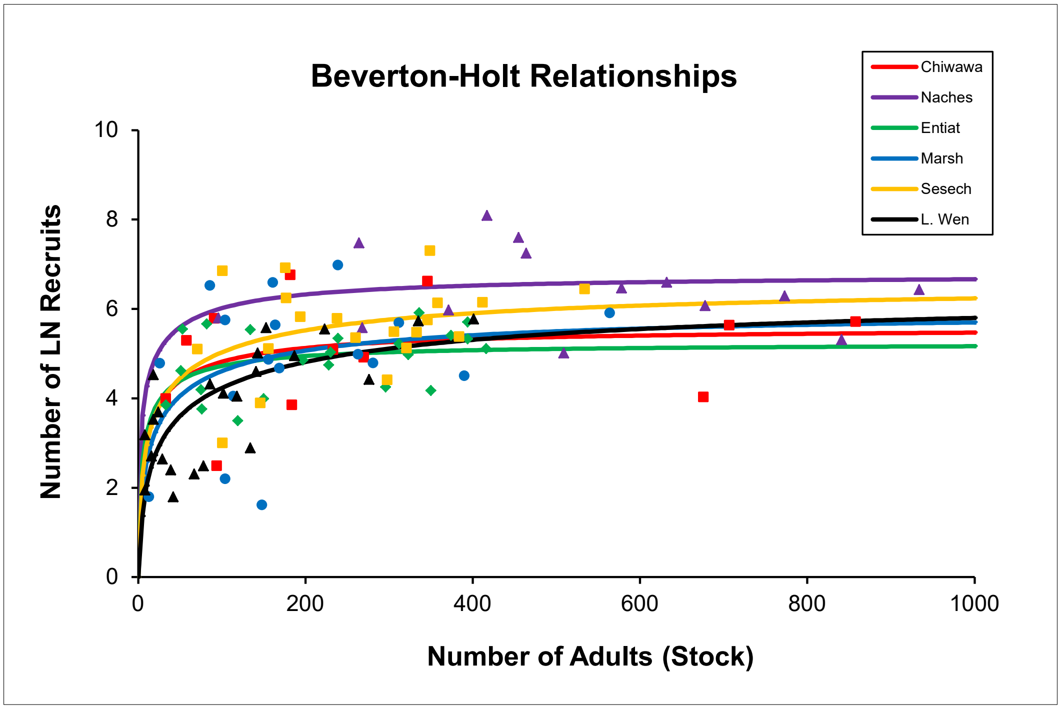


Figure 20. Scatter plot of the number of spawners and natural log adult recruits and fitted Beverton-Holt curves to the Chiwawa (supplemented population) and reference (un-supplemented) populations.

Table 31. Randomization test results comparing the equality of Beverton-Holt curves and equality of parameter values (α and β). Randomization tests were based on 10,000 Monte Carlo samples. Equality or curves was based on 500 points along the x-axis (spawner abundance axis).

Curves tested	Curve inequality randomization P-value	Parameter inequality		
		Model Parameter		Randomization P-value
		Chiwawa	Reference	
Chiwawa v. Naches	0.036	α = 264.25	α = 870.62	0.777
		β = 113.79	β = 112.24	0.963
Chiwawa v. Entiat	0.746	α = 264.25	α = 186.34	0.960
		β = 113.79	β = 65.33	0.954
Chiwawa v. Marsh	0.850	α = 264.25	α = 381.79	0.944
		β = 113.79	β = 281.04	0.891

Curves tested	Curve inequality randomization P-value	Parameter inequality		
		Model Parameter		Randomization P-value
		Chiwawa	Reference	
Chiwawa v. Sesech	0.272	α = 264.25	α = 689.31	0.821
		β = 113.79	β = 351.59	0.869
Chiwawa v. L. Wenatchee	0.654	α = 264.25	α = 568.69	0.864
		β = 113.79	β = 725.87	0.751

Smooth Hockey Stick Relationships

Smooth hockey stick curves differed significantly between the Chiwawa and Naches populations and the Chiwawa and Sesech populations (Figure 21; Table 32). There was no significant difference in curves between the Chiwawa and the other reference populations. Most of the parameters in the smooth hockey stick model did not differ significantly among the populations (Table 32). However, the productivity parameter β did differ significantly between the Chiwawa and the Naches and the Chiwawa and Little Wenatchee populations. The β parameter for the Naches was significantly greater than the Chiwawa, while the β parameter for the Little Wenatchee was significantly less than the Chiwawa.

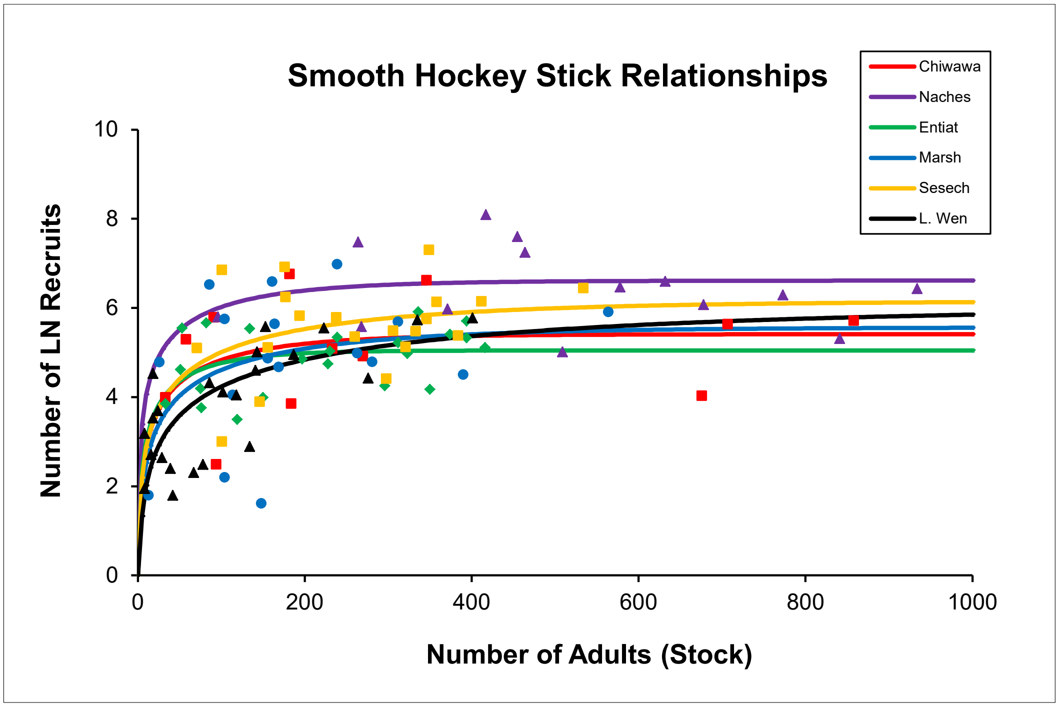


Figure 21. Scatter plot of the number of spawners and natural log adult recruits and fitted smooth hockey stick curves to the Chiwawa (supplemented population) and reference (un-supplemented) populations.

Table 32. Randomization test results comparing the equality of smooth hockey stick curves and equality of parameter values (α and β). Randomization tests were based on 10,000 Monte Carlo samples. Equality or curves was based on 500 points along the x-axis (spawner abundance axis).

Curves tested	**Curve inequality randomization P-value**	**Parameter inequality**		
		Model Parameter		**Randomization P-value**
		Chiwawa	**Reference**	
Chiwawa v. Naches	0.000	α = 5.41	α = 6.61	0.000
		β = 1.84	β = 5.99	0.000
Chiwawa v. Entiat	0.999	α = 5.41	α = 5.05	0.999
		β = 1.84	β = 2.17	0.999
Chiwawa v. Marsh	0.999	α = 5.41	α = 5.56	0.999
		β = 1.84	β = 1.27	0.999
Chiwawa v. Sesech	0.000	α = 5.41	α = 6.15	0.000
		β = 1.84	β = 1.80	0.999
Chiwawa v. L. Wenatchee	0.990	α = 5.41	α = 6.02	0.999
		β = 1.84	β = 0.75	0.000

Comparing different stock-recruitment curves and their parameters did not provide strong evidence that the supplementation program has negatively affected the productivity of the Chiwawa population.

Analysis without Reference Populations

In some cases, suitable reference populations may not exist to compare with supplemented populations. It is therefore important to have alternative analyses to assess supplementation effects. In this section, we describe methods that can be used to assess supplementation effects when suitable reference populations are not available. We discuss before-after comparisons, correlation analysis, and comparisons to standards as alternatives when reference populations are unavailable.

Before-After Comparisons

Before-after analyses compare population metrics (spawner abundance, NORs, and productivity) before supplementation to those during supplementation. In this case, data collected before supplementation represent the reference condition. The assumption is that population trajectories measured during the pre-supplementation period would continue in the absence of supplementation. We compared trends in abundance and productivity, mean abundance and productivity, and stock-recruitment relationships before and after supplementation.

Trend Analysis

Comparing trends before and after supplementation can be used to assess the effects of supplementation. Here, we compared the slopes of trends of spawner abundance, NORs, and productivity before and during supplementation using t-tests. If the hatchery program is successfully supplementing the natural spring Chinook population, the trend for spawner abundance and NORs during supplementation should be greater than the slope during the pre-supplementation period. For productivity, the slope during the supplementation period should increase or remain the same as that during the pre-supplementation period.

Visual examination of trends of Chiwawa data indicates that spawner abundance, NORs, and productivity decreased during the pre-supplementation period, but increased during the supplementation period (Figure 22). Only the changes in NOR trends were significant (Figure 22). This was true for both transformed and untransformed data.

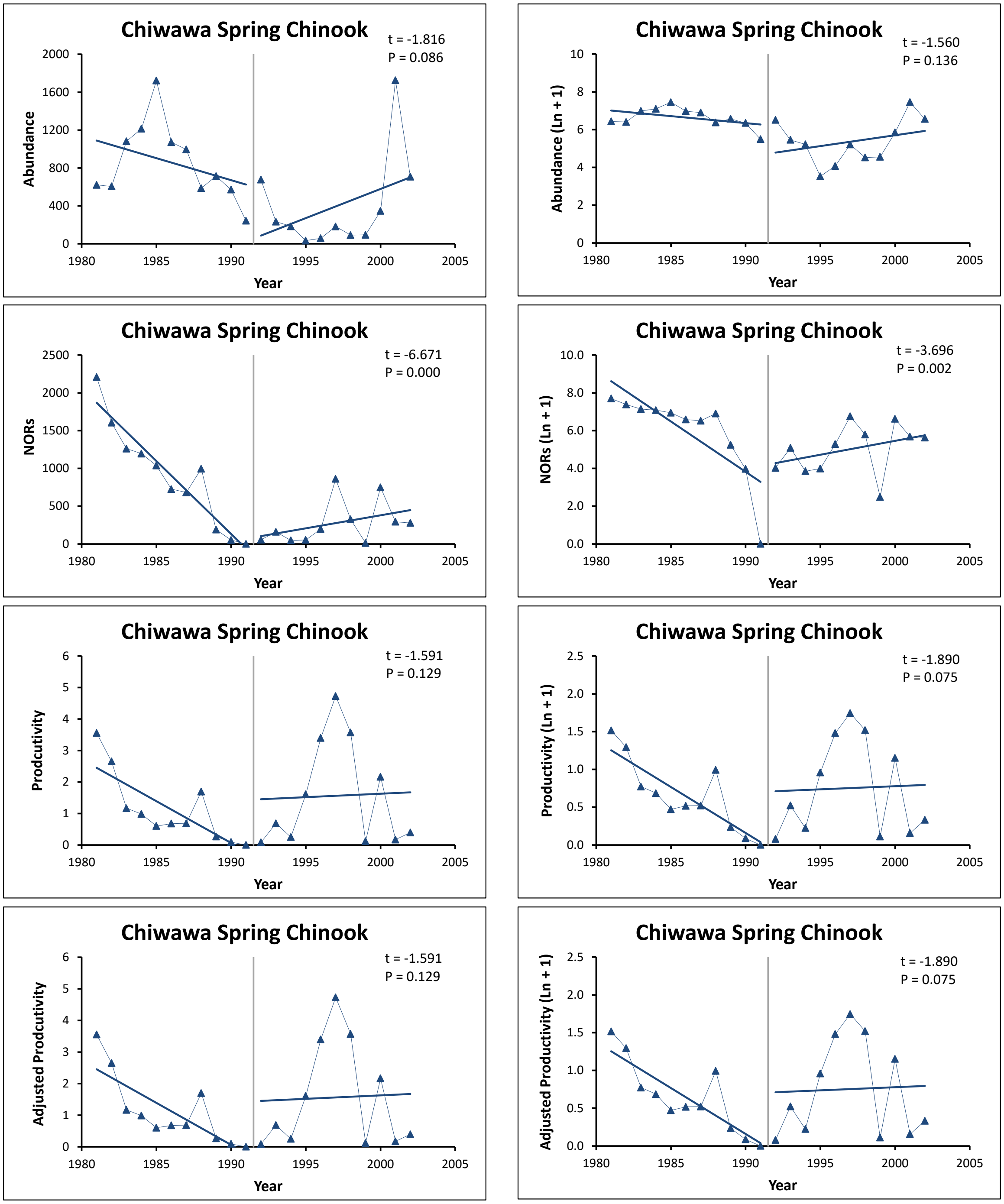


Figure 22. Trends in Chiwawa spring Chinook spawner abundance, natural-origin recruits (NORs), productivity (adults recruits per spawner), and adjusted productivity (adjusted for density dependence) before and during supplementation. The vertical lines in the figures separate the pre- and post-supplementation periods. Figures on the left show untransformed data; figures on the right include natural-log transformed data. Figures include results of t-tests comparing slope of trends before and during supplementation.

Analysis of Mean Scores

We also compared mean spawner abundance, NORs, and productivity data before and after supplementation. If the hatchery program is successfully supplementing the natural spring Chinook population, mean spawner abundance and NORs during the supplementation period should be greater than the pre-supplementation period. For productivity, the mean productivity during the supplementation period should be equal to or higher than the pre-supplementation period. We tested the following statistical hypotheses.

Spawner Abundance and NORs:

> Ho: Mean spawner abundance and NORs before supplementation $\geq$ Mean spawner abundance and NORs during supplementation.
>
> Ha: Mean spawner abundance and NORs before supplementation $<$ Mean spawner abundance and NORs during supplementation.

Productivity (Recruits/Spawner):

> Ho: Mean productivity before supplementation $\leq$ Mean productivity during supplementation.
>
> Ha: Mean productivity before supplementation $>$ Mean productivity during supplementation.

We tested before-after supplementation effects using a one-tailed Aspin-Welch unequal-variance test. We also used a randomization test, based on 10,000 Monte Carlo simulations, to assess differences in spawner abundance and productivity before and during supplementation. The randomization procedure only allowed the testing of two-tailed hypotheses. Therefore, we generated 95% confidence intervals on the mean difference (μ_{pre} - μ_{post}) using bootstrapping methods to determine if the significant result from the randomization test was in the right direction. We generated 5,000 bootstrap samples to calculate confidence intervals.

Mean spawner abundance during the supplementation period was significantly less than the pre-supplementation spawner abundance (Table 33). Mean spawner abundance decreased 46% between the pre- and post-supplementation periods. Likewise, mean NORs decreased significantly between the two periods (Table 33). On the other hand, productivity increased slightly, but not significantly, between the pre- and post-supplementation periods (Table 33). This was true for both adjusted and transformed productivity data.

Table 33. Statistical results comparing mean scores of spawner abundance, natural-origin recruits (NORs), and productivity (using both untransformed and natural-log transformed) before and during supplementation of Chiwawa spring Chinook. Randomization tests were based on 10,000 Monte Carlo samples and 95% CI were based on 5,000 bootstrap samples.

Population metric	Mean scores		Test on means			
			Aspin-Welch test		Random test P-value	Bootstrap 95% CI
	Before	During	t-value	P-value		
Abundance	856	393	2.383	0.986	0.028	112 - 843
LN Abundance	6.6	5.4	3.304	0.997	0.004	0.56 – 1.99
NORs	905	275	2.846	0.993	0.009	214 – 1034
LN NORs	6.0	5.0	1.197	0.876	0.250	-0.40 – 2.54
Productivity	1.13	1.56	-0.721	0.759	0.479	-1.55 – 0.73
LN Productivity	0.64	0.75	-0.450	0.671	0.649	-0.55 – 0.35
Adj Productivity	1.12	1.56	-0.721	0.759	0.477	-1.54 – 0.71
LN Adj Productivity	0.64	0.75	-0.450	0.671	0.652	-0.57 – 0.34

Analysis of Stock-Recruitment Curves

The third method compared stock-recruitment curves of the Chiwawa population during supplementation with those generated before supplementation. Specifically, we tested whether the regression parameters were equal between the pre- and post-supplementation periods, and whether the fitted curves coincided between the two time periods. We used the methods described earlier to fit the Ricker, Beverton-Holt, and smooth hockey stick curves to the two data sets. We tested the following hypotheses.

Parameter equivalence:

Ho: Stock-recruitment parameters (α and β) of the pre-supplementation period = Stock-recruitment parameters of the supplementation period.

Ha: Stock-recruitment parameters (α and β) of the pre-supplementation period ≠ Stock-recruitment parameters of the supplementation period.

Curve equivalence:

Ho: Modeled stock-recruitment curves from the pre-supplementation period = Modeled stock-recruitment curves from the pre-supplementation period.

Ha: Modeled stock-recruitment curves from the pre-supplementation period ≠ Modeled stock-recruitment curves from the pre-supplementation period.

We were only able to fit stock-recruitment curves to the post-supplementation data. Non-linear regression was unable to converge on a solution using only pre-supplementation data. Therefore, we were unable to use this method to test supplementation effects on the Chiwawa spring Chinook population. If we could have fit curves to both the pre- and post-supplementation periods, we would have used two-sided randomization tests to evaluate the null hypotheses of equal model parameters and that fitted curves coincided.

Before describing correlation approaches, it is important to note that comparing before-after data can sometimes be misleading. For example, the spawner abundance, NORs, and productivity data presented in Figure 22 suggest that supplementation is increasing the abundance and productivity of spring Chinook in the Chiwawa Basin. However, when we compared these trends to those from reference populations during the same time periods (Figures 9-11), it becomes clear that supplementation was not responsible for increasing the trends in spawner abundance, NORs, and productivity of the Chiwawa population. Thus, whenever possible, it is wise to compare before-after data with a reference population.

Correlation Analyses

A simple way to see if the supplementation program is increasing or decreasing productivity is to assess the association between the proportion of adult spawners that are made up of hatchery adults (pHOS) and productivity (recruits/spawner). If the supplementation program is working as planned, the increase in hatchery fish spawning naturally should increase the productivity of the population. It should not decrease the productivity of the population.

We tested the association between pHOS and adult productivity[23] using Pearson correlation. During the pre-supplementation period, productivity averaged 1.13 recruits/spawner; during the supplementation period, productivity averaged 1.39 recruits/spawner. This increase in productivity did not appear to be strongly correlated to pHOS (Figure 23). Correlation analysis showed that there was no significant association between pHOS and productivity, even though productivity increased with increasing pHOS.

---

[23] Note that the analysis could also include juvenile productivity (e.g., smolts/spawner).

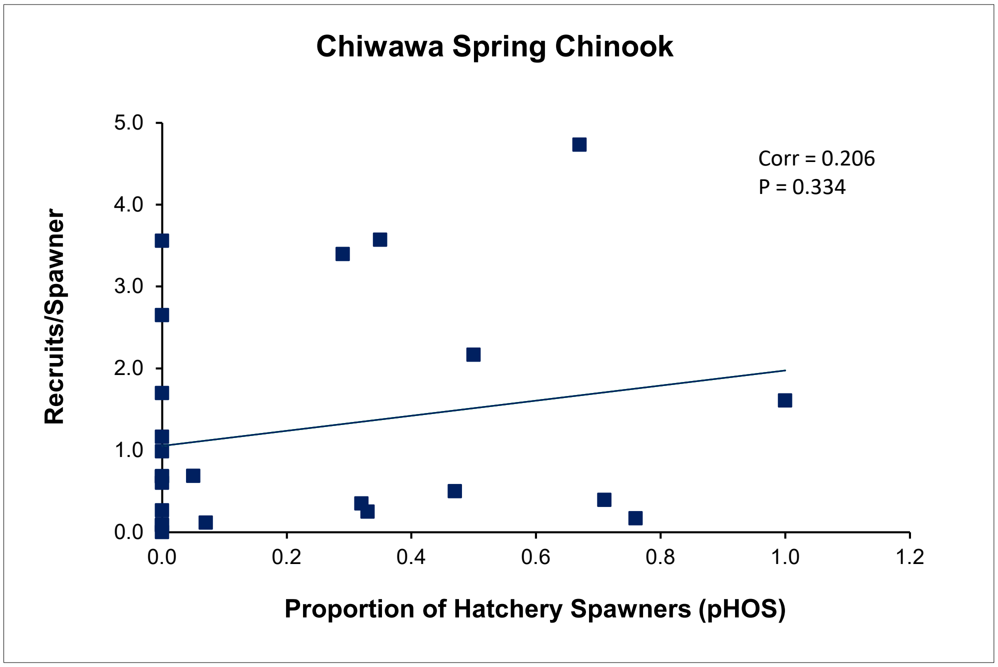


Figure 23. Association between the proportion of spawners that are made up of hatchery adults (pHOS) and the number of natural-origin recruits. The Pearson correlation coefficient (Corr) and its P-value (P) are shown in the figure.

The association between pHOS and productivity can also be assessed by testing the correlation between pHOS and the residuals from stock-recruitment curves fitted to the Chiwawa spawner and natural-origin recruitment data. This approach removes the effects of density dependence on the relationship between pHOS and productivity. A significant negative association provides evidence that hatchery-origin spawners may not be as productive as natural-origin spawners.

The Ricker, Beverton-Holt, and smooth hockey stick models were fit to the Chiwawa stock and recruitment data (including {S, R} data from both the pre- and post-supplementation period, 1981-2004) using methods described earlier. Residuals were calculated by subtracting the predicted recruitment values from the observed (modeled) values. Pearson correlation then tested the association between pHOS and the residuals from each model.

Although there was a negative trend in residuals with increasing pHOS, suggesting that hatchery-origin spawners may not be as productive as natural-origin spawners, the association was not significant (Figure 24). Thus, based on these analyses, there is no strong evidence that the supplementation program has significantly benefited or harmed the natural spring Chinook population.

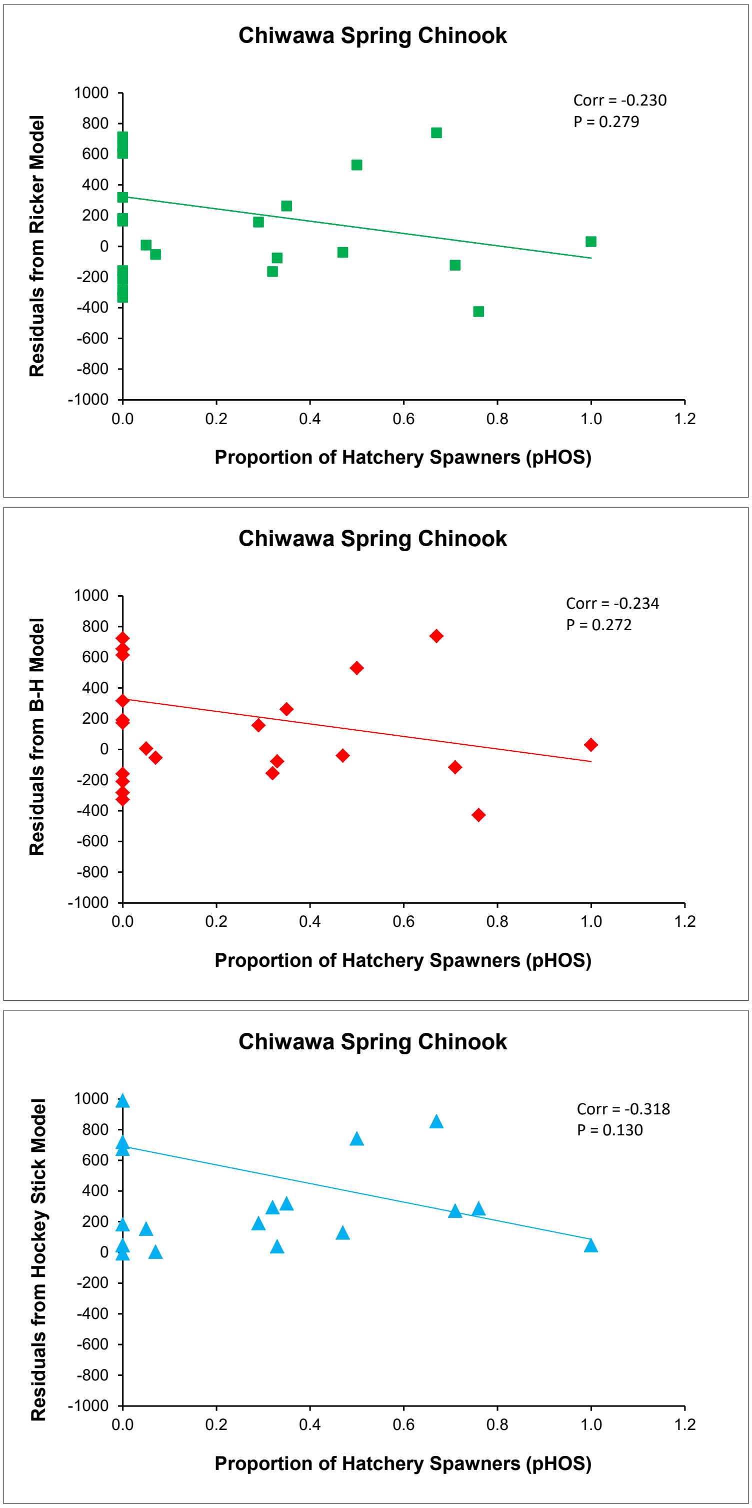


Figure 24. Association between the proportion of spawners that are made up of hatchery adults (pHOS) and the residuals from Ricker, Beverton-Holt (B-H), and smooth hockey stick stock-recruitment models. The Pearson correlation coefficient (Corr) and its P-value (P) are shown in the figures.

Comparison to Standards

In those cases in which suitable reference populations are not available and there are no pre-supplementation data, the investigator is left with comparing population parameters to relevant standards. Standards can include performance of natural-origin fish in similar environments (a type of reference condition), mitigation requirements, quantitative objectives of the program, Biological Assessment and Management Plan (BAMP) values, or other appropriate standards. An example of a statistical hypothesis would be:

Ho: Productivity (Recruits/Spawner) of the supplemented population ≥ standard productivity.

Ha: Productivity (Recruits/Spawner) of the supplemented population < standard productivity.

For these analyses to be useful, the standards must be based on biological reality.

Conclusions and Recommendations

Hatcheries are an important component of fish production within the Upper Columbia Basin. The goal of some of these programs is to supplement natural production in declining populations. The supplementation programs generally use both hatchery and natural (spawned and reared in nature from either wild or hatchery parents) adults for hatchery broodstock. These programs are designed to supplement natural populations by increasing natural reproduction while preventing the establishment of a domesticated hatchery stock. Thus, the programs should increase total spawning escapement and NORs, and not reduce the productivity of the natural population. Measuring the success of these programs is challenging and expensive.

In this paper, we described methods that can be used to determine if supplementation programs are achieving some of their goals. This paper focused on the use of reference populations to determine if the supplementation programs increase total spawning escapement, NORs, and maintain or increase productivities. In some cases, suitable reference populations may not be available (e.g., we found no suitable reference populations for Upper Columbia steelhead and sockeye). In these cases, alternative methods are needed to assess supplementation effects. We also described these alternative methods in this paper.

Identification of Reference Populations

Finding suitable reference populations that match well with supplemented populations is a difficult and time-consuming process. Our three-step selection process included identification of populations with similar life-history characteristics, few or no hatchery spawners, a long time series of accurate abundance and productivity estimates, and similar freshwater habitat impairments and out-of-basin effects. Those populations that met these criteria were then examined for their relationship with the supplemented population (in this case, the Chiwawa spring Chinook population). Several criteria were scored, including pNOS, correlation, trend, and effect size. Reference populations with total weighed scores of 81 or greater were selected as suitable reference populations.

This selection process provided a valuable framework for selecting suitable reference populations for supplemented populations. Interestingly, we found that a given reference population may match well with one parameter of the supplemented population (e.g., spawning escapement), but not for all parameters (e.g., not NORs or productivity). The reason for this may be related to errors in the estimation of population parameters and/or differential factors limiting population parameters of supplemented and reference populations. Therefore, depending on the parameter analyzed, a different suite of reference populations may be needed.

An important assumption in the use of reference populations is that the supplemented and reference populations that tracked each other before supplementation would continue to track each other in the absence of supplementation. Given that the reference populations did not match the Chiwawa population on all criteria examined, and some reference populations tracked the Chiwawa population more poorly than others, there may be some uncertainty as to whether differences observed between the supplemented and reference populations during the supplementation period are associated with the hatchery program, or other unaccounted factors. For example, any large-scale change (man-made or natural) within the reference or supplemented population could affect our ability to assess the effectiveness of the supplementation program.

To account for some of these uncontrollable factors, we recommend the use of a "causal-comparative" approach to strengthen the certainty of our inferences. This approach relies on correlative data to try and make a case for causal inference. We recommend that the following state variables be measured and tracked within the supplemented and reference populations: mean annual precipitation, total and riparian forest cover, road density, impervious surface, and alluvium. These variables can be used to describe differences in water temperatures at different life stages (pre-spawning, egg incubation, and summer rearing) and substrate characteristics, including fine sediments and embeddedness. These state variables can be used to help explain possible changes in spawner abundance, NORs, and productivity that are independent of supplementation. In addition, the use of multiple reference streams reduces the possibility that man-made changes to a single reference stream will influence the interpretation of the results.

Analyses with Reference Populations

Using reference populations, we evaluated the effects of supplementation on natural-log transformed and untransformed total spawning escapement, NORs, and productivity by comparing trends, analyzing mean differences, ratios, and rates, and comparing stock-recruitment curves and their parameters. For trend analysis, we compared the slopes of the trends between each supplemented/reference pair before and during supplementation. If the hatchery program is successfully supplementing the natural population, trends in spawner abundance and NORs should deviate significantly during the supplementation period (i.e., the slope of the supplemented population should be greater than the slopes of the reference populations during the supplementation period), but not during the pre-supplementation period. For productivity, the slope of the supplemented population, relative to the reference population, should increase or remain the same.

Because trend analysis only tests the slopes of the trend lines, it does not test for differences in elevations of the trend lines, additional analyses were needed to determine if supplementation increased spawner abundance, NORs, and productivity of the target population without changing the slopes of the trend lines. To do this, we derived three different response variables using natural-log transformed and untransformed spawner abundance, NORs, and productivity data. The first derived variable included difference scores, which were calculated as the difference between paired treatment and reference data (T-R). The second included ratios, which were calculated as the ratio of paired treatment and reference data (T/R). Finally, we calculated the differences in annual changes in paired treatment and reference population data (ΔT-ΔR). If the hatchery program is successfully supplementing the natural population, the mean difference or ratio score of paired spawner abundance data and NORs during the supplementation period should be greater than the pre-supplementation period. For productivity, the mean difference or ratio score during the supplementation period should be equal to or higher than the pre-supplementation period.

As a final set of analyses, we compared the stock-recruitment curves of the supplemented population (using stock and recruitment data only from the supplementation period) to the reference populations (using all available stock and recruitment data). Specifically, we tested whether the regression parameters were equal between the supplemented population and the reference populations, and whether the fitted curves coincided between populations. Here, we were most interested in comparing the productivity parameters in the models.

Surprisingly, these different analyses yielded similar results when they were applied to the Chiwawa spring Chinook and reference population data. Trend analysis was unable to detect a significant difference in trends between the supplemented and reference populations during the supplementation period. Even though we measured an increasing trend in spawner abundance, NORs, and productivity in the supplemented population during the supplementation period, these same parameters trended upward in the reference populations. Likewise, we were unable to detect a significant supplementation effect using difference scores, ratios, and differences in annual changes. However, we found the results from analysis of mean differences of annual change difficult to interpret and they may be insensitive to treatment effects. A simpler analysis, which is also easier to interpret, is to use trend analysis. Finally, comparing stock-recruitment curves and their parameters did not provide strong evidence that supplementation has affected the productivity of the natural population.

Based on these results, we do not recommend using difference scores of annual change (ΔT-ΔR), nor do we recommend comparing stock-recruitment curves and their parameters. As noted above, difference scores of annual change are difficult to interpret and may be redundant with trend analysis. Testing stock-recruitment curves and their parameters appears redundant with testing differences in productivity using difference scores or ratios. In addition, the analyses are computer intensive and do not appear to be very sensitive to changes.

There was little difference in results using difference scores and ratios. It appears that ratios may be more sensitive to change than difference scores (e.g., we found significant differences in some comparisons using ratios but not with difference scores), but ratios can be more difficult to interpret than difference scores. Nevertheless, we recommend the use of ratios in future analyses.

Correcting for Density Dependence and Carrying Capacity

The analyses described so far assumed that the density of spawners or recruits did not affect the survival and productivity of fish. However, without controlling for density effects, productivity of the population would continue to decline with increasing abundance. This scenario could occur in supplementation programs that increase the number of spawners, and could result in lower productivities relative to reference populations. In addition, lower productivities may be caused by differential environmental carrying capacities rather than the capacity of the supplemented fish to produce offspring. Therefore, we described two different methods for deriving density-corrected estimates of productivity. The first controlled the effects of density on productivity by partitioning observed productivities into density-independent and density-dependent productivity. These productivities were then combined in a single test. The second method corrected for differences in carrying capacities between the supplemented and reference populations. This was accomplished by calculating the percent saturation of NORs, which was estimated as the ratio of observed NORs to the maximum number of NORs that the habitat could support.

We fit Ricker, Beverton-Holt, and smooth hockey stick models to stock and recruitment data to estimate the maximum number of NORs (NORs at carrying capacity) and the maximum number of spawners needed to produce maximum NORs. We fit models to the supplemented and reference populations. Using information-theoretic criterion and evaluating the precision of estimated parameters, we found that the smooth hockey stick model provided the best estimates of maximum NORs and spawners. We used these modeled values to estimate density-independent and density-dependent productivities, and saturation of NORs.

Statistical analyses, using difference scores and ratios of adjusted Chiwawa spring Chinook productivity data, found no significant effects of supplementation on the productivity of the supplemented population. Indeed, the results from correcting for density dependence were similar to those without correcting for density dependence. This is in part because the abundance of the supplemented and reference populations has been below their respective carrying capacities in most years. This was clearly demonstrated in the analyses of NORs corrected for carrying capacity. In the supplemented population, the mean fraction of the carrying capacity filled with NORs decreased significantly during the supplementation period. In other words, the carrying capacity was filled with more NORs during the pre-supplementation period than during the supplementation period, which is contrary to the goal of supplementation. By comparison, two of the reference populations showed a similar decrease in saturation, while the other two reference populations actually increased in saturation. Analyzing the saturation scores using BACI-design analyses indicated that two of the four pairings differed significantly. That is, the percent saturation of the supplemented population decreased significantly relative to two reference populations.

Because productivity can be affected by the abundance of spawners and recruits, we recommend that future analyses comparing supplemented and reference populations adjust for density-dependent effects and differential carrying capacities. Although we detected only slight differences between adjusted and unadjusted results, as supplemented stocks recover, it will become more important to adjust productivities to account for density dependence. Importantly, the analyses using percent saturation placed NORs in the context of the carrying capacity of the environment. This will help managers determine if supplementation programs are filling or over-filling the capacity of the habitat with NORs.

As we noted earlier, analyses using productivities adjusted for density dependence assume that there is a spawner abundance at which density-independent effects end and density-dependent effects begin. In reality, density-dependent effects occur at low spawning abundance and intensify as spawning abundance increases. We did not account for these increasing density-dependent effects at lower spawner abundances. This is an area that needs additional attention.

Analyses without Reference Populations

Because of the rigorous criteria we used to select reference populations, it is likely that reference populations may not exist for making comparisons with supplemented populations. For example, we used the criteria described in this paper to identify reference populations for supplemented steelhead and sockeye populations in the Upper Columbia Basin. We were unsuccessful in identifying any suitable reference populations. Therefore, in the absence of suitable reference populations, it is important to have alternative methods for assessing supplementation effects. We described three different types of analyses one can use to assess supplementation effects in the absence of reference populations. They include before-after comparisons, correlation analysis, and comparisons to standards.

Before-after analyses compare population metrics before supplementation with those during supplementation. In this case, data collected before supplementation represent the reference condition. The assumption is that population trajectories measured during the pre-supplementation period would continue in the absence of supplementation. We compared trends in spawner abundance, NORs, and productivity before and after supplementation. In addition, we compared mean scores in these three parameters before and after supplementation. Finally, we attempted to compare stock-recruitment parameters before and after supplementation. The hypotheses examined were that the spawner abundance and NORs would be greater during the supplementation period, and that productivities would not decline during the supplementation period.

Trend analysis indicated that the all three Chiwawa spring Chinook population parameters trended downward during the pre-supplementation period, but trended upward during supplementation. On the other hand, mean spawner abundance and NORs were lower during the supplementation period than during the pre-supplementation period. Mean productivities increased, but not significantly, during the supplementation period. We were unable to compare pre- and post-supplementation stock-recruitment curves because we were unable to fit stock-recruitment models to the pre-supplementation data.

We used correlation analyses to determine if the proportion of hatchery-origin fish that spawn naturally on the spawning grounds (pHOS) increased productivity. In addition, we used correlation to assess the association between pHOS and the residuals from stock-recruitment relationships. A significant negative association provides evidence that hatchery-origin spawners may not be as productive as natural-origin spawners. The analysis indicated that the productivity of Chiwawa spring Chinook increased with increasing pHOS, but the association was not significant. In contrast, there was a negative association between pHOS and the stock-recruitment residuals, but again the association was not significant. The latter analysis accounts for density-dependent effects.

In concert, the before-after comparisons and correlation analyses do not provide conclusive evidence that the supplementation program has increased spawner abundance and NORs, or that it has significantly reduced the productivity of the supplemented population. Although increasing the number of hatchery fish on the spawning grounds appears to reduce NORs and productivity, mean productivity actually increased during the supplementation period compared to the pre-supplementation period.

It is important to note that relying on only one set of analysis could result in drawing a wrong conclusion. For example, if we had only conducted trend analysis, we may have concluded wrongly that the Chiwawa spring Chinook supplementation program significantly increased spawner abundance, NORs, and productivity in the supplemented population. The analysis of mean scores and correlations indicates that the supplementation program has not increased spawner abundance or NORs in the supplemented population. Therefore, in the absence of suitable reference populations, we recommend that analyses include the evaluation of trends, means scores, and correlations. By conducting more than one set of analyses, one can use weight-of-evidence to assess the effects of supplementation programs.

Under the scenario that there are no reference populations or pre-supplementation data, one is left with comparing population parameters to relevant standards. These standards could come from mitigation requirements, quantitative objectives, or published or unpublished standards. One could also use correlation to evaluate the association between productivity and pHOS, but this requires a wide range in pHOS values to be most effective. A more extreme approach, which probably would not gain much traction with managers, is to shutoff the supplementation program for some time and then evaluate the effects of the program in a before-after design. The Entiat spring Chinook hatchery program provides a unique opportunity to evaluate this type of management decision.

Some Concerns and Limitations

No matter how hard we try to explain different sources of variation in population data, we are limited by the quality of the data. Teasing out the effects of supplementation requires long time series of population data. Because funding levels and methods change over time, the quality (i.e., accuracy and precision) of the data also changes over time. Importantly, the population parameters examined in this paper (spawner abundance, NORs, and productivity) are rarely measured directly in the field. That is, other population metrics, such as numbers of redds, number of fish counted at weirs or dams, scales, tags, etc., are sampled in the field. These metrics are then used to calculate spawner abundance[24], NORs, and productivity, often based on assumptions about fish/redd, pre-spawning loss, marking rates, and sampling rates. This has a tendency to increase the variability in the data independent of supplementation programs. In our studies, we can only control sampling within the supplemented populations, and even that is limited by available funding. We have no control over the sampling within reference populations. Thus, we have to assume that sampling within the reference populations will continue and that sampling effort will remain comparable to that in the supplemented populations.

In our analyses, we included both the Entiat and Little Wenatchee populations as references for the Chiwawa population. In the analyses, we treated them as equivalent to the other reference populations. That is, the statistical procedures used to compare the supplemented population to each reference population were identical. This is appropriate. However, the interpretation of the results must be different when comparing the Entiat and Little Wenatchee to the supplemented population, because they are populations that were influenced by hatchery fish. As noted earlier, the Entiat spring Chinook hatchery program has been discontinued. Therefore, it provides a unique type of reference where the comparison changes from both populations being supplemented to only one population being supplemented. For the Little Wenatchee, nearly all the strays came from the Chiwawa program. Straying should stop or be greatly reduced with the change in water supply to the Chiwawa Rearing Ponds. In sum, one must be careful in how they interpret these test-reference results.

Finally, it is important to point out that for this paper, we conducted 463 statistical tests. Because we set our Type I error rate at 0.05, by random chance alone, we may have incorrectly rejected about 23 null hypotheses. Inasmuch as this work was designed to evaluate different ways to analyze test-reference data, the number of future analyses will be greatly reduced based on the results from this work. However, if the Type I error rate is a concern to managers, researchers can

---

[24] The smooth hockey stick model, which we used to estimate density-dependent correction factors for productivity and NORs, is sensitive to errors in spawner escapement estimates. Therefore, it is important to use accurate and precise estimates of spawner escapement.

use a lower error rate, such as $\alpha = 0.01$. Another option is to analyze test-reference data graphically. Although this is subjective, there are no statistical analyses and therefore no concerns with violating assumptions of statistical tests, including temporal correlation. We believe researchers should use the statistical procedures recommended in this report to support graphic analysis.

References

Burnham, K. P. and D. R. Anderson. 2002. Model selection and multimodel inference: a practical information-theoretic approach. Second edition. Springer, New York, N.Y.

Froese, R. 2008. The continuous smooth hockey stick: a newly proposed spawner-recruitment model. Journal of Applied Ichthyology 24:703-704.

Galbreath, P., C. Beasley, B. Berejikian, R. Carmichael, D. Fast, M. Ford, J. Hesse, L. McDonald, A. Murdoch, C. Peven, and D. Venditti. 2008. Recommendations for broad scale monitoring to evaluate the effects of hatchery supplementation on the fitness of natural salmon and steelhead populations. Final report of the Ad Hoc Supplementation Monitoring and Evaluation Workgroup, Portland, OR.

Gotelli, N. and A. Ellison. 2004. A primer of ecological statistics. Sinauer Associates, Inc., Sunderland, MA.

Hays, S., T. Hillman, T. Kahler, R. Klinge, R. Langshaw, B. Lenz, A. Murdoch, K. Murdoch, and C. Peven. 2006. Analytical framework for monitoring and evaluating PUD hatchery programs. Final report to the HCP Hatchery Committee, Wenatchee, WA.

Hilborn, R. and C. Walters. 1992. Quantitative fisheries stock assessment: choice, dynamics, and uncertainty. Chapman and Hall, New York, N.Y.

ISRP and ISAB. 2005. Monitoring and evaluation of supplementation projects. ISRP and ISAB Report 2005-15, Northwest Power and Conservation Council, Portland, OR.

Jorgensen, J., J. Honea, M. McClure, T. Cooney, K. Engie, and D. Holzer. 2009. Linking landscape-level change to habitat quality: an evaluation of restoration actions on the freshwater habitat of spring-run Chinook salmon. Freshwater Biology 54:1560-1575.

Murdoch, A. and C. Peven. 2005. Conceptual approach to monitoring and evaluating the Chelan County Public Utility District hatchery programs. Final report. Prepared for the Chelan PUD Habitat Conservation Plan's Hatchery Committee, Wenatchee, WA.

Noakes, D., D. Welch, and M. Stocker. 1987. A time series approach to stock-recruitment analysis: transfer function noise modeling. Natural Resource Modeling 2: 213-233.

Pearsons, T. and G. Temple. 2010. Changes to rainbow trout abundance and salmonid biomass in a Washington watershed as related to hatchery salmon supplementation. Transactions of the American Fisheries Society 139:502-520.

Quinn, T. and R. Deriso. 1999. Quantitative fish dynamics. Oxford University Press, New York, N.Y.

Ricker, W. 1954. Stock and recruitment. Journal of the Fisheries Research Board of Canada 11:559-623.

Skalski, J. R. and D. S. Robson. 1992. Techniques for wildlife investigations, design and analysis of capture data. Academic Press, New York, NY.

Smith, E. P., D. R. Orvos, and J. Cairns. 1993. Impact assessment using the before-after-control-impact (BACI) model: concerns and comments. Canadian Journal of Fisheries and Aquatic Sciences 50:627-637.

Stewart-Oaten, A., J. R. Bence, and C. W. Osenberg. 1992. Assessing effects of unreplicated perturbations: no simple solutions. Ecology 73:1396-1404.