State of California
The Natural Resources Agency
Department of Fish and Wildlife

Fish Bulletin 182

Integration of Steelhead Viability Monitoring, Recovery Plans and Fisheries Management in the Southern Coastal Area

By

David Boughton[1]

Jennifer Nelson[2]

Michael K. Lacy[3]

2022

[1] National Marine Fisheries Service, Southwest Fisheries Science Center, Santa Cruz, CA 95060

[2] California Department of Fish and Wildlife, Bay-Delta Region, Aptos, CA 95003

[3] California Department of Fish and Wildlife, Fisheries Branch, West Sacramento, CA 95605

A sea-run gravid female steelhead (65 cm) and a sexually mature resident Rainbow Trout (25 cm) over a freshly created redd in the process of spawning in Maria Ygnacio Creek, near Santa Barbara, California, March 15, 2017. Photograph by Mark H. Capelli, National Marine Fisheries Service.

SUMMARY

California's Coastal Monitoring Plan (CMP) is a design-based plan to collect statistically valid, ecologically meaningful data on the status of salmonid fishes inhabiting California's coastal watersheds. Statistical validity comes from formal development of a sampling frame and sampling scheme for stream reaches and fish. Ecological meaning comes from a conceptual basis in high-level indicators of fish stock viability: abundance, productivity, spatial structure and diversity (McElhany et al. 2000). However, in the original technical formulation by Adams et al. (2011), monitoring methods for the northern coastal area were considerably more developed than for the coastal area from the Pajaro River southward. Key impediments in the southern area stemmed from (1) the episodic flow regime characteristic of the area's river systems, (2) the sparse distribution of the salmonid *Oncorhynchus mykiss*, and (3) the need to distinguish rare anadromous forms from the more common resident form of *O. mykiss*. Here we update and expand the original vision of Adams et al. (2011) for the southern area. We formulate a closer integration of the monitoring plan with Federal recovery plans and propose other modifications to the design to make it more practical and efficient. Proposed modifications include stratifying sampling by targets of estimation identified in recovery plans (biogeographic population groups, selected "backbone" populations), conducting electrofishing surveys instead of snorkel surveys during the low-flow season, modifying the sampling frame to include "short reaches" for low-flow surveys, and incorporating an additional stage of sampling in the low-flow season to identify the proportion of habitat that is unsuitable due to lack of surface flow. In addition, we recommend flexibility for abundance monitoring, deploying redd surveys or counting stations, depending on which is best suited to field conditions of a given biogeographic area. We also recommend flexibility with respect to methods used in Life Cycle Monitoring stations, populations where smolt production, redd surveys, and adult counts are made in combination. Finally, we provide explicit indicators for diversity monitoring, including anadromous fraction, a key need for monitoring viability of steelhead populations in the southern monitoring area. We believe the modifications will allow a leaner, more information-rich monitoring scheme that is practical to implement. Implementation will require some methodological development, especially for refinement of the design for counting stations. In addition, we outline how the sampling framework can support a broader vision of combining data with life-cycle models, to learn how to establish productive fish stocks in the coming years.

CONTENTS

INTRODUCTION

The purpose of California's Coastal Monitoring Plan (CMP) is to collect statistically valid, ecologically meaningful data on the status of salmonid fishes inhabiting California's coastal watersheds (Adams et al. 2011). Collecting such data is difficult but important, because it provides objective data to close the feedback loop between management actions and the response of fish stocks, including recovery actions for threatened and endangered stocks. Such data support status review updates conducted every five years by NMFS for stocks on the Federal Endangered Species List and can also support fisheries management of both listed and unlisted stocks.

Statistical validity comes from formal development of a sampling frame and sampling scheme for stream reaches and fish, from which inference is made about specific ecological indicators (Adams et al. 2011). Ecological meaning comes from a conceptual basis in high-level indicators of fish stock viability: abundance, productivity, spatial structure and diversity, commonly referred to as "Viable Salmonid Population parameters," or simply "VSP parameters" (McElhany et al. 2000). The targets of estimation for these parameters are specific populations or groups of populations, developed in technical guidance (Boughton et al. 2007, Spence et al. 2008, Williams et al. 2008) of Federal recovery plans for coastal salmonids.

REASONS FOR AN UPDATE

In the original technical formulation of the CMP, Adams et al. (2011) developed a detailed strategy, design, and methodology for the northern coastal region of California.[1] Scientific work on salmonids in the northern region (Gallagher et al. 2010) had produced a series of workable yet rigorous methods for implementation, especially for Coho Salmon. Although Adams et al. (2011) also dealt with the southern coastal region,[2] the approach there was less developed, even though only one species of salmonid, steelhead (*Oncorhynchus mykiss*) is native to the region. In the southern region, scientific uncertainty about steelhead ecology, low and patchy abundances of steelhead, and arid ecoregions where streams exhibit an episodic flow regime, meant that a workable and rigorous monitoring strategy could not yet be fully identified. Instead, Adams et al. (2011) identified a general approach using counting stations suitable for estimating adult abundance, and for other VSP parameters outlined a system of low-flow snorkel surveys and life-cycle monitoring stations similar to the northern monitoring area. To date, these recommendations have not been implemented broadly in the southern monitoring area.

1 Coastal stream networks from Aptos Creek in Monterey Bay, north to the border with Oregon, including tributaries of San Francisco Bay west of Carquinez Strait.

2 Coastal stream networks from the Pajaro River (inclusive) in Monterey Bay, south to the border with Mexico.

Here we update this outline into a more detailed strategy, design, and methodology for the southern monitoring area. We have three objectives: (1) promote closer integration of the monitoring plan with specific metrics and targets of estimation identified during recovery planning (Boughton et al. 2006, 2007); (2) facilitate greater methodological flexibility; and (3) foster closer integration of viability monitoring and management in order to facilitate recovery and open a path to delisting.

We outline how these updates support the original vision of Adams et al. (2011), but are more practical, meaningful, and operationally efficient. Practicality includes considerations of performance of field methods within the challenging environmental conditions of the southern monitoring area. Meaningfulness stems from direct integration with specific VSP indicators identified in technical guidance (Boughton et al. 2007), especially regarding diversity. Operational efficiency includes strategies to reduce the level of effort or resources needed to make high-quality estimates of the indicators. New methods are proposed when they seem practical to develop and are either necessary or would produce large gains in operational efficiency. Going forward, many specifics about methods and implementation will still need to be refined by the Science Team for the Coastal Monitoring Plan, in consultation with regional staff of California Department of Fish Wildlife and external partners.

SOUTH COAST RIVERS AND STREAMS

A key motivation for an update is to tailor monitoring to the distinct ecoregions of the southern area, where climate is more arid than the north coast. Upland vegetation in the southern monitoring area is mostly oak woodlands, grassland, and chaparral (Figure 1), contrasting with the heavily forested watersheds to the north. Runoff is more immediate and stream flows are highly episodic (Figure 2), frequently reworking channels and transporting large bed loads of sand and gravel after storms (Kondolf et al. 2013, Harrison et al. 2017, Harrison et al. 2018). During the hot, dry summer, many streams are prone to seasonal drying (Figure 3), though creeks in high-rainfall areas are generally perennial and surprisingly cool in summer (Figure 4). Bedrock outcrops can force large deep pools that provide refugia for fish in otherwise inhospitable conditions (Figure 5).

FIGURE 1. An arid chaparral watershed in Ventura County in late summer. The line of green alders marks the course of Lion Creek, a tributary to Sespe Creek that maintains cool surface flow during the summer and supports numerous juvenile *O. mykiss.*

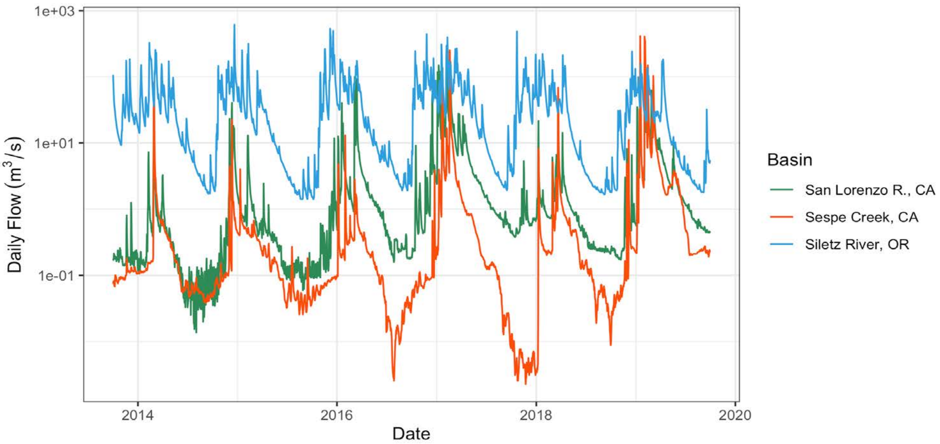

FIGURE 2. Recent "episodic" stream flows in unregulated Sespe Creek (Ventura County) compared to streams draining more humid watersheds of comparable area to the north. Flows in Sespe Creek annually vary by 4 to 5 orders of magnitude, stay only briefly at high flows during the wet season, and reliably decline to very low flows in most summers. San Lorenzo River is on the central California coast in Santa Cruz County, and Siletz River is on the central Oregon coast.

FIGURE 3. An intermittent reach in Horse Creek, tributary to Arroyo Seco in Monterey County, just prior to seasonal loss of all surface flow (D. Boughton 06/30/2006). Juvenile *O. mykiss* are commonly observed in such reaches prior to drying or in residual pools after drying (David Boughton, National Marine Fisheries Service, June 30, 2006).

FIGURE 4. A perennial run in the heavily shaded Tassajara Creek, draining a chaparral subbasin within the Santa Lucia Mountain Range (David Boughton, National Marine Fisheries Service, June 12, 2006). The watershed of Tassajara Creek receives significantly more winter rainfall than the nearby Horse Creek pictured in Figure 3.

FIGURE 5. A large bedrock force-pool on the Santa Ynez River at Red Rocks (Mark H. Capelli, National Marine Fisheries Service, May 19, 2002). The bedrock outcropping on one side forces a constriction at high flows, which in turn increases bed scour relative to unforced reaches. The result is a deep excavated pool that is resilient to loss of surface water during low-flow season. Such pools often provide refugia for juvenile *O. mykiss*, adult Rainbow Trout, and other aquatic species, but are prone to heating up in the sun.

Juvenile steelhead in the area commonly use habitats where continuous surface flow is lost in the summer, including intermittent streams with residual pools (Boughton et al. 2009) and estuaries that form seasonal lagoons (Bond et al. 2008). It is important that a monitoring strategy include these types of habitat (see terminology in Table 1). A somewhat surprising feature of the region is that the larger mainstem rivers are often intermittent and sensitive to drought, whereas small upland creeks that drain into them are often drought-resistant perennial streams (Boughton et al. 2009, Booth et al. 2013). This pattern is widespread and attributable to basin-level controls on water and sediment dynamics (Montgomery and Buffington 1997, Buffington et al. 2004, May and Lee 2004).

TABLE 1. Flow terminology for freshwater streams[1] and estuaries[2] in the southern monitoring area.

Term	Definition	Usage by *O. mykiss*
Streams		
Perennial	Reach maintains continuous surface flow year-round during a typical year	Extensive use
Intermittent	Reach loses continuous surface flow in dry periods. May have sections of surface flow or residual pools	Extensive use
Ephemeral	Surface flow only in direct response to precipitation	Not generally used
Estuaries		
Lagoon	Seasonal closure from the sea by wave-built sand spits that periodically breach (most estuaries in the area)	Extensive use
Bar-built	A geologically structured sand bar creates a semi-enclosed bay (Elkhorn Slough, Morro Bay)	Extensive use

[1] from Nadeau and Rains (2007).

[2] from Emmett et al. (2000). Although "lagoon" and "bar-built" are often used interchangeably to describe California's many estuaries that become closed off by sand bars during the dry season, Emmett et al. (2000) draw a distinction between the two types, with lagoons describing estuaries that close off seasonally--the vast majority on the California coast--and bar-built referring to estuaries separated from the ocean by substantial sandbars that are stable at geologic time scales, but that often maintain a continuous connection at a gap, such as Morro Bay or San Diego Bay. Rarer types are drowned river valleys (San Francisco Bay) and tectonically formed estuaries (Tomales Bay).

Adams et al. (2011) outlined distinct northern and southern monitoring areas because the preferred method for monitoring abundance in the northern area was not suitable for much of the southern area. Methods used during the steelhead migration season (late November through May) must be compatible with the large, turbid, episodic stream flows typical of this season, while methods used during the low-flow season must accommodate the warm conditions, extremely shallow streams, and widespread drying of stream channels from June through October. Methods must also be suitable for application in remote wilderness areas and accommodate highly dynamic channel behavior and large bed-loads.

At the same time, the distinction between the southern and northern monitoring areas is more of emphasis than kind, with chaparral and grassland found inland in the northern monitoring area, and pockets of redwood and other coniferous forest occurring in the southern monitoring area in suitable microclimates. Thus, we consider that methods used in the northern monitoring area may be suitable for some parts of the southern area.

STEELHEAD IN THE SOUTHERN MONITORING AREA

The only native salmonid in the southern area is *Oncorhynchus mykiss*, though strays from other species occasionally enter rivers (Skiles et al. 2013), and non-native brown trout (*Salmo trutta*) have been established in some systems. Populations of *O. mykiss* typically express both freshwater-resident and anadromous forms (Boughton et al. 2006), traditionally known as Rainbow Trout and steelhead, respectively (Page et al. 2013). Only the anadromous form is considered to be at risk of extinction, due to its extreme rarity and inconsistent expression in populations of *O. mykiss* in the southern coastal area.

TABLE 2. ESA-listed salmonids in the southern monitoring area.

Distinct Population Segment (DPS)	Freshwater Geographic Range	Status	Recovery Plan
South-central California coast steelhead	Pajaro River south to but not including the Santa Maria River	Threatened	NMFS (2013)
Southern California steelhead	Santa Maria River to Tijuana River	Endangered	NMFS (2012)

Federal recovery plans identified two Distinct Population Segments (DPSs) of steelhead in the southern area (Table 2). Each DPS—a legal term for the entity being protected under the Endangered Species Act—comprises only the anadromous portion of a corresponding ESU—a biological term. ESU, or Evolutionarily Significant Unit, is a biological concept developed by Waples (1991) to describe groups of populations within a species that share a common evolutionary legacy and are substantially reproductively isolated from other such groups. For *O. mykiss*, ESUs are typically composed of both anadromous and freshwater-resident forms, because in any given geographic region the two forms generally can and do interbreed, and also share an evolutionary legacy (Clemento et al. 2009, Pearse et al. 2014). However, only the anadromous component is listed as a protected DPS under the US Endangered Species Act. To provide some clarity to this potentially confusing situation, we recommend adopting the terminological conventions in Table 3.[3]

[3] We depart slightly from Page et al. (2013), who use Rainbow Trout as the common name for the entire species, while we confine it to populations or ESUs composed solely of individual Rainbow Trout.

TABLE 3. Terminological conventions recommended for *Oncorhynchus mykiss*.

Term	Convention
Oncorhynchus mykiss	A species of Pacific salmonid, composed of both anadromous and freshwater-resident individuals that all spawn in freshwater creeks and rivers.
Steelhead	Individuals: *O. mykiss* that are anadromous (individuals that migrate to and spend one or more seasons in the ocean), here used to mean adult steelhead. Populations/ESUs: Contain steelhead individuals and possibly Rainbow Trout individuals.
Rainbow Trout	Individuals: *O. mykiss* individuals that are freshwater-resident (individuals that complete their life-cycle in freshwater), here used to mean adult Rainbow Trout. Populations/ESUs: Composed solely of Rainbow Trout individuals.
Juvenile *O. mykiss*	Immature fish whose fate as steelhead or Rainbow Trout cannot yet be established.
Anadromous waters	Stream reaches that are potentially accessible to migrating adult steelhead (not blocked by complete natural barriers such as waterfalls nor complete anthropogenic barriers such as dams). *O. mykiss* in anadromous waters are not necessarily anadromous themselves.
ESU	An evolutionarily significant unit (see Waples 1991). In *O. mykiss*, frequently composed of both steelhead and Rainbow Trout.
DPS	A distinct population segment of a species listed as threatened or endangered under the US Endangered Species Act. In the southern monitoring region, only the steelhead component of each *O. mykiss* ESU is listed as a threatened or endangered DPS.

The natural structure of each ESU can be conceived as hierarchical, comprising biogeographic groupings of ecologically similar populations, in turn comprising demographically distinct populations, which in turn are composed of mixtures of interacting steelhead and Rainbow Trout. Biogeographic population groups (BPGs) developed in technical guidance (Boughton et al. 2007) assumed that gradients in broad-scale ecological controls (climate, topography, ecosystem) lay mostly along two axes: north-south, and coastal-inland (Figure 6). Each distinct coastal basin was assumed to support a single demographically distinct population of steelhead, except in the large Salinas River basin where three populations were thought to be distributed among Gabilan Creek, Arroyo Seco, and a set of southern Salinas tributaries.

Figure 6. Distribution of biogeographic population groups (BPGs) for steelhead in the southern coastal area, developed in the technical guidance for recovery plans (Boughton et al. 2007).

The coastal population groups live in streams along the ocean-facing slopes of the coastal mountains, which tend to be relatively cool in summer, wet in winter, and confined to a narrow zone directly influenced by the maritime climate along the coast. The inland population groups live in streams typically draining larger inland valley systems, of much larger geographic extent but where cool, wet conditions are typically confined to mountainous areas or the fog zone near the coast. The mountain refugia for the species tend to be separated from the coast by long rivers with intermittent surface flow (Boughton et al. 2006). In the far south, even the mountains provide only widely scattered refugia for the species (Abadia-Cardoso et al. 2016). Any effective monitoring strategy must successfully accommodate these features.

The long-term expression of the steelhead life history appears to depend on Rainbow Trout being present at some level in each population. This dependency arises from the tendency described earlier for perennial creeks to be separated from the ocean by intermittent rivers sensitive to drought. In dry years the lack of surface flow in migration corridors may prevent adult steelhead from migrating inland to spawn, while perennial flows in upland reaches continue to support reproducing populations of Rainbow Trout (e.g., Carmel River in 2014; MPWMD 2015). When wet conditions return, it appears that Rainbow Trout populations can regenerate anadromous runs of steelhead. For example, the Rainbow Trout in Santa Clara River produced smolts for decades, despite a diversion dam blocking steelhead from returning to spawn (Boughton et al. 2006). Hydrologic limits to anadromy are most severe southeast of Point Conception but occur throughout the southern monitoring area. In the northern coastal area, such ecological dependence of steelhead on Rainbow Trout has not been established, but in the southern area it is difficult to ignore. Thus the viability indicators identified in technical guidance (Boughton et al. 2007) involved Rainbow Trout as well as the steelhead, even though only the latter is given formal protection under the Federal Endangered Species Act.

Recent work has improved our understanding of the genetic architecture underlying these mixed coastal populations of steelhead and Rainbow Trout. Analyses of neutral genetic variation has shown that steelhead populations are genetically distinct from each other, with a historic pattern of increasing distinctness with geographic distance that has been somewhat disrupted in the modern age by the construction of dams and other barriers to migration (Clemento et al. 2009). Sometime in evolutionary history, a large section of chromosome 5 appears to have undergone a large inversion, a rearrangement in which a segment of a chromosome is reversed end to end (Pearse et al. 2019). Currently, all coastal populations of *O. mykiss* in California so far examined contain both the ancestral form and the recent inverted form of chromosome 5 (Pearse et al. 2014). There are a number of interesting features of this chromosomal rearrangement.

First, the inversion appears to prevent genetic recombination between the ancestral and recent versions of chromosome 5 (hereafter called the A and R haplotypes of Omy5). Recalling that each fish has two Omy5 chromosomes, it appears that genetic recombination can occur in the genesis of RR fish and of AA fish, but not in the genesis of AR fish (Pearse et al. 2019)—the set of genes encompassed by the inversion thus sort together during reproduction, passing down to offspring as a single "supergene" distinct from the ancestral supergene. Moreover, many of the genes in this inverted section of chromosome are associated with circadian rhythms, sensitivity to photosensory cues, the timing of age at maturity, and other traits associated with life-history variation (Pearse et al. 2019). These two features allow the A and R haplotypes to adaptively diverge in response to selection on life-history, while still being maintained together in the same population of *O. mykiss* (Pearse 2016).

Second, the two kinds of supergenes do, in fact, appear to be associated with different expression of life-history. For example, Pearse et al. (2019) found that in a small steelhead population in Big Sur, juvenile females with the AA and AR genotypes were much more likely to migrate to the ocean than females with the RR genotype. Juvenile males with the AA and RR genotypes were similar to the females, but the male AR genotype was much less likely to migrate than the female AR genotype. This last observation is consistent with adaptive evolution of contrasting life-history strategies in males and females: female fitness is more associated with large body size than is male fitness, because of the energetic demands of manufacturing eggs versus sperm. Thus, females should be more likely than males to pursue anadromy because *O. mykiss* can generally achieve larger size at maturity in the ocean than in freshwater, and this provides more of a fitness benefit to females than to males (Pearse et al. 2019). In an independent study in the South Fork of the Eel River on the north coast, Kelson et al. (2019) made similar observations, finding that the expression of the downstream-migrant phenotype was associated both with being female and with having the A haplotype. In their smaller sample they did not detect a difference in the migration rate of AR females versus AR males, but they did find that in general the migration frequency of the AR genotype was intermediate between the RR and AA genotypes. Thus, even when the A haplotype is rare in a population, so that AA individuals are unlikely to occur, anadromy is still visible to natural selection due to its partial expression in AR individuals; and likewise, for freshwater-residency and the R haplotype.

Third, the regional distribution of the two Omy5 haplotypes across coastal populations is consistent with their link to migratory phenotypes. Throughout the California coast, subpopulations above and below dams are generally each other's closest relatives when viewed from the perspective of neutral genetic variation, but are highly divergent in their frequencies of the A and R supergenes—the A haplotype is relatively common below dams, where fish have migratory access to the ocean, and R is usually more common above dams, where anadromous migrants cannot return to reproduce (Clemento et al. 2009, Pearse et al. 2014, Pearse et al. 2019).

For example, Apgar et al. (2017) examined haplotype frequencies in 39 steelhead populations in coastal California, and found that frequency of the A haplotype at a sample site was associated with the site's degree of impact from migration barriers. Relative to similar sites without migration barriers, the frequency of the anadromous haplotype was most strongly affected by sites with complete barriers to anadromy that were longstanding (naturally occurring, such as waterfalls; -31% effect when present). The next strongest effect was of complete barriers that were more recently imposed (anthropogenic barriers; -18% effect when present), followed by recent partial barriers (-2% per barrier), with the weakest effect from longstanding (natural) partial barriers (-0.5% per barrier). In addition, the migration distance itself (river kilometers between the sample site and the ocean) had a negative effect on frequency of the anadromous haplotype. Overall, these five predictors explained 75% of the variation in haplotype frequency across populations (R^2=0.75). In Bay-Area populations as well, haplotype frequencies showed substantial evolutionary differences between the groups of fish above and below dams, despite the groups being each other's closest relatives (Leitwein et al. 2017). At a set of nine reservoirs, the A haplotype was significantly more frequent in the group below the dam (71% versus 50%, $p < 0.05$), but more variable above the dam, where it was associated with the volume of the reservoir impounded by the dam (R^2=0.69, $p < 0.01$). This last observation suggests that the A haplotype can be maintained not only by access to the ocean, but also by access to a large reservoir with capacity to support a migratory phenotype (sometimes called an "adfluvial" life-history).

Finally, although the A and R haplotypes are forms of adaptive genetic variation linked to anadromy and residency, respectively, they probably do not capture all the genetic variation associated with heritability of life-history (Pearse 2016, Kelson et al. 2019). Moreover, the Omy5 haplotypes may also contain adaptive variation associated with other traits such as growth and maturation timing (O'Malley et al. 2003, Nichols et al. 2008), and environmental factors such as realized growth also play a role in life-history trajectories (Satterthwaite et al. 2009, Ohms et al. 2014, Kendall et al. 2015). Indeed, the mean size at which fish initiate downstream migration—that is, the way life-history responds to environmental factors such as food availability—is itself subject to natural selection (Phillis et al. 2016). So, while there is a link between frequency of the A haplotype in a population and its expression of anadromy, numerous other genetic and environmental factors also play a role in expression of anadromy.

The ubiquitous co-occurrence of the anadromous and resident forms poses special challenges to monitoring: Adult Rainbow Trout and steelhead can usually be distinguished from one another visually by size, morphology, and coloration, but immature juveniles generally cannot, prior to maturation or smolting (transformation into a salt-water tolerant migrant). Nor do the Omy5 haplotypes of juvenile fish unambiguously indicate the life-history it subsequently pursues.

That said, the frequency of the A haplotype in the juvenile fish of a population does seem to be a lagging indicator for the past successful expression of the anadromous life-history in a population. Indeed, Apgar et al. (2017) suggested that the recovery potential of a population was indicated by the difference between the measured frequency of the A haplotype, and the frequency that would be expected based on predictive natural factors such as migration distance and occurrence of natural barriers. Thus they viewed haplotype frequency as an indicator for both the recent past expression of anadromy, and the future potential expression of anadromy. Pearse (2016) remarked that even though life-history of individual fish cannot be inferred from the haplotypes and their constituent genetic alleles, in general "population-level inference based on the frequencies of specific ...[adaptive alleles] could potentially be used to identify populations in which a particular trait [such as anadromy] is favoured" (see also Funk et al. 2012). In short, the frequency of the A haplotype appears to be a useful broad-scale indicator for the degree to which the anadromous life-history has been more favored by natural selection in the recent past. Because of the probabilistic association and the intermediate expression by the heterozygote AR genotype, the indicator would be expected to change gradually, integrating selective effects over multiple generations of the fish (Apgar et al. 2017). Thus, it seems likely to be a much less noisy indicator for anadromy than annual counts of adult steelhead, which tend to fluctuate greatly from year to year.

VIABILITY INDICATORS

Conceptually, the VSP parameters of abundance, productivity, spatial structure and diversity can be validly estimated at any level of the hierarchy—ESU, BPG, or individual population (McElhany et al. 2000)—though practically speaking it will seldom be feasible to monitor all parameters sufficiently to make estimates at all levels, and for a given parameter, different indicators may make sense for different levels of the hierarchy. Below as we work through the various design questions for comprehensive monitoring, we make a distinction between routine monitoring—data and indicators updated annually as part of baseline monitoring of status and trends—and ad hoc monitoring—additional data collected within a specific BPG or population over a limited time period, to address a specific management question.

For routine monitoring, Federal recovery plans (NMFS 2012, 2013) and technical guidance (Boughton et al. 2007) identified specific biological targets and specific viability indicators, which we briefly review below. Targets of estimation refer to the biological entity—population, BPG, or entire ESU—for which a particular indicator is estimated. Viability indicators, in turn, are the specific quantity being estimated for the target.

TARGETS OF ESTIMATION

Following Shaffer et al. (2002), the recovery plans specified criteria for diversity and spatial structure at the ESU level using three concepts: representation, redundancy, and resiliency. This framework provides a basis for selecting specific populations and population groups to be targets of estimation, which in turn dictates where and how much sampling is needed to estimate those targets.

Representation is the idea that species have adapted to a diverse set of habitats and prey throughout their geographic range, and so successful species protection needs to represent this diversity in a set of protected, viable populations. Redundancy is the idea that multiple redundant populations should represent each component of habitat diversity, to ensure that no catastrophe or bad year simultaneously extirpates all populations in a particular type of habitat. Representation and redundancy for steelhead was addressed by recommending multiple "backbone" populations be targeted for viability within each BPG, with the level of redundancy (number of backbone populations per BPG) specified to protect against an extreme wildfire scenario (see Table 4, middle column). Viability monitoring of the ESU as a whole thus depends on monitoring relevant VSP criteria in each of these backbone populations.

TABLE 4. Minimum number of backbone populations per BPG recommended for monitoring at the population level[1]

Biogeographic Population Group	Population-level sampling needed for:	Achieved if monitoring:
Interior Coast Range	4 populations	All populations in BPG
Carmel Basin	1 population	All populations in BPG
Big Sur Coast	3 populations	Core 1 populations[2]
San Luis Obispo Terrace	5 populations	Core 1 populations[2]
Monte Arido Highlands	4 populations	All populations in BPG
Conception Coast	3 populations	Core 1 populations[2]
Mojave Rim	3 populations	All populations in BPG
Santa Monica Mtns	3 populations	Core 1 plus Core 2[2]
Santa Catalina Gulf Coast	8 populations	Core 1 plus 4 more populations [2]

[1] Supports assessment of representation and redundancy as defined in Boughton et al. (2007), their Table 6

[2] Core 1 and Core 2 populations as specified in recovery plans, reproduced here in Appendix A.

Resiliency is the idea that a viable population should be sufficiently large and diverse to persist in the face of normal environmental variation such as fluctuations in rainfall and marine conditions. The scientific guidance sought to ensure resiliency by recommending that backbone populations be located in watersheds with larger productive capacity, and by protecting the expression of three general life-history forms: the resident form, the anadromous form, and a third "lagoon-anadromous" form in which juveniles rear in estuary/lagoon habitat prior to smolting (Bond et al. 2008, Hayes et al. 2008). In addition, technical guidance identified the relevant VSP indicators to be adult abundance and anadromous fraction, the latter defined as the proportion of adults that are steelhead. Finally, it was recommended that populations designated as part of the backbone be located in watersheds possessing drought refugia.

Thus population-level monitoring is not necessary for the entire ESU, only for a selected set of backbone populations. In the Interior Coast Range, Carmel Basin, Monte Arido Highlands, and Mojave Rim BPGs, the number of backbone populations "saturates" the BPG, constituting the entire BPG. In these cases, all populations must be monitored (Table 4). For other BPGs there is flexibility. A starting point is suggested by the Federal recovery plans (NMFS 2012, 2013), which established a "Core 1/2/3" scheme for prioritizing recovery actions (see Appendix A for a list). Selecting Core 1 populations for the monitoring backbone is sufficient from the Conception Coast northward, while additional populations are needed for the two southern-most BPGs (Table 4). In this paper we provisionally adopt this scheme, but suggest adjustments be considered during implementation: Core 1 populations are defined as those with high priority for recovery actions, but this may omit populations that are currently well protected and productive, and thus more logical candidates for the backbone. Once a final set of backbone populations is identified for monitoring, those populations should be consistently monitored over the long term.

Intensive monitoring of backbone populations is necessary but not sufficient for viability monitoring. Viability depends on maintaining spatial structure, productivity, and diversity more broadly within each biogeographic population group (Boughton et al. 2007). Thus, we recommend the indicators for these VSP parameters have targets of estimation at the level of population groups (Table 5). This provides comprehensive coverage for the southern area but allows less-intensive sampling than in the backbone populations.

TABLE 5. Viability indicators and Targets of Estimation for steelhead in the southern area.

Viability Indicator	Estimation Target	VSP Parameters[1]
Adult Abundance (steelhead + Rainbow Trout)	Backbone Populations[2]	Abundance, Productivity
Anadromous Fraction	Backbone Populations[2]	Diversity, Productivity
Juvenile Distribution	BPGs	Spatial Structure
Population Density	BPGs	Productivity, Spatial Structure
Ocean Conditions	BPGs	Productivity
Freshwater Conditions	BPGs	Productivity
Genetic Diversity[3]	BPGs	Diversity
Biogeographic Diversity	None[4]	Diversity, Spatial Structure
Life-History Diversity (3 life-history types)	BPGs	Diversity
Drought Refugia	BPGs	Productivity, Spatial Structure

[1] Broad-scale spatial structure is established by designation of backbone populations and BPGs.

[2] See Table 4.

[3] Not included in Boughton et al. (2007) but recommended here based on new information in Pearse et al. (2014) and Apgar et al. (2017).

[4] Established by designation of backbone populations in recovery plans.

ABUNDANCE

In salmonid recovery plans, abundance was defined as the annual number of spawning adults (Boughton et al. 2007, Spence et al. 2008, Williams et al. 2008), but only in the southern area did it explicitly include both steelhead and Rainbow Trout (Table 5), due to the heightened interdependence of the two forms. Status reviews for each ESU compare the 20-year mean of abundance to viability criteria defined for backbone populations. This ensures that the indicator spans periods of both strong and poor marine survival of steelhead.

PRODUCTIVITY

Productivity refers to two related aspects of viability—the long-term trend in abundance, and the tendency for abundance to recover after short-term disturbances. The importance of monitoring long-term trend was discussed by Adams et al. (2011). As in the northern area, data collected for abundance can also be used to estimate trends over time. Trends in anadromous fraction (see below) are also an important aspect of productivity and can be estimated from data collected for abundance, provided the data distinguish steelhead from Rainbow Trout.

For the second aspect of productivity—sometimes called resiliency—Table 5 specifies four additional indicators: population density, ocean conditions, freshwater conditions, and existence of drought refugia.

Population density is defined as fish per unit area of stream or estuaries and is an indicator of whether the species is saturating available habitat. In recovery plans for the northern area, criteria for adult population density were incorporated into abundance targets for specific populations (Spence et al. 2008). For the southern area, Boughton et al. (2007) identified population density as an indicator that required further research. Arriaza et al. (2017) found in the Carmel River that the relationship between spawner abundance and juvenile abundance the following fall was strongly density-dependent, suggesting the summer low-flow season imposes a bottleneck on freshwater density. We therefore conclude that juvenile density in the low-flow season is the most relevant indicator of habitat saturation. Low-flow juvenile density should be routinely monitored with BPGs as the target of estimation.

The indicator for ocean conditions is survival during the smolt-to-spawner phase of life-history, as in the northern monitoring area. Ocean survival can be estimated using a system of life-cycle monitoring stations, as recommended by Adams et al. (2011). Similarly, freshwater condition and juvenile rearing capacity can be monitored via the complementary indicator of spawner-to-smolt ratio. Additional indicators of freshwater condition are juvenile size distributions and growth rates, which correlate with smolting rate and future marine survival (Bond 2006, Satterthwaite et al. 2009, Satterthwaite et al. 2012, Arriaza et al. 2017).

Drought refugia were highlighted in the technical guidance as an important factor allowing populations to recover rapidly from drought. Here we define them simply as stream reaches that maintain surface flow and temperatures suitable for *O. mykiss* during even the driest conditions.

DIVERSITY

Diversity indicators identified in the recovery plans were anadromous fraction, biogeographic diversity, and life-history diversity. Adams et al. (2011) also emphasized ad hoc studies of genetic diversity. As described earlier, more recent findings on the genetic architecture of anadromy (Pearse et al. 2014, Apgar et al.

2017, Leitwein et al. 2017, Kelson et al. 2019, Pearse et al. 2019) support routine genetic monitoring of Omy5 haplotypes as a lagging indicator for trends in the expression of anadromy. The collection of tissue samples for haplotype monitoring can also be curated and stored in a permanent collection to support additional ad hoc genetic analysis as warranted. Such analyses include changes in other allele frequencies over time, estimates of effective population size, diversity as measured from neutral alleles, inbreeding, and others.

Anadromous fraction is simply the annual proportion of adults from each backbone population that are steelhead. Thus, anadromous fraction can be estimated using the data collected for abundance, provided the data distinguish Rainbow Trout from steelhead. As noted in Adams et al. (2011), this is more complicated than one might initially think, and requires development of novel methodology. However, the estimation of anadromous fraction and its trend over time is absolutely critical since steelhead are the component of the ESU that is on the Federal endangered species list, their adult numbers are thought to be extremely low, and improvement in their status is the focus of recovery efforts.

Biogeographic diversity does not require explicit monitoring, because sufficient biogeographic diversity was established through the designation of backbone populations (Williams et al. 2016).

Life-history diversity was defined in recovery plans as the persistence of each of three life-history types: freshwater-resident, fluvial-anadromous, and lagoon-anadromous, which we recommend for routine monitoring. Under life-history diversity, Adams et al. (2011) also listed fecundity, sex ratio, age and size structure, habitat utilization patterns, emigration age and timing, maturity timing, adult spawning timing, and other traits of populations. While data on these traits may be useful for specific management goals, they are not identified as viability criteria in the recovery plans so we do not propose they be part of routine monitoring. However, these and other characteristics may be worthwhile components of ad hoc monitoring for specific management questions.

SPATIAL STRUCTURE

Spatial structure was partly addressed in the recovery plans by designation of BPGs and backbone populations. Data collection for only backbone populations (Table 5), however, would omit contributions of the broader set of populations to spatial structure. In contrast, Adams et al. (2011) recommended that spatial structure be monitored in the southern area via estimates of juvenile occurrence at broad spatial scales. We recommend this latter approach be implemented for routine monitoring, with BPGs as targets of estimation.

This provides estimates for each BPG, establishing a comprehensive yet practical level of monitoring status and trends for each ESU. However, when spatial structure is poor or downward trending this level of resolution will likely be insufficient to uncover underlying causes. McElhany et al. (2000) conceptualized spatial structure foremost as an indicator of viability at the level of local populations, emphasizing that within-population processes—such as redundancy in local breeding groups and movement of fish between them—were the essence of good spatial structure. The sampling effort required to estimate spatial structure at this level of resolution is likely quite large. We treat it as prohibitive for comprehensive routine monitoring, but desirable for ad hoc monitoring of specific populations to improve management strategies.

MONITORING PLAN ORGANIZATION

Most of the indicators described above can be routinely monitored during the low-flow season, but adult abundance and anadromous fraction require data collection during the winter migration season, typically late November through May. Marine and freshwater condition in addition require estimates of smolt production from March through May and possibly June. Figure 7 illustrates the general strategy for collecting these data, also including a capability for ad hoc studies for research or management, using a similar statistical framework and indicators. In Table 6 we define some important statistical terms we use throughout the bulletin.

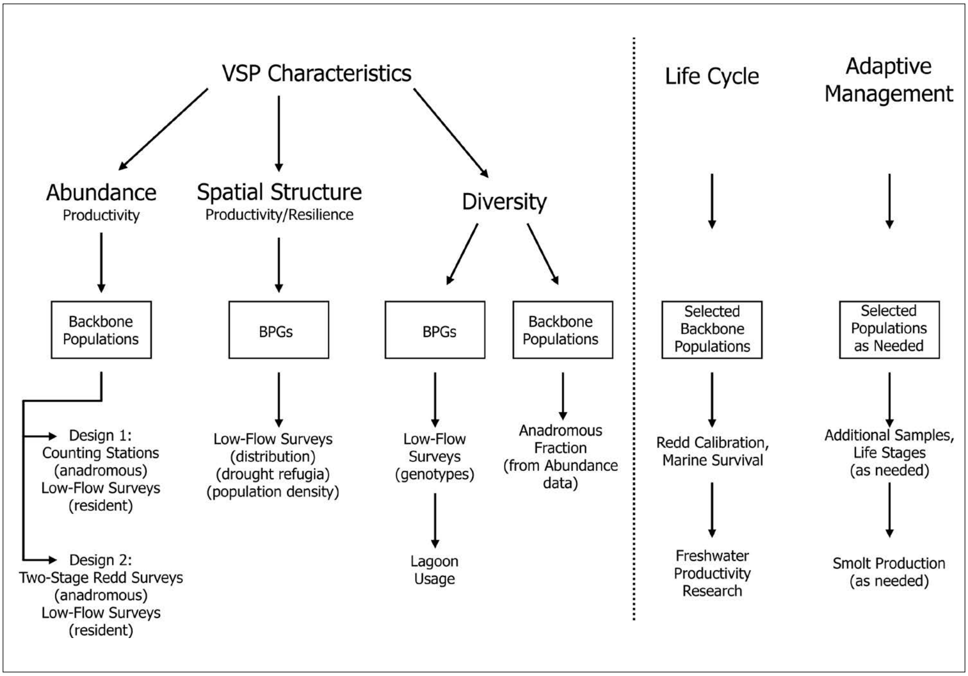

Figure 7. Organization of the monitoring plan for the southern area; modified from Adams et al. (2011). Boxes show targets of estimation for different indicators.

Low-flow surveys, a generalization of the snorkel surveys described by Adams et al. (2011), are used to collect data on a variety of indicators. Adams et al. (2011) focused solely on juvenile distribution, which is a potent indicator of spatial structure, but is not terribly informative for the southern area, where juvenile *O. mykiss* are widely distributed but cannot be resolved into the common Rainbow Trout form and the rare steelhead form. Additional indicators in Table 5 that can provide a richer dataset from low-flow surveys include juvenile density, distribution of drought refugia, juvenile use of sloughs and estuaries for rearing, and the proportion of juvenile fish with the anadromous Omy5 genetic haplotype. In general, targets of estimation for these indicators are BPGs, requiring less-intensive sampling than when the target is at the population level. When desired, sampling can be intensified for specific populations to support ad hoc studies.

Abundance and anadromous fraction, in contrast, are more intensively sampled to produce population-level estimates, but only in backbone populations. Smolt production, which in combination with adult abundance supports estimation of marine and freshwater condition, is confined to a small number of life-cycle monitoring stations that have a similar role as in the northern area. For operational efficiency we recommend these stations be situated in backbone populations, which are already foci of data collection for adult abundance. Life-cycle monitoring stations are also expected to attract research on freshwater and estuarine productivity, especially on factors limiting the productivity of the anadromous form of the species.

For practicality and operational efficiency, we recommend flexibility in the use of field methods. Adams et al. (2011) emphasized uniformity of methods, under the assumption that all methods produce some unknown level of bias in estimates and estimates from different years or populations are more comparable if they have similar bias due to common methodology. While we agree with this reasoning, we note the assumption is not always realistic. Variation in site conditions can produce heterogeneous biases (e.g., in large rivers versus small creeks), or more seriously, prevent data from being collected at all in certain watersheds or types of water year.

In trend assessment, the most important comparison is among years rather than among places. Therefore, we recommend the CDFW Science Team and regional staff select different methods as needed to tailor sampling to specific site conditions, but then stick to a given method in a given stream basin to support valid trend estimation over time. To promote consistency across basins, however, the flexibility should be limited to a short list of vetted or promising methods, with standard operating procedures for a given method. In this bulletin we consider a preliminary short-list, with redd surveys and various types of counting stations as alternatives for abundance monitoring, and electrofishing and snorkel surveys as alternatives for low-flow surveys. Later in the update we summarize pros and cons of these alternative methods and also discuss emerging methods such as eDNA, drone surveys, or remote sensing of habitat.

TABLE 6. Some statistical terms on the representativeness of an indicator.

Statistical Term	Plain-Language Description[1]
Estimator	A formula or mathematical procedure used to produce an estimate from data
Error	Random differences between an estimate and true value over many samples
Bias	Systematic differences between an estimate and true value over many samples
Variance	Magnitude of expected error
Precision	The inverse of variance (smallness of expected error)
Robustness	Ability of an estimator to produce unbiased estimates and error terms even when underlying assumptions of the estimator are violated
Parsimonious	Involving fewer assumptions or simpler reasoning

[1] For more technical definitions of these terms consult standard statistical texts.

HYDROLOGICAL SAMPLE FRAME

The viability indicators derive from two fundamental types of data: counts of fish migrating past a fixed point (counting stations, smolt production); and in-situ metrics of fish distributed among the reaches of a stream network (redd surveys, low-flow surveys). For the latter, Adams et al. (2011) outlined a two-stage sampling design: Stage 1 generates a probability sample of stream reaches, and stage 2 samples fish data from the sample of reaches. For this we need a sample frame—a digital representation of the stream network that divides it into a finite set of distinct reaches.

Methods used to delineate reaches are outlined in Adams et al. (2011) and described in greater detail by Garwood and Ricker (2011). A digital representation of the stream network is divided into reaches about 1–3 km in length, with boundaries between reaches marked by obvious on-the-ground landmarks such as tributary confluences, gully junctions, or bridge crossings whenever practical. Over the decadal scale of a monitoring program, the stream channel is expected to move around in response to water and sediment dynamics. The use of landmarks to mark reach boundaries generates a more consistent marker than, say, GPS coordinates. The target length of 1–3 km was identified as an appropriate scale for redd surveys in this region, but this target length can be adjusted to conditions for other regions, as long as it is kept consistent within sampling strata (BPGs). Since the sample frame is to be used in perpetuity, even stream reaches with moderate to poor habitat conditions and those above anthropogenic barriers are included in the frame, with the notion that as recovery efforts occur, habitat conditions will improve, and barriers will be remediated. Natural limits to anadromy, which generally consist of waterfalls or impassable bedrock chutes, define the maximum extent of the sample frame.

Below we outline some enhancements to this frame development and the stage-one sampling strategy, to improve practicality and operational efficiency.

LONG AND SHORT REACHES

For sample reaches, the target length of 1–3 km is appropriate for redd surveys, but longer than is practical or necessary for low-flow surveys in the southern area. For both electrofishing and snorkeling, the remoteness of sites and shallow conditions during the low-flow season often preclude sampling a reach length greater than 100–200 m in a day. For some regions, longer reaches up to 400 m may be feasible to sample in a day, and so the target length for short reaches is flexible. As with long reaches, one would want to have a consistent target length within any given BPG and avoid situations where the length tends to be longer in larger or more accessible channels than in remote tributaries. Otherwise, the final estimates will be biased.

There is little benefit in choosing a target length that takes two or more days to sample, because data collected in the second and subsequent days will tend to be similar to (spatially autocorrelated with) the data collected on the first day. In contrast, data collected at a completely different sample reach on the second day would have less spatial autocorrelation, and thus greater information content. More information for the same amount of effort is therefore gained by sampling shorter reaches during low-flow surveys. The key criterion here is not the length of the reach per se, but the length that can typically be surveyed in one day.

A simple method to implement this idea is to design the digital sample frame from the outset as a system of short reaches nested within the standard long reaches (Figure 8). A sample of long reaches is used for redd surveys, and a sample of short reaches is used for low-flow surveys. The nested or hierarchical relationship of the two frames is achieved simply by subdividing each long reach into a set of short reaches, so that the relationship between the two frames is precisely specified.

A key design choice is whether the long and short reaches should be independently sampled, or the short reaches treated as a subsample of the long reaches. In general, subsampling should be used when the goal is to examine correlations within a common target of estimation—for example when characteristics of spawning habitat are subsampled within a sample reach used for redd surveys. If the target of estimation is different—for example collecting data for a population using long-reaches and for a BPG using short reaches, it makes more sense to draw an independent sample. The reason is that subsampling is a form of cluster sampling, which usually provides less precision than a simple random sample, and two independent samples should thus have a greater total information content than subsamples of the same reaches. This also gives greater operational flexibility by

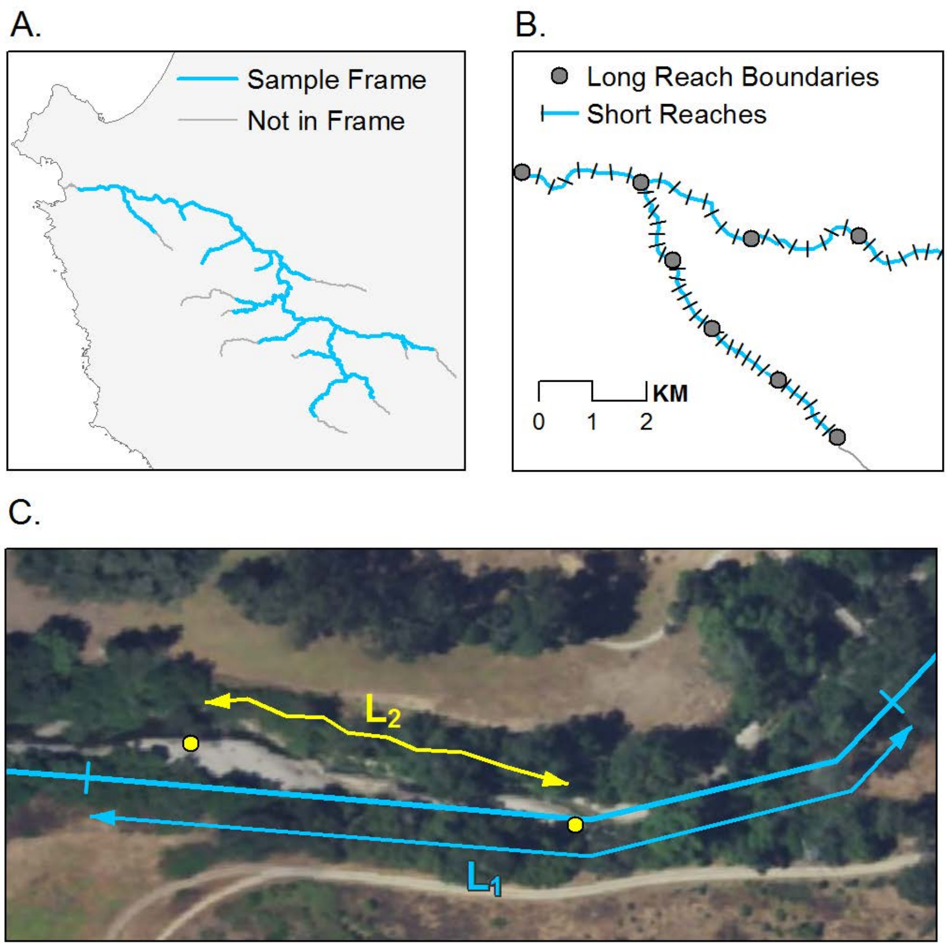

FIGURE 8. Elements of a sample frame, using a provisional frame from the Carmel BPG as an example. A. Entire sample frame for the target of estimation, showing natural limits to anadromy and excluding the estuary. B. System of short reaches nested within long reaches. C. One particular short reach in the GIS (blue segment within hash marks) versus the "same" reach defined by block nets on the ground (yellow dots). The downstream block net (left) is sited at the first unit boundary (pool/riffle transition) upstream from the end of the GIS reach. The upstream block net is the closest unit boundary to a point 100m upstream of the first block net. L1 and L2 show the relative lengths of the GIS reach and on-the-ground sample reach, respectively. The fish density measured within the block-netted section (L2) is assumed to apply to the entire length of the GIS reach (L1).

decoupling the two sampling strategies and is more easily reconciled with the two-stage/two-phase strategy for dealing with intermittent streams during low-flow surveys (described in the next section). However, a downside is that reach-sampling typically involves a process of contacting landowners and obtaining permission to sample, and the use of independent samples rather than subsampling substantially increases the number of landowners that must be contacted.

TABLE 7. Terminology for the sampling frame.

Term	Description
Reaches	
Long Reach	1-3 km stream segment, with endpoints defined by natural landmarks such as tributary junctions. Used for redd surveys.
Short Reach	100-200m stream segments, nested within long reaches, with endpoints defined by GPS coordinates. Used for low-flow surveys.
Habitat Unit	Individual pools, runs, and riffles within a short reach.
Fish-Sampling Reach	A set of habitat units within a short reach, usually 85m – 115m long, with endpoints defined by unit boundaries. Used for sampling fish and to represent fish density in the encompassing short reach (see Figure 8).
L_1	Length of the short reach, according to the GIS (see Figure 8).
L_2	Length of the fish-sampling reach, as measured in the field (see Figure 8).
Two-Phase Design	
Dry Reach	A sample reach determined to have lost all surface water by the time of sampling in the low-flow season.
Wet Reach	A sample reach determined to have maintained surface flow or residual pools since the previous wet season.
Sampleable Reach	A wet reach that can be sampled for fish (permission to access, suitable for collecting valid data).
d	Number of dry reaches.
w	Number of wet reaches.
n	Number of wet, sampleable reaches.
Stratification	
Sampling Frame	The complete set of long and short reaches available for sampling.
Target of Estimation	A portion of the sampling frame for which an estimate is desired.
Sample Reach	A long or short reach randomly selected for redd- or fish-sampling.

Developing new sample frames, which is currently in progress for the southern area, thus involves two steps: development of long reaches as described by Garwood and Ricker (2011), and subdivision of each long reach into short reaches. Although boundaries of long reaches will be marked by natural landmarks as described above, we expect this not to be practical for short reaches. Natural breaks at this finer resolution tend to be unit boundaries (i.e., between pools, riffles, and runs), which are dynamic and not consistent from season to season. We thus recommend that short-reach boundaries be defined digitally as GPS coordinates, but that a convention be established to identify field boundaries defined by natural units (Allen 2014). For example, in sampling the Carmel River we have navigated to the GPS coordinate at the downstream end of the short reach, and then begun sampling (established a block net) at the first unit boundary upstream (Figure 8). Duffy (2006) has described specific methods for identifying natural units. For clarity, key terms are defined in Table 7 and the scheme is illustrated in Figure 8.

TWO-PHASE SAMPLING IN ARID AREAS

For low-flow surveys, the sparse distribution of *O. mykiss* in large arid watersheds risks a situation where the sample consists entirely of dry fishless reaches, even though the species is present. This is an extreme form of negative bias that would provide misleading data. Our recommended solution to sparseness is to introduce a two-phase sampling strategy (Thompson 2012, §14.3). After initial stage-1 sampling (selection of sample reaches), a second phase of stage 1 sampling is conducted in mid to late summer, to determine basic suitability of each sample reach. Here, basic suitability refers simply to the existence of any surface water in the reach, including isolated pools, that has the potential to contain fish; and complete lack of surface water indicates zero fish, a safe assumption. Each reach is sampled sequentially from a randomized list for stage 1, until a prescribed number (sample size) are found to pass the screening for basic suitability. Then, in stage 2, the wet reaches are sampled for fish using one of the low-flow survey designs described on page 21.[4] Estimators for the wet fraction and its variance are given in Appendix B. For clarity, key terms are defined in Table 7. For targets where dry reaches are not found in the sample, the two-phase/two-stage design simply collapses back to the standard two-stage design.

A feature of this approach is that the identity and geographic extent of drought refugia will emerge over time from the low-flow surveys. The location and adequacy of drought refugia can be further clarified if desired by additional targeted low-flow sampling during multi-year droughts, to determine if the designation of backbone populations needs to be adjusted to protect the species.

[4] Occasionally a wet reach may dry out between phase 1 and 2. In this case, it is reclassified as a dry reach, and new sample reaches are drawn from the randomized list until a new wet reach is obtained for fish-sampling.

STRATIFICATION BY TARGETS OF ESTIMATION

Adams et al. (2011) described a general sampling strategy in which stream reaches throughout the range of salmonids were sampled at a uniform sampling rate (proportion of reaches sampled). In principle, estimates of status for any individual population or group of populations can be made after sampling, by using only the samples drawn from that particular target. Although this allows for a completely general monitoring strategy, it is not very efficient operationally when targets vary greatly in size (vary greatly in the number of stream reaches available for sampling). The precision of an estimate is generally dependent on two things: the variability among samples and the number of samples. For the latter, a uniform sampling rate will tend to either generate more samples than necessary for large targets or fewer samples than necessary for small targets.

We can observe this problem in the southern monitoring area, where the targets of estimation listed in Table 5 vary by a factor of 20 in the amount of stream habitat (Table 8). In Table 8 we explore a scenario where it has been established ahead of time that at least $n = 40$ wet reaches are needed to make an estimate with adequate precision. The sampling intensity could be set at a level to ensure this sample size for the largest BPG (Table 8, Column 3). But then five of the eight other BPGs have samples of 5 or fewer reaches on average, not anywhere close to enough for estimation. Or the sampling intensity could be set to ensure a sample of 40 reaches in the smallest BPG (Table 8, Column 4), but this expands to a total sample of 2847 reaches for the entire southern area. This level of effort seems unrealistic: If field crews sampled one reach per day for three months (60 days), such an effort would require about 50 crews.

TABLE 8. Sample size (number of reaches) for uniform sampling versus stratification by targets of estimation.

Biogeographic Population Group (BPG) (N to S)	Total Stream Km[1]	Uniform Sampling Intensity		Stratified, All BPGs[3]
		Largest BPG[2]	All BPGs[3]	
Interior Coast Range	12631	39	792	40
Carmel	646	2	41	40
Big Sur Coast	810	2	51	40
San Luis Obispo Terrace	1613	5	101	40
Monte Arido Highlands	12967	40	813	40
Conception Coast	1045	3	66	40
Santa Monica Mountains	638	2	40	40
Mojave Rim	7012	22	440	40
Santa Catalina Gulf Coast	8029	25	503	40
Total Sample Size (Reaches)		140	2847	360
ESU estimates possible		2 of 2	2 of 2	2 of 2
BPG estimates possible		2 of 9	9 of 9	9 of 9

[1] Compiled from CDFW 100k stream hydrography. For simplicity, sampling scenarios assume proportion of reaches that are wet and below natural barriers are similar among BPGs.

[2] Numbers of reaches sampled if $n = 40$ and the minimum target of estimation is the largest BPG (Monte Arido Highlands).

[3] Numbers of reaches sampled if $n = 40$ and all BPGs are targets of estimation.

In contrast, if we relax the requirement for uniform sampling, then sampling can be stratified by targets of estimation, with each BPG providing exactly 40 samples (Table 8, Column 5). Now the total sample is only 360 reaches, more than 85% smaller than a uniform sampling rate aimed at the same set of estimates; or only 6 field crews sampling one reach per day for three months for the entire southern coastal region.

For this reason, we recommend that Stage 1 sampling be stratified by the targets of estimation in Table 5. In low-flow surveys, the number of wet reaches needed would be about the same in each BPG, producing estimates of comparable precision. Similarly, in abundance estimation, each backbone population would have roughly the same number of reaches sampled for Rainbow Trout (Design 1) or redds (Design 2). This means that the sampling rate will vary greatly among targets, with higher rates in small geographic areas such as the Big Sur Coast, and lower rates in large geographic areas such as the Mojave Rim.

In BPGs where the abundance of Rainbow Trout is monitored using Design 1, the stratification scheme for the low-flow surveys must accommodate multiple targets of estimation (Table 9). The prescribed number of wet reaches to sample (n and n_A in Table 9) will depend on the desired precision in the estimate and the variability among wet reaches, which can be identified after preliminary sampling work. In practice, the larger basins will probably exhibit greater variability among reaches simply because larger basins tend to have greater heterogeneity in habitat conditions. To achieve a given precision, a larger sample of reaches may be necessary in larger basins, which can be determined empirically after a year or two of data have been collected.

Table 9. Stratification scheme for low-flow surveys.

Target of Estimation	Purpose	Sample size[1]
Each biogeographic population group (BPG)	Density, Genotypes, Occurrence, Drought Refugia	n per BPG
Each backbone population within a BPG	Rainbow Trout Abundance	Draw additional samples for n_A per population

[1] The number of wet, sampleable reaches required to make the estimate at the desired precision: $\boldsymbol{n}$ for occurrence, density, and genetic indicators, and $\boldsymbol{n_A}$ for abundance of Rainbow Trout.

SPATIAL BALANCE AND ROTATING PANELS

Two strategies for operational efficiency implemented by Adams et al. (2011) were spatially balanced samples of reaches, and assigning reaches to rotating panels. The first, spatially balanced random sampling, provides a more even spatial distribution of reaches than simple random sampling. This decreases spatial autocorrelation in the dataset and effectively increases the information content for a given sample size. We embrace this approach but, like Adams et al. (2011), describe estimators (Appendix B) that assume a simple random sample for simplicity. These estimators provide accurate estimates but will tend to overestimate sampling error. More sophisticated estimators are available (e.g., Kincaid et al. 2019) if desired for specific analyses.

Rotating panels accommodate the dual goal of the CMP to estimate both status and trends of indicators. An independent sample of reaches each year tends to increase the precision of status, while repeat sampling of reaches from year to year tends to increase precision of trend, by allowing estimators to disentangle spatial variation from temporal variation. Adams et al. (2011) proposed a rotating panel scheme in which reaches are systematically assigned to different panels with different schedules for revisits. One panel is revisited every year; other panels are revisited at staggered multiyear intervals, which expands the total number of reaches that can be sampled while still retaining the statistical benefits of revisits. We embrace this approach but propose that the 46-panel scheme of Adams et al. (2011) be greatly simplified for implementation in the southern area. Some targets of estimation are too small (too few sample reaches) to accommodate this level of complexity; in other cases, the sample size (number of reaches sampled) may be too small. In Table 10 we provide a simple scheme for ensuring that each panel is assigned at least 5 sample reaches, and that the target of estimation is not assigned a more complex scheme than it can usefully accommodate. We expect that in general the required sample size to achieve acceptable precision will lie somewhere between 20 and 80 reaches, but a more exact number cannot be identified until an initial dataset is collected. For each target we recommend starting with 30 reaches (short or long, depending on indicator) and adjusting as needed in subsequent years to achieve the desired precision.

TABLE 10. Tailored Rotating Panel Scheme for Small Targets of Estimation.

Number of Sample Reaches		Rotating Panel Scheme[3]	Description
Frame Size[1]	**Sample Size[2]**		
20 or fewer	**100%**	Census	Perform 100% survey every year. Data still reported by sample reach.
21 to 100	**20 to 80**	Annual Random Sample	New spatially balanced sample every year. Sample 30% of frame or 20 reaches, whichever is greater.
> 100	**20 to 80**	3-Year Rotating Panel	1 panel revisited every year; 3 panels revisited every third year.

[1] The total number of reaches available for sampling within a given target of estimation. Targets are generally the set of reaches supporting a given population, or a given biogeographic population group, depending on the indicator.

[2] Recommended range of sample sizes for a given target; the precise number chosen will depend on the desired error variance for the estimate, which cannot be known until an initial dataset is collected. We recommend starting with 40 reaches and then after a few years, increase or decrease to achieve desired sampling variance.

[3] Recommended rotating panel scheme for a given frame size, with equal numbers of reaches per panel.

SLOUGHS AND ESTUARIES

Due to their distinct spatial characteristics, a separate frame should be developed for the entire set of estuaries used by *O. mykiss*, as well as coastal channels and sloughs. Examples of the latter can be found along the coast south of Santa Clara River or north of Arroyo Grande and Salinas River, but the full set needs to be identified. We recommend that low-flow surveys include sloughs and estuaries because the lagoon-anadromous form of *O. mykiss* that depends on these habitats is a key element of life-history diversity, whose expression is necessary for species viability (Boughton et al. 2007). Stage-1 sampling in sloughs and estuaries can use a similar finite-sampling frame as stream reaches, but each sampling unit will be a small area of water rather than a small length of channel. They would be sampled as a distinct target of estimation for each BPG.

LOW-FLOW SURVEYS

Low-flow surveys focus on the distribution, density, and genetic diversity of juvenile *O. mykiss* in streams and estuaries, as well as the distribution of drought refugia. The targets of estimation are the Biogeographic Population Groups. However, where abundance monitoring uses counting stations, targets of estimation also include Rainbow Trout abundance in backbone populations (see p. 26). Two field methods—electrofishing and snorkel surveys—are suitable for low-flow surveys, but both methods detect fish imperfectly and thus require calibration via strategies to estimate detectability. Below we review the pros and cons of the two field methods and three popular calibration methods, and also the integration of calibration methods using double-sampling (Thompson 2012).

DESIGN 1: ELECTROFISHING

Temple and Pearsons (2007) describe the general method of electrofishing. In the shallow streams of the southern area, backpack electrofishers are typically carried to the sample reach and used to systematically sample live fish, moving from the bottom to the top of the sample reach. The electrical field in the water, generated by the electrofisher, temporarily stuns stream fish, allowing them to be located and netted by the electrofishing crew. Because fish may detect the electric current and escape upstream or down before being stunned, block nets are typically set up beforehand at either end to close-off the reach during the procedure. Fish are collected live in buckets during the procedure, counted, and released back to the stream after data collection.

Rarely are all fish captured, so the count resulting from a single pass of the electrofisher is a negatively biased estimate of abundance and density. However, if care is taken to implement uniform sampling intensity across reaches,[5] counts of captured fish tend to be highly correlated with more accurate estimates of

5 Commonly quantified as seconds of electrofisher operation per wetted area of stream.

abundance, such as closed-population mark-recapture (Jones and Stockwell 1995, Bertrand et al. 2006, Foley et al. 2015, Korman et al. 2016). As a result, even though single-pass electrofishing produces biased counts, the counts can contain substantial information about true density, and can be calibrated (bias-corrected) using double-sampling.

Genetic monitoring of the Omy5 haplotypes is readily incorporated into electrofishing. During handling, juvenile steelhead are subsampled for tissues via a caudal fin-clip. Genetic analysis of fin clips provides the fraction of the sampled fish that are homozygous (AA) or heterozygous (AR) for the anadromous haplotype (Pearse et al. 2014). These samples are then used to estimate proportions of the haplotype at the level of BPG (see estimator in Appendix B).

Advantages of electrofishing over snorkeling are that fish are directly handled, so samples for genetic analysis can be collected and adult Rainbow Trout can be identified at close range. Additional useful data such as size, weight, and age via scale samples, can also be collected to support more detailed ad hoc population studies as needed. Data on other fish species can also be collected if desired. Electrofishing is a flexible, information-rich method for low-flow surveys.

Disadvantages include non-lethal impacts on the fish such as stress or occasional spinal injuries that reduce subsequent growth rate, and typically a low level of mortality on the order of 1% or lower (~1% in Ainslie et al. 1998; usually ~0.1% in our experience). To reduce such impacts, permitting agencies currently restrict electrofishing for steelhead to times when stream temperatures are less than 18° C, which in the southern area tends to restrict sampling to the morning when water temperatures are still cool. In warmer areas electrofishing may not be feasible even in the morning. Permitting may also be limited due to impacts on other aquatic vertebrates, including frogs, but at a given power density the impacts on frogs are typically no greater than on fish (Gilbert et al. 2017). Threatened and endangered species such as the California red-legged frog (*Rana draytonii*) typically require special permitting, training, and procedures to avoid harm while sampling via electrofisher.

Unfortunately, backpack units generally become ineffective samplers in water deeper than 1 meter, due to attenuation of the electrical field as well as inability of operators to safely wade into deeper waters or reach structurally complex habitat. Deep pools with extensive areas deeper than 1 meter are important habitat for Rainbow Trout, and for juveniles during years of low summer flows (Boughton et al. 2009, Hwan and Carlson 2016), so omission of such areas will tend to bias estimates low, particularly for abundance of Rainbow Trout. We recommend that sample reaches with deep pools be identified during the screening for basic suitability and then surveyed using snorkel surveys (see Anadromous Fraction, p. 29).

DESIGN 2: SNORKEL SURVEYS

O' Neal (2007) describes the general method of snorkel surveys. One or more divers lie face-down in the stream channel at the bottom of the sample reach, and systematically work their way upstream identifying and counting fish. As with single-pass electrofishing, the resulting counts are typically biased low due to imperfect detection, but highly correlated with more accurate estimates of density. Bias is expected to covary with water clarity, so that negative bias is likely higher in the rainy season when turbid stream conditions are more common. At extremely high densities the bias can switch to positive due to accidental double-counting of fish, but we believe this situation will be rare in the southern monitoring area.

Snorkel surveys in long reaches were recommended by Adams et al. (2011) to sample occurrence. Unfortunately, much of the habitat in the southern area is so shallow that snorkeling long reaches is impractical. Riffles with mean depths less than 10 or 20 cm are extremely common and often support substantial numbers of age 0 steelhead (e.g., Boughton et al. 2009). Sampling can be effective when crawling in the channel and placing only one eye underwater, carefully looking upstream, and identifying fish before getting close enough to spook them. But this approach requires extreme care and diligence, is laborious, and in practice is unworkable for distances greater than 100 or 200 m. At the same time, 100 or 200 m appears sufficient to usually detect the species when it is present; Boughton et al. (2005) estimated detection failure while snorkeling 100 m reaches to be only 1.75%. The need to sample shallow habitat is not unique to southern California; Constable and Suring (2015) found that when estimating densities of three species of juvenile salmonid in coastal Oregon, a sample of habitat down to 20 cm deep gave less biased estimates with smaller confidence limits than a sample of habitat down to 40 cm deep.

Advantages of snorkel surveys over electrofishing include their lower impact on habitat, targeted fish species, and other species of aquatic vertebrates including threatened and endangered species such as California red-legged frog. The method is suitable for pools deeper than 1 m, and simpler to implement, requiring no foot transport of heavy equipment, setting of block nets, nor restrictions due to water temperature. A disadvantage is that calibration requires a separate method at a subset of reaches (see double-sampling below). Snorkeling may also pose health risks to survey crews when streams are polluted or contain pathogens.

A key disadvantage in practice is that fish are not directly handled, which limits the type of data that can be collected. In particular, tissue samples cannot be collected for genetic monitoring. In addition, the need for separate counts of adult Rainbow Trout and juveniles requires development of practical field criteria for distinguishing them visually.

CALIBRATION

The two sampling methods described above provide biased estimates that cannot be directly interpreted as fish density. Calibration refers to procedures to bias-correct the estimates by collecting data on nuisance parameters—in our case, on per-fish detectability. If the fish count from a reach is s, and the probability of detection is p, then an unbiased estimate of the true count is simply s/p (Thompson 2012). However, estimation of p requires additional sampling strategies, and some strategies for estimating p are more robust than others.

For electrofishing, detectability (catchability) is generally estimated by depletion or mark-recapture designs. The sample reach is first isolated via block-nets, so that catchability is not confounded with fish movement into or out of the reach. In a depletion design, the electrofishing crew then repeatedly passes through the reach using consistent sampling effort, holding the captured fish and recording separate counts for each pass. The decline in the counts with each pass allows for an estimate of catchability (Carle and Strub 1978). In mark-recapture, fish captured in the first pass are immediately marked and released. After waiting to allow marked fish to mix with the uncaptured fish, a second pass is conducted and the proportion of marked fish in the second sample provides an estimate of p (Seber 1982).

Fish usually vary in catchability, which violates an assumption of equal catchability that both estimators rely on. Compared to depletion, mark-recapture tends to be more robust to this violation, because with each pass the same group of fish are being sampled and thus the mean catchability stays the same. For depletion, in contrast, the more catchable fish tend to be caught sooner, and the successive passes are increasingly enriched by less catchable fish. Peterson et al. (2004) found that the resulting bias averaged -12% relative to mark-recapture. Other advantages of mark-recapture relative to depletion are the elimination of both a third pass and the need to hold fish until completion of sampling. Historically, a major disadvantage of mark-recapture was a requirement to wait 24 hr until the second pass, to allow marked and unmarked fish to mix. However, Temple and Pearsons (2006) showed that valid estimates could be obtained by waiting only 3 hr, allowing completion of the procedure within a single day. This shorter waiting time also reduced additional biases commonly caused by failure of block nets or escape of fish through the nets during the longer waiting time. We conclude that mark-recapture is clearly the preferred method for calibration. However, it is possible that sufficiency of the 3-hr waiting time from Temple and Pearsons (2006) may be sensitive to local conditions that impede movement and redistribution of fish after they are released from the first pass. The extreme low-flows characteristic of the southern monitoring area may pose this problem, possibly requiring a tailored methodology to counteract it. Thus, a clear research need is to identify and validate specific protocols both for wait time, and for releasing fish from the first pass in a way that encourages mixing (for example, distributing their release throughout the sample reach).

For snorkel surveys, detection rate can be estimated by replicating snorkel surveys in the same sample reach or across a set of reaches, and assuming variation in counts is due to binomial sampling error (Royle 2004). Unfortunately, estimates using this method are not robust, especially when detection rate is low or the number of replicated counts is low (Dennis et al. 2015). Thus, we do not recommend that snorkel surveys be replicated as a calibration method. Instead, a subset of snorkel surveys should be calibrated by pairing them with mark-recapture electrofishing, using a double-sampling design.

DOUBLE-SAMPLING

Snorkel surveys and single-pass electrofishing provide biased indicators of relative density, while mark-recapture electrofishing provides a relatively robust, unbiased method for estimating true density. Hankin and Reeves (1988) pointed out that at the broad scale of entire river basins, spatial variation in density is often much greater than spatial variation in detectability. In such cases operational efficiency (accuracy per unit effort) can often be greatly enhanced by increasing the total number of sample units (stream reaches) through use of the simpler, biased method; and then bias-correcting the result via a smaller sample of calibration reaches where both methods are applied. Thompson (2012) refers to this design as double-sampling, and the biased indicator as an auxiliary variable. Double sampling is useful when the quantity to be estimated is linearly related to the auxiliary variable, tends to be zero when the auxiliary variable is zero, and is harder or more expensive to sample than the auxiliary variable. Hankin and Reeves (1988) recommended combining double-sampling with a particular stratification scheme (by individual habitat units and watershed locations); as described earlier our approach here differs by limiting stratification to targets of estimation.

In double-sampling, a sample of reaches is first screened for basic suitability until *n* wet, sampleable reaches have been identified. Of these *n* reaches, *k* are then randomly selected for double-sampling, and the remaining *n* – *k* reaches are sampled solely for the auxiliary variable. Estimators for density are given in Appendix B.

The most cost-effective proportion *k*/*n* depends on the cost of the two methods and the variances of the estimates (Thompson 2012, p. 186),

Eq. 1 $$optimal\ \frac{k}{n} = \sqrt{\frac{C_1}{C_2}\left(\frac{\widehat{Var}(r)}{\widehat{Var}(\hat{y}_k) - \widehat{Var}(r)}\right)}$$

where C_1 is the cost for sampling the auxiliary variable at one reach, C_2 is the cost for obtaining an unbiased mark-recapture estimate at one reach, and the variance terms are the ratio and sample variances given in Appendix B. Note that if *r*, the ratio between mark-recapture estimates and auxiliary counts, is too variable among the reaches, the auxiliary counts are not as informative and the value of k begins to approach n, at which point the advantage of the cheaper biased method disappears. Similarly, if the biased method is not cheaper the advantage also disappears, although it might still be retained for other reasons such as reducing negative impacts on steelhead or other aquatic species. The variances cannot be known prior to sampling, but variances from past years of sampling can be used as an approximation. Dauphin et al. (2019) observed that calibration relationships change over time, so regardless of cost estimates, sufficient double-samples should be collected each season to make final estimates.

Detectability of fish varies across sites for a variety of reasons, including both environmental factors such as habitat complexity and visual conditions, and factors associated with the sampling crew, such as variation among crewmembers in skill at catching or visually identifying fish. The higher the variability in fish detection among sites, the lower the likelihood that double-sampling is a worthwhile cost-saving strategy—a strategy of making unbiased mark-recapture estimates at every sample site may perform just as well in terms of precision achieved as a function of number of reaches sampled.

For this reason, we do not expect double sampling to be useful for Design 1 electrofishing, where capture rates can vary greatly among reaches as a function of habitat complexity and crew experience. In addition, the time saved by conducting a single pass rather than a two-pass mark-recapture session will rarely be enough to allow a second reach to be sampled the same day, due to limitations imposed by travel time and water temperature. Thus double-sampling is unlikely to allow a larger sample of reaches to be surveyed by the same number of crew-days.

The situation is different for snorkel surveys, which may be used because electrofishing imposes an unacceptable impact, or because multiple reaches may indeed be feasible to survey in a single day by a given field crew. These surveys would need to be calibrated, and as described above our preferred method is mark-recapture using electrofishing. So even when snorkeling is used to either expand the sample size or limit the negative impacts of electrofishing, a calibration sample of electrofished sites must still be included.

SAMPLING LAGOONS, SLOUGHS AND ESTUARIES

Compared to streams, the field methods for sampling lagoons, sloughs and estuaries are relatively undeveloped. Seine pulls at a series of sites within a given estuary are traditional methods for sampling fish communities (Fierstine et al. 1973, Horn 1980, Gilchrist et al. 1992, Roth et al. 2000, Alley 2017, Hagar Environmental Science (HES) 2017), but the exact sites are typically selected non-randomly for ease of access, and do not represent a unit within a finite sampling frame. Seining samples of lagoons, sloughs and estuaries are rarely calibrated to account for detectability, but mark-recapture has sometimes been used over two-day sampling sessions to estimate abundance, for example juvenile steelhead in Pescadero Lagoon, San Mateo County (Jankovitz and Diller 2019 and earlier reports). This approach could be codified and integrated with a first-stage sampling scheme using a finite sampling frame. Smaller estuaries would probably comprise a single sample unit, but larger estuaries or complex systems of sloughs might better be monitored using a sampling frame of multiple units. Here is a key need for methodological development.

ABUNDANCE AND ANADROMOUS FRACTION

For monitoring abundance, Adams et al. (2011) recommended fixed counting stations in the southern area and biweekly redd surveys in the northern area. We review both these designs as potentially useful in the southern area and also consider strategies to collect data on anadromous fraction. For counting stations, we recommend Rainbow Trout be estimated from low-flow surveys. For redd surveys, we recommend collection of auxiliary data during the surveys so that redds of steelhead can be distinguished from those of Rainbow Trout.

DESIGN 1: COUNTING STATIONS AND LOW-FLOW SURVEYS

This design has two elements: estimation of steelhead abundance using counting stations, and estimation of Rainbow Trout abundance using the low-flow surveys. These two estimates are then combined to estimate total adult abundance and anadromous fraction. Adams et al. (2011) discussed counting stations set up on stream channels close to the ocean to directly count migrating steelhead but did not address the Rainbow Trout component.

A variety of methods for counting stations have been tried or proposed, including fyke nets, acoustic cameras, photographic systems such as the Vaki Riverwatcher, resistance-board weirs, and approaches using passive integrated-transponder tags (PIT tags) (Adams et al. 2011, Williams et al. 2016). Unfortunately, an optimal design for counting stations remains elusive. We consider the most promising designs (Table 11) and identify outstanding issues that would benefit from methodological research. In our view, no one design is likely to be generally superior; instead, particular designs will be favored for particular site conditions. The methods likely involve different inherent biases in their estimates; and thus, once a deployment decision is made, a given method should be used consistently for a given population, to support valid trend estimation.

TABLE 11. A taxonomy of potential designs for counting stations.

Method	Description	Key Issues
Traps	Traditional deployments of fyke nets, resistance-board weirs.	Often incompatible with episodic flow regimes and high bed loads.
Visual-light Video Cameras	Vaki Riverwatcher, overhead video monitors, other systems based on visual light.	Unsuitable for turbid conditions, episodic flow regimes. Interpretation of data as run size.
Sonar Cameras	DIDSON, ARIS sonar-based imaging systems.	Species identification is problematic. Interpretation of data as run size.
Combination Capture/ Sonar	Sonar camera paired with weekly capture session to estimate species proportions.	Low abundance and/or overlapping season may reduce effectiveness of species ID.
Combination Video/Sonar	Overhead video cameras paired with sonar camera.	Partial weir may be necessary. Deposition of bed load may obscure video. Turbidity may still be problematic.
PIT-tagging Approaches	Estimate of smolt production paired with tag-based estimate of marine survival.	Requires a large sample of tagged smolts or juveniles, and effective monitoring stations.

Weirs. Adams et al. (2011) argued that established methods using traps (fyke nets, resistance-board weirs) were likely to be incompatible with the episodic flow regime in the southern region. Resistance-board weirs and other types of trap require instream infrastructure that is vulnerable to failure from extreme flow events, as well as high bedloads and debris transport during those events (Zimmerman and Zabkar 2007). In the more arid regions, these flow events may provide the only opportunities for adults to move upstream, such that the equipment is least functional at exactly the time it is needed. Weirs can be pulled or secured in advance of high flows, and reinstalled when flows come down, but they then miss a large unknown fraction of the migrants.

It is sometimes believed that as a high-flow event proceeds, steelhead often stop migrating due to high water velocities well before the weir itself is overtopped or blown out. Counts of migrant steelhead in a fish ladder on the Carmel River suggest this is unlikely (Figure 9)—steelhead do appear to eventually stop migrating at high flows, but only after flows exceed 4000 ft^3/s which is well past when a typical weir is overtopped. Streams amenable to resistance-board weirs are therefore probably rare in the flashy stream systems of the southern monitoring area, although it is worth considering whether an engineering solution can be found by updating the design of resistance-board weirs.

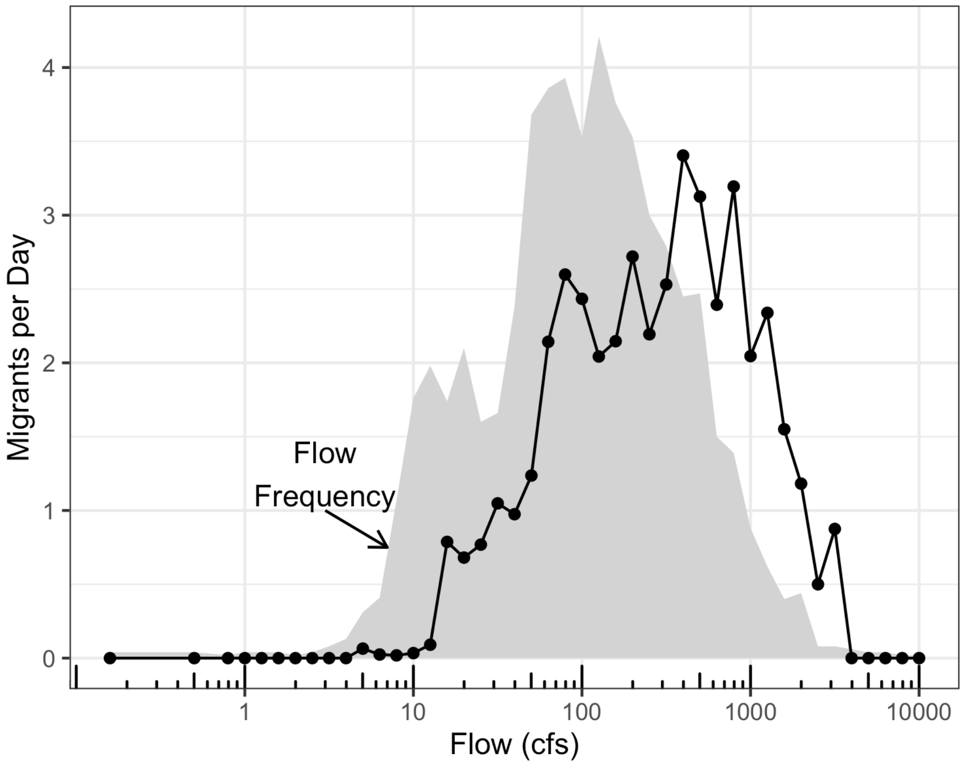

Figure 9. Adult migrants per day for different flows in the Carmel River, for the 11,032 steelhead ascending the San Clemente Dam fish ladder from 1993 through 2015. Daily flows are from USGS gauge 11143250 near Carmel, tabulated for the typical migration season of January through May. Units for flow frequency are 100 days per 23 years.

Sonar Cameras. Current technology favors the use of sonar cameras, in part because they require minimal instream infrastructure and can usually "see" fish during turbid conditions. Sonar cameras include dual-frequency identification sonar cameras (DIDSON cameras) and adaptive resolution imaging sonar (ARIS cameras). These cameras generate and process sonar signals to produce imagery suitable for enumerating adult steelhead as they migrate upstream past the monitoring site (for more details see Pipal et al. 2010b, Pipal et al. 2012, Atkinson et al. 2016). Deployed low in the watershed downstream of spawning areas and remaining in the river for the entire spawning season, the data from the cameras can provide an estimate of steelhead escapement for the season even at very high or very low abundance. The lack of an infrastructure that might impede steelhead movement behavior also gives this design the lowest direct impact on the steelhead themselves. However, partial weirs are sometimes deployed to direct fish into the portion of the channel where the acoustic camera can retrieve the highest-quality image (Denton et al. 2014, Denton et al. 2015)

One of the key issues with using the sonar cameras is the ability to distinguish steelhead from other species of similar shape and size, and to apportion the counts accordingly. Although no other native anadromous salmonids occur in the southern area, there are a variety of introduced species that might be confused with steelhead in the imagery. For example, large carp (within the size range of steelhead) have shown up in sonar camera datasets for the Ventura River (K. McLaughlin, pers. comm.), and large striped bass probably occur in DIDSON datasets collected for Carmel River steelhead in recent years (C. Hamilton, pers. comm.) Certain large-bodied native species such as Sacramento sucker and pikeminnows may also sometimes be observed moving upstream during the migration season for steelhead. The problem of species apportionment is likely relevant throughout river systems of the Interior Coast, Carmel, Monte Arido Highlands, Mojave Rim, and Santa Catalina BPGs.

Sonar/Capture Combination. One solution to the species apportionment problem is to conduct periodic capture sessions of migrating fish during the spawning season. This strategy has been used successfully in the Elwha River in Washington State, where weekly tangle net surveys were conducted at several reaches in the lower river (Denton et al. 2014, Denton et al. 2015). Captured migrating fish provided a quantitative estimate of the proportions of different species moving upstream in a given week, which were then used to adjust the total DIDSON counts into counts of steelhead and other salmon species for that week. Tangle-nets are designed to snare fish from the nose or jaw as opposed to the gills, allowing ensnared fish to be released with low immediate mortality and relatively low post-release mortality; Ashbrook et al. (2007) give an overview of the method. Tangle nets are just one potential capture method for sampling proportions; other methods such as hook-and-line sampling (Atkinson et al. 2016) or quantitative eDNA sampling might also be considered.

In the Elwha River, this strategy provided escapement estimates of steelhead with pleasingly high precision (coefficient of variation of 2.7% in 2014 and 7.5% in 2015). The high precision probably depended on the three salmonid species in the river having only modest overlap in migration timing (G. Pess, pers. comm.). Weekly proportions estimated from the tangle-net surveys were usually either close to 1 or close to 0, with intermediate values only occurring during shorter periods when the run of one species was waning and the next species was waxing. Introduced species opportunistically exploiting high-flow events may not exhibit such regularity.

Other factors to consider are (1) the extreme rarity of anadromous steelhead in some watersheds, which might prevent the estimation of meaningful proportions from capture ratios; (2) whether it is appropriate to assume equal catchability of steelhead and introduced species; and (3) the potential impact of capture itself on the adult steelhead.

Sonar/Video Combination. A second potential solution to the species apportionment problem is to combine a sonar camera with a conventional video system (Killam 2012). The sonar camera is used to make complete counts continuously, and a video system is used to make counts and species identifications only in daytime during the clearer water conditions between storms. Apportionment during the turbid conditions could be made either by (1) assuming a similar proportion of species during turbid and clear conditions; or by (2) using the concurrent video and sonar images to identify species criteria in the sonar images (e.g., image size, shape, tail-beat frequency), and using them to apportion the images collected during turbid conditions.

Killam (2012 and annual reports in subsequent years) described a successful design deployed in tributaries of the upper Sacramento River. In that design a partial weir is used to focus upstream migrants through a narrow portion of the stream channel. The bed of the constriction is covered with a light-colored plastic sheet to provide visual contrast, and a video camera is suspended overhead to record data. A sonar camera is deployed just upstream or downstream.

We suspect the partial weir of Killam's (2012) design may prove problematic in many southern California streams for the reasons described previously, but the design could potentially be modified by eliminating the weir and suspending multiple video cameras laterally across the channel. The relative affordability of video equipment makes this cost-effective. A key challenge is deploying a high-contrast bed cover in such a way to minimize scour or deposition of bedload during high flow events, either by site selection or by design.

PIT-Tag Designs. Finally, abundance of migrating adults can be estimated using PIT tags (Boughton 2010). Passive integrated-transponder tags (PIT tags) are small, cheap electronic tags that can be implanted in fish. They encode unique ID numbers that can later be decoded by electronic monitoring stations as the fish pass by. A monitoring station would be established on the river near its estuary, where it would detect and identify adult tagged steelhead as they migrate upstream. There are two basic PIT-tag designs for monitoring abundance, depending on whether fish are originally tagged as juveniles or smolts.

With smolts, established methods such as rotary-screw traps are used to estimate smolt production (see next section), and some or all of the captured smolts receive PIT tags prior to release. The monitoring station near the river mouth is then used to detect the proportion of tagged smolts that survive and eventually return to spawn (marine survival). Steelhead abundance is then estimated as the product of smolt production and marine survival.

With juveniles, the low-flow surveys are used to tag a sample of juveniles and to estimate total juvenile abundance. The monitoring station detects both the outgoing tagged smolts and the eventual return of tagged adults, which allow estimates for both smolting rate (probability that a juvenile fish transforms to a downstream migrant) and marine survival. Steelhead abundance is estimated as the product of juvenile abundance, smolting rate, and marine survival (Boughton 2010).

Successful design and instream deployment of antennae for the PIT tag monitoring stations are key elements of this approach. Typically, the maximum detection distance between a tag and the station's antenna is on the order of 30–80 cm depending on the type and size of the tag. Careful antenna design is needed to maximize detection rate under typical river conditions (Ibbotson et al. 2004, Ramirez et al. 2013, Magie et al. 2015, Baker et al. 2017). Even so, detection rates are generally less than 100% and multiple antennae are necessary to estimate detection rates via mark-recapture models (Boughton 2010). The estimations of smolt production, marine survival, and adult run size are then special cases of robust-design mark-recapture models (Pollock 1982). In addition, many types of instream antennae have their own version of the weir problem (Figure 9), being vulnerable to snagging of debris, washout during high flows, or getting buried under deposited bedloads.

One promising antenna design is an easily constructed slack-loop design now in use by Oregon Department of Fish and Wildlife on the MacKenzie River (Tranquilli 2010), and currently being tested on the Carmel River in California. The antenna is a pass-through loop of wire deployed across the river and secured only at the banks. The lower limb is heavy gauge wire and lies loosely on the bottom of the channel, while the top wire is lighter gauge and surrounded by plastic tubing that floats it about 5–10 cm below the water surface. The wires are left slack to allow downstream billowing with the current. The design appears to have promise because the slack loop tends to shed snagged debris and can be deployed in such a way that when it does get blown out, it can be redeployed even at relatively high flows by field crew working at the riverbank.

A key limit on the overall precision of both the juvenile- and smolt-based PIT-tag designs is the number of tagged adults that eventually return and get detected (Boughton 2010). Around 30 to 100 tagged returns are needed to achieve a coefficient of variation of 30% in the estimate of adult abundance for a given year. When smolts are tagged, or juveniles are tagged and marine survival is relatively low (~0.33%), around 30 are needed and the rule of thumb for tagging effort is

Eq. 2 $$\text{Tags needed} = \frac{4}{s_1 s_2 (0.0075CV)^{1.7}}$$

where s_1 is the expected smolting rate, s_2 is the expected marine survival, and CV is the desired coefficient of variation for adult abundance, expressed as a percentage. For smolt-tagging designs simply set $s_1 = 1$. When juveniles are tagged and marine survival is expected to be higher than 0.33%, the above rule-of-thumb should be multiplied by $3\log_{10}(300s_2)$.

Anadromous Fraction. Counting stations only provide data on adult steelhead, but data on adult Rainbow Trout can be obtained from the low-flow surveys described earlier. The main modifications needed for the low-flow surveys are a suitable stratification scheme to support dual targets of estimation (Table 9) and field criteria for distinguishing adult Rainbow Trout from juvenile *O. mykiss*.

The standard target of estimation for low-flow surveys are the broad biogeographic groups of populations mapped in Figure 6, but the target for adult abundance are usually smaller drainage basins supporting individual backbone populations. Therefore, additional reaches will need to be sampled from these smaller targets to estimate Rainbow Trout abundance.

As for field criteria, given what we know about the flexibility and complexity of *O. mykiss* life-history, there can be no simple, unambiguous field criteria for distinguishing adult Rainbow Trout from immature juveniles, particularly since some parr—especially males—may be reproductively active at a small size, while larger fish may appear to be adults but transform to smolts rather than reproduce. Rather than conclude that it is simply infeasible to estimate abundance of adult Rainbow Trout, we recommend a pragmatic two-pronged approach. First, during low-flow surveys, use fork length to assign fish to first-order categories. Second, various methods can be used to develop a rating curve for maturity as a function of size and sex.

For the first prong, findings from Scott Creek steelhead, a population just north of the monitoring area (e.g., Hayes et al. 2011, their Figs. 2 and 4) suggest the following first-order rating scheme:

Fork Length < 150 mm:	**Probably immature**; mostly immature *O. mykiss*, but can include precociously breeding adults, especially males.
Fork Length 150 – 200 mm:	**Enigmatic**; fish that may have forgone smolting and are maturing into Rainbow Trout, but presmolts and precocious breeders are common.
Fork Length > 200 mm:	**Probably adult Rainbow Trout;** but can include fish that will later smolt, especially if rearing in a high-growth habitat such as a lagoon.

For the second prong, various methods are available to construct a locally-tailored rating curve for maturity as a function of fork length and potentially sex. The definitive way to establish adulthood is to kill the fish and dissect the gonads. For example, in a study of steelhead partial migration, Ohms et al. (2014) determined stages of gonadal development histologically, using light microscopy and protocols from Nagahama (1983). A disadvantage of this approach is the lethal sampling, but in practice a small subsample of fish could be sacrificed to develop a rating curve for each sex. The least intrusive way to accomplish this while still maintaining statistical rigor would be to sample fish sparsely from across the entire southern area during routine low-flow surveys. The underlying assumption of this sparse approach is that the same rating curve describes different, finer-scale targets of estimation, preventing local tailoring of the rating curve.

A less destructive method is to sample live fish and use a small hand-held ultrasound device to identify females with mature gonads; Bangs and Nagler (2014) successfully developed and used this approach to identify mature female cutthroat trout (*Oncorhynchus clarkii*). After measurement, the live fish are returned to the stream. Although non-lethal, a limitation of the method is that it cannot identify mature males and requires an assumption that males and females have similar maturation curves.

Finally, genetic methods could potentially be used to construct locally-tailored rating curves (Kelson et al. 2019, Pearse et al. 2019). Both the Omy5 haplotypes and the sex of individual *O. mykiss* can be determined from genetic analysis of a tissue sample. In a cohort of juvenile fish, the sex ratio is initially close to 50:50 and the distribution of Omy5 genotypes (AA, AR, RR) is independent of sex, but due to the association between sex and expression of migratory phenotype (Pearse et al. 2019), as the cohort grows through the typical smolting size the non-migratory fish remaining in the stream decrease in proportion of RR fish, increase in proportion of AA fish, and the proportion AR males and females diverges. Pearse et al. (2019, their Fig. 3) show that this change in genetic composition begins around 100 mm fork length, but they measured fish length the fall before the smolting season so subsequent growth is not reflected. To construct a locally tailored rating curve using genetic methods, the simplest design would be to nonlethally sample genetics from a cohort of fish at repeated intervals from the beginning to the end of smolting season and identify the size at which the change in genetic composition occurs. This size threshold could be used to estimate the abundance of Rainbow Trout, though not necessarily adult Rainbow Trout if there is a substantial delay between forgoing smolting and maturing into an adult.

A final concern with counting Rainbow Trout in low-flow surveys is the potential bias caused by deep pools. Backpack electrofishers are rarely able to effectively sample pools deeper than 1 m, but these are precisely the types of habitats where adult Rainbow Trout tend to aggregate. Thus, an electrofishing design will miss many of these fish, leading to an underestimate of Rainbow Trout abundance. However, snorkel surveys are often a highly effective method for obtaining estimates of fish abundance in deep pools.

For low-flow surveys based on snorkel surveys (Design 2, p. 36), the existing design could therefore be used with a small modification. In this design, a first sample of reaches is drawn for snorkel surveys, and then a second calibration sample is drawn from the first for double-sampling with an electrofisher. The modification: If a reach in the calibration sample has a pool too deep to electrofish, return it to the first sample and draw another reach from the first sample in its place. Repeat this procedure until an “efishable” set of calibration reaches is obtained. The key assumption of this modification is that any visual conditions that affect detectability of fish during snorkel surveys is not systematically different between the efishable and unfishable reaches.

What should be the criterion for efishability? The importance here is not a precise determination, but a simple rule of thumb that can be consistently applied in the field by diverse personnel. We suggest a short reach be rated as efishable if no more than 20% of the reach by length has a deepest point greater than 1.0 m. This can be determined by a quick wade-through of the reach checking deep points with a stadia rod, during the initial screening visit to the reach when it is surveyed for surface flow. If sampleable, the actual paired snorkel survey and electrofishing sample are subsequently conducted at a later date.

For low-flow surveys based on electrofishing (Design 1, p. 34), a slightly more complex strategy is needed. We propose that a short reach determined to not be efishable would still be sampled, by receiving a snorkel survey. A calibration sample would also be drawn in this case, specifically to calibrate these reaches too deep to electrofish. The calibration sample would be double-sampled, with snorkel surveys preceding mark-recapture electrofishing.

Pros and Cons. Of the various types of counting stations, the sonar/video combination is especially promising. If partial weirs—used to protect the instream deployment of the sonar device—can either be eliminated or adapted to withstand south-coast site conditions, it is the only design for a counting station that simultaneously (1) has the potential to resolve the species-apportionment problem; (2) does not require infrastructure as vulnerable to episodic flow events as cross-channel antennae or weirs; (3) can likely monitor fish throughout an episodic flow event; 4) involves no capture of anadromous fish or impacts on redds; and 5) can be used to monitor very small run sizes. The sonar/tangle-net combination has all these advantages except for its impacts on steelhead captured during the weekly netting operation. Other methods such as quantitative eDNA or hook-and-line sampling might also be suitable for estimating apportionment but have not been tried.

Cuthbert et al. (2014) used a video-only method at the Salinas River Diversion Facility, counting adult steelhead using a Riverwatcher visual-light camera system paired with a resistance-board weir. Their experience illustrates a key weakness of the video- and sonar-based approaches: the interpretation of the data is somewhat ambiguous. Counts are reported as numbers of upstream migrants and numbers of downstream migrants. Because steelhead may move up and down a stream corridor multiple times before or after spawning, additional assumptions are necessary to convert such counts into estimates of anadromous abundance (Pipal et al. 2010b, Pipal et al. 2012).

A disadvantage of sonar-based designs under current practice is the amount of personnel time required for post-processing the imagery. This is an issue even in BPGs where the only large-bodied fish can be safely assumed to be *O. mykiss*, because large-bodied invasive fish might colonize and go completely undetected and indeed could be mistaken for improvement in steelhead runs. However, current image analysis technology is advancing rapidly. It is likely feasible to develop automated post-processing methods that are highly accurate and low-cost to operate.

PIT-tagging approaches are attractive because they eliminate the ambiguity of converting upstream and downstream counts into anadromous abundance. Individual fish are uniquely identified by their tag ID, which resolves the ambiguity. In addition, in the process of estimating anadromous abundance, these designs also estimate smolt production and marine survival, which are necessary in life-cycle monitoring stations and are useful performance metrics in highly managed river systems. Key disadvantages are the large tagging effort necessary at the juvenile or smolt stage, especially when marine survival is low; and the need for steelhead runs in at least the dozens to hundreds to achieve the needed sample of tagged adults, which is not useful where steelhead runs appear to be in the single digits (Dagit et al. 2020). In addition, the method is unlikely to perform as well as the sonar + video design for low abundances. Finally, the viability of the method has not yet been demonstrated within the context of the episodic flow regime of the southern monitoring area.

DESIGN 2: REDD SURVEYS AND REDD ASSIGNMENT

Gallagher and Gallagher (2005) describe redd surveys, and Gallagher et al. (2007) provide detailed protocols. Two-stage redd surveys involve drawing a probability sample of long reaches (stage 1), at which biweekly redd surveys are then conducted for the duration of the spawning season (stage 2). Biweekly is generally sufficient to census redds in each sample reach, given the typical time that redds stay recognizable as such to surveyors.

These surveys give estimates of redd abundance for the target of estimation, but to estimate adult abundance, two additional calibrations must be made. First, redd abundance must be converted to spawner abundance, using an independent estimate of mean spawners per redd. As in the northern area, this independent estimate is typically obtained from life-cycle monitoring stations (see next section). Second, to estimate anadromous fraction, redds must be assigned to steelhead versus Rainbow Trout.

Local Feasibility. Adams et al. (2011) did not recommend redd surveys for the southern area for two fundamental reasons. First, redd surveys are only practical for watersheds meeting two rather restrictive conditions: long reaches must be readily accessible to field staff during the wet season, and typical flow conditions must be suitable for the field surveys themselves. Many of the larger watersheds of the inland areas are clearly too remote or hazardous for staff to visit biweekly in the wet season. Other watersheds may be accessible but have flows too deep or hazardous to conduct redd surveys, especially at the peak of the migration season (Feb – Mar). For example, redd surveys have been conducted in the Carmel River mainstem for many years, but biweekly surveys were usually infeasible for much of the rainy season due to high water, swift currents, and turbid conditions. Although redds could be surveyed sporadically, they were not suitable for making an estimate of total abundance.

Second, in some areas the anadromous form of the fish is so rare and sparsely distributed that a probability sample of reaches would usually find zero redds, even when steelhead are present, leading to incorrect inference. For example, redd surveys have been successfully conducted in Ventura River annually since 2013 (summarized in Williams et al. 2016) using a complete census of all reaches rather than a probability sample. The few redds found were spatially clumped, supporting the idea that a probability sample could have easily missed the few reaches where redds occurred. At a counting station on the Salinas River, Cuthbert et al. (2014) found consistently low numbers of steelhead ascending the river (annual counts of 13, 17, 43, and 0 steelhead over 2011- 2014). With some simplifying assumptions, we can make a rough estimation of the chances of detecting such small runs using redd surveys[6]. The chance that a random sample of 40 reaches would turn up zero anadromous redds is 64%, 55%, and 20% for the first three years respectively, with chances of zero detection even higher if redds tend to be aggregated into particular sample reaches.

On the other hand, biweekly redd surveys have been successfully carried out in smaller coastal watersheds, where targets of estimation are much smaller and more accessible during the winter. In southern California, small watersheds within the Conception Coast (Arroyo Hondo, Carpinteria Creek) and Santa Monica Mountains (Arroyo Sequit, Malibu Creek and Topanga Creek) have been successfully surveyed with a 100% sample (i.e., a census of all reaches) in one or more years (K. MacLaughlin, pers. comm.). Two tributaries of the larger Santa Clara River system (Santa Paula and Sisar Creeks) have also had index reaches successfully surveyed over multiple years. In these settings, access by field crews, moderate flows, and shorter total lengths of channel made a biweekly census of redds feasible.

Anadromous Fraction. The basic redd survey requires additional methods to assign individual redds to steelhead versus Rainbow Trout mothers. An obvious possibility is using redd size (length by width) or timing for the assignment. Gallagher et al. (2010) showed on the northern coast that redd size and timing of formation could be used to distinguish species (*O. mykiss*, *O. kisutch* and *O. tshawytscha*), mostly because of distinct run timing. They did not try to distinguish redds of steelhead from those of Rainbow Trout. In the Elwa River of Washington State, McMillan et al. (2015) showed that redd size, timing, and also the substrate size of the redd tailspill were all effective at discriminating Rainbow Trout from steelhead. The size-distributions of redds for steelhead versus Rainbow Trout did not overlap: the smallest steelhead redds were always larger than the largest trout redds. However, McMillan et al. (2015) noted that the clean discrimination of life history types was likely due to the small size of adult Rainbow Trout in the Elwa (74% no larger than 24 cm FL). They thought the lack of

[6] Simplifying assumptions: The number of anadromous redds is half the number of anadromous spawners, the redds are maximally dispersed in space (no reach has more than one redd), the Salinas basin has roughly 1120 km of potential habitat accessible to anadromous fish (from Boughton et al. 2006), and the sample reaches are 2 km long. Under such assumptions the chance of observing zero redds is $\prod_{i=40}^{i=1} \frac{561-r-i}{561-i}$ where r is the total number of anadromous redds in the basin.

intermediate sizes might simply be an artifact of an impassable dam that had recently been removed but had previously blocked interbreeding. In other basins such as the Deschutes River (Zimmerman and Reeves 2000), adult body sizes, redd sizes, tailspill substrate, and spawn timing all showed much more overlap between Rainbow Trout and steelhead. They would thus be less effective for classifying redds, or at least require local calibration due to variation among different river systems. In a review, Kendall et al. (2015) note that fully anadromous individuals generally mature around 500–1100 mm in length, while freshwater residents tend to mature at 100-350 mm length, but they also note that maturation schedules are under strong local selection and that both life-history types show a wide range in size and age at maturation. Indeed, rivers capable of supporting large Rainbow Trout are expected to select for the resident life-history (Satterthwaite et al. 2010).

So, redd size, date of formation, and substrate size might each be useful for assigning a life-history to the mother that constructed the redd. But they will require local calibration and validation, and may turn out not to be useful, especially if steelhead and Rainbow Trout frequently interbreed and produce adults of intermediate sizes, or if freshwater habitat such as large bedrock pools can support large adult Rainbow Trout. Both phenomena appear to be common in the southern monitoring area. Even so, timing might prove useful if most Rainbow Trout spawn in the summer, as was found in both the Elwa and Deschutes examples above, but this would then require redd surveys to continue through August and perhaps longer (e.g., see Figure 3 in McMillan et al. 2015). These two rivers in the Pacific Northwest have summer flows typically supported by snowmelt, and differences in timing may not be as pronounced in our Mediterranean climate where streamflows become very low and warm in summer and are not particularly amenable to incubation after mid-June.

Another potential method for assigning maternal life history is to follow up the redd survey by sampling fish densities five or six weeks later after young-of-the-year have emerged from the redd. The typically higher fecundity of anadromous females relative to resident females may show up in snorkel counts per unit of streambed area, allowing the two kinds of mothers to be distinguished via the density of their offspring—densities of fish in reaches visited by anadromous adults are often an order of magnitude higher than those with resident only (B. Spence, pers. comm.) In addition, genetic sampling of one or a few captured young-of-the-year could be used to identify the Omy5 haplotype, further refining the life-history assignment while avoiding lethal take or disturbance of the redd during incubation. These follow-up snorkel surveys of identified redds could be incorporated into biweekly surveys by observing density of fry on the third survey after the redd was first identified.

The least ambiguous method for identifying maternal life history from a redd appears to be microchemical analysis of eggs or fry. Concentration ratios of certain trace elements (strontium, barium, calcium) in the egg yolk or otoliths (earbones) of the fry are indicators of whether the mother matured in freshwater or saltwater. Ratios of these elements systematically differ between freshwater and marine water, and their ratios in otoliths have been widely used and validated for determining the life history of adult freshwater and diadromous fish (Elsdon and Gillanders 2004, 2005, Brown and Severin 2009, Gibson-Reinemer et al. 2009). Similarly, the core of the otolith, typically laid down during the embryonic stage from materials in the egg, can be used to determine whether the mother matured (formed eggs) in freshwater or seawater (Volk et al. 2000, Zimmerman and Reeves 2000, Donohoe et al. 2008, Mills et al. 2012, McMillan et al. 2015), and so can be used as an indicator for maternal life-history. Although microchemistry of eggs themselves is less studied, it has helped identify maternal life-history for brown trout (*Salmo trutta*) in New Zealand (Waite et al. 2008, Kristensen et al. 2011, Gabrielsson et al. 2012).

Unfortunately, microchemical analysis means destructive sampling of eggs or fry. Removal of a small number of eggs or fry from a subsample of redds might have acceptable impacts,[7] but broader disturbance to redds while sampling is a significant concern. To be useful here, sampling methods for individual eggs or fry, with minimal impacts on the rest of the redd, would need to be developed, demonstrated, and permitted.

Thus, there is currently no satisfying method for assigning redds to steelhead versus Rainbow Trout mothers. For areas where redd surveys might be implemented, we suggest a two-pronged strategy moving forward.

First, implement first-order criteria based on redd size (area of pit and tailspill). Following Kendall et al. (2015), redd sizes corresponding to a mother with a fork length less than 35 cm would be classified as Rainbow Trout, a fork length greater than 50 cm as steelhead, and intermediate sizes as enigmatic. Riebe et al. (2014) estimated a rating curve for redd area as a function of female length (Figure 10), using a composite dataset on *O. mykiss*, *S. trutta*, and *S. salar* (Atlantic Salmon) from three different environmental settings in Britain (Crisp and Carling 1989). The curve suggests the following first-order rating scheme:

Redd Size < 0.95 m^2:	**Probably Rainbow Trout**.
Redd Size 0.95 – 2.2 m^2:	**Enigmatic**; could be either life-history type.
Redd Size > 2.2 m^2:	**Probably Steelhead;** but could be a large Rainbow Trout.

7 Eggs have low "reproductive value," meaning a lot of natural mortality occurs between the egg stage and stage of adult reproduction. Thus, the vast majority of eggs that would get sampled would die anyway prior to maturing into a reproductive adult, and so their removal has negligible effect on adult abundance. However, a sampling method that disturbs most eggs in a redd would have a much greater, non-negligible effect.

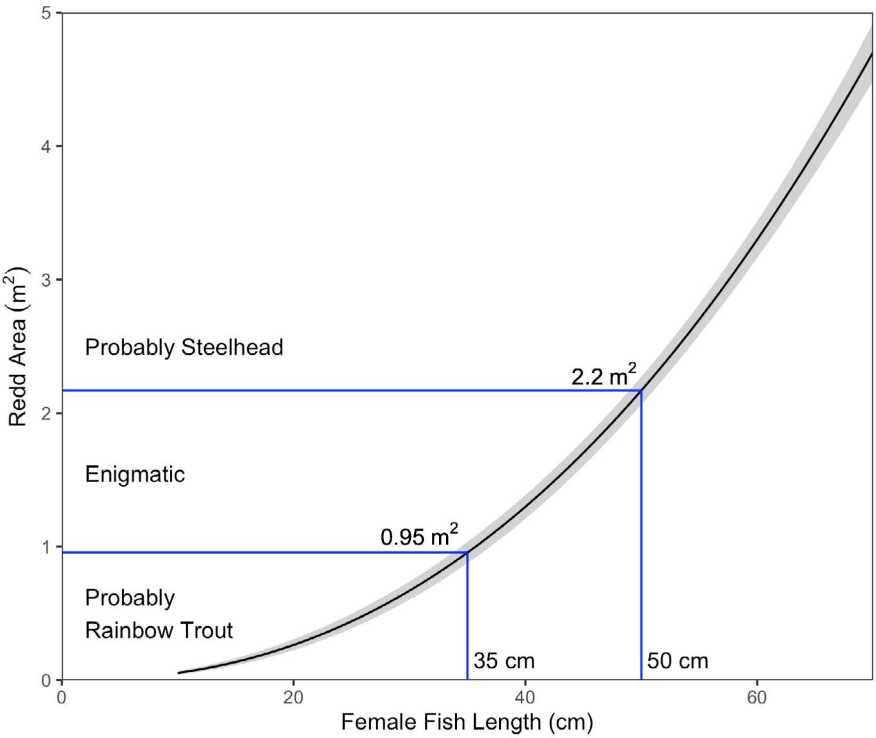

FIGURE 10. First-order scheme for assigning life-history to female creators of redds based on size. The curve was estimated by Riebe et al. (2014) from data on *O. mykiss*, *S. salar* and *S. trutta* collected by Crisp and Carling (1989) at three sites in the British Isles. Shading indicates uncertainty (50% c.i.) in the prediction line; prediction intervals are larger. Length criteria for female fish are general findings from a range-wide review of life-history diversity in *O. mykiss* (Kendall et al. 2015).

Second, if this scheme proves insufficient (e.g., a large proportion of redds are enigmatic, or the relationship between fish size and redd size is too noisy), develop methods to use microchemistry to assign maternal life-history to each redd. If successful, this methodology could eventually be integrated into redd surveys using a double-sampling approach: All redds would be measured for size, a random subset of redds would be sampled for eggs or fry, and the correspondence between microchemistry and size and/or timing would be combined with the complete size distribution of redds to develop a rating curve. A noisy relationship of fish size and redd size might require local tailoring of the rating curve if the noise stems from aspects of the channel geomorphology, geology, or flow regime.

Pros and cons. Advantages of redd surveys include incidental data collection on the spatial distribution of spawning adults and the characteristics of spawning habitat. Disadvantages include potential impacts on the species from microchemistry sampling, the problem of sparseness in large watersheds described earlier, and the possibility that an extended survey season is needed to estimate abundance of Rainbow Trout. Moreover, identifying the small redds of Rainbow Trout can be ambiguous, as they can be mistaken for footsteps, dabbling duck feeding depressions, or other disturbances. Another disadvantage is that anadromous fraction is estimated for mothers only, rather than for all adults. In general, anadromous fraction is expected to be lower for males than females due to contrasting mating strategies, so anadromous fraction from redd surveys is likely biased high. Potentially a calibration technique could be developed based on sex-ratio balancing (Ohms et al. 2019), but this requires methodological development.

Two key disadvantages relative to counting stations are the high labor needs to conduct biweekly redd surveys over an extensive area for 4-9 months, and the limited feasibility of implementing redd surveys in many areas, especially the large interior watersheds. In addition, it seems unlikely that the spawner-to-redd ratio, estimated in a particular population to convert redd abundance to spawner abundance, would be the same across all populations, and thus may introduce bias and underestimate sampling error. A key advantage, on the other hand, is implementation of redd surveys in coastal areas with numerous small watersheds independently connecting to the ocean. Rather than requiring a counting station on each small creek, a redd survey could be deployed at a broader spatial scale and calibrated by adult counts from a single counting station on a single creek system.

LIFE-CYCLE MONITORING STATIONS

Following Adams et al. (2011), we recommend a system of life-cycle monitoring stations (LCMs). These stations provide data to:

1. Estimate marine survival, an indicator for ocean conditions;
2. Estimate freshwater productivity (smolts per adult), an indicator of freshwater conditions;
3. Serve as a focal point for research on limiting factors of productivity; and
4. In areas with redd surveys, collect calibration data on spawner-to-redd ratios.

LCMs have three core elements for routine monitoring, and one optional element, each providing a particular kind of estimate (Table 12). Redd surveys are the optional element, only being required to calibrate redd abundances in areas where redd surveys are deployed at a broader spatial scale. Additional types of data collection, not covered here, would be deployed ad hoc for productivity research.

TABLE 12. Core elements of a life-cycle monitoring station.

Kind of Estimate:	Needed For:	Potential Methods:
Adult Abundance/ Anadromous Fraction	Marine Survival, Spawner-to-Redd Ratios, Freshwater Productivity	Resistance-Board Weir Sonar-Based Methods PIT-Tag Designs
Redd Abundance*	Spawner-to-Redd Ratios	Redd Survey*
Smolt Production	Marine Survival, Freshwater Productivity	PIT-Tag Designs Rotary-Screw Traps Fyke Nets
Vital rates of life-stages, cohorts, subpopulations	Productivity Research	Redds (Repeated Obs.) PIT-Tag Designs Growth-at-age Studies BACI Sampling Designs

* Not necessary in BPGs using counting stations for abundance monitoring, because estimates of spawner-per-redd are not needed to estimate abundance.

For estimating adult abundance and anadromous fraction, LCMs can draw from the same short-list of methods described earlier (Table 11). Note that even in areas monitored with redd surveys, a counting station or PIT-tag approach must be deployed in the LCM itself to estimate spawner-to-redd ratios.

Estimates of redd abundance are only necessary if the LCM is being used to calibrate a broad-scale redd survey. In general, the same methods are used as at the broad scale, but a greater sampling intensity for reaches, or even a 100% sample, will often be desirable in the LCM, to provide a tighter estimate of spawners-per-redd.

For estimating smolt production, two general methods are available: traditional downstream migrant traps and PIT-tag designs. Both methods usually involve some form of mark-recapture that allows estimates of trap efficiency or tag detection rate.

With downstream migrant traps, this is typically accomplished by tagging some fraction of captured smolts and releasing them upstream of the trap. The proportion of marked, released smolts that are trapped a second time provide data for estimating trap efficiency. Volkhardt et al. (2007) describe use of rotary-screw traps and inclined-plane traps for this purpose and provide mark-recapture estimators for various types of implementations. Typically, trap efficiency varies through time as a function of flow conditions, in which case a stratified design may provide tighter estimates (Bjorkstedt 2005).

PIT-tagging approaches, described earlier as a method for estimating adult abundance, can also be used to estimate smolt production. In this method, low-flow surveys are used to estimate total juvenile abundance in the populations, and a random or systematic sample of the juveniles is tagged during the survey. Smolting rate is then estimated as the proportion of tagged juveniles that emigrate the following spring. Smolt production is the product of juvenile abundance and smolting rate; an estimation procedure is given in Appendix B.

TABLE 13. Designs for Life-Cycle Monitoring Stations.

Design	Methods For:		
	Juveniles	Smolts	Anadromous Adults
Traditional	-	Trapping[1]	Weir
Acoustic Camera	-	Trapping[1]	DIDSON/ARIS
Existing Fishway	-	Trapping[1]	Visual at Fishway
Juvenile-Tagging	PIT-tagging[2]	PIT-monitoring[3]	PIT monitoring[3]
Smolt-Tagging	-	Trapping & PIT-tagging[1]	PIT-monitoring[3]
Tagging + Acoustic	PIT-tagging[2]	PIT-monitoring[3]	DIDSON/ARIS[4]
Tagging + Weir	PIT-tagging[2]	PIT-monitoring[5]	Weir
Tagging + Fishway	PIT-tagging[2]	PIT-monitoring[6]	Visual at Fishway

[1] With mark-release-recapture to estimate trap efficiency.
[2] In samples of short reaches during low-flow surveys.
[3] With multiple antennae to estimate detection rate.
[4] DIDSON or ARIS needed when annual number of returning tagged steelhead is less than 30.
[5] Possibly integrated into weir.
[6] If mounted on fishway, design is only feasible if all smolts are known to go through fishway. Use multiple antennae to estimate detection rate.

The various methods for estimating smolt production and anadromous abundance lead to a diversity of LCM designs (Table 13). Note that redd surveys can be added to any one of these if needed to complete the design. In our view, which design is selected for a given site is based on two considerations: (1) identifying suitable locations to implement particular methods, and (2) whether a PIT-tag design is also desired for integration with research or management efforts (p. 44).

Identifying a suitable location is key to the success of a given method. For example, successful siting of an acoustic camera involves finding a channel where (1) flow conditions discourage "milling" of adult fish and encourage steady upstream movement, and (2) the channel cross-section matches the shape of the sonar beam to prevent dead zones where migrants go undetected (Pipal et al. 2010a, Pipal et al. 2010b). Identifying methods that are appropriate for a particular site and thus provide good data is in our view more important than maintaining a uniformity of methods across all LCMs.

One of the juvenile-tagging designs might be desirable in a management or research context, to provide a flexible and powerful approach for identifying specific limiting factors on smolt production. If so, selecting this design to meet basic LCM functions can potentially form the basis for partnerships with local stakeholders, and may be a more economical use of funding.

One complicating factor stems from using marine survival as an indicator for ocean conditions. The marine survival of steelhead typically depends strongly on their size at ocean entry (Bond 2006, Arriaza et al. 2017), which in turn depends on freshwater or estuarine conditions. Therefore, some variation in marine survival may in fact reflect changes in juvenile growth patterns during the freshwater phase (Arriaza et al. 2017). Thus, we recommend that both methods for smolt production (juvenile-tagging, downstream-migrant traps) collect data on body size during handling of the fish. This would allow estimates of marine survival to be standardized by size-at-ocean entry, providing a more powerful index of ocean conditions.

IMPLEMENTING THE MONITORING PLAN

Up to this point we have reviewed the various viability indicators, methods for estimating them, and strategies to improve practicality and operational efficiency. Here we summarize the resulting monitoring scheme and how to implement it to generate a comprehensive set of indicators.

ESTABLISHING DATA MANAGEMENT

At the outset of implementation, a system for the flow and curation of data needs to be carefully designed. Adams et al. (2011) emphasized this point and, in reviewing this paper, one of the authors (P. Adams) emphasized it again as his most important feedback. Here we provide some broad principles for effective data management. Data management can be defined as an administrative process in which data is compiled, validated, stored, protected, and made accessible and trustworthy to data users. Here are some key principles:

Store raw data, not estimates. Once collected and validated (checked for errors), a given dataset is, in principle, forever the same; but the methods used to analyze it will likely change over time or different methods will be used for different questions. This simple fact dictates a data-management scheme where data is stored in as raw and uninterpreted form as possible (Figure 11). Data generated from field activities undergo QA/QC and standardization at the back end (prior to being added to the database), but estimation of indicators occurs at the front end, after querying the database and applying a standard estimation procedure to the resulting data. This allows diverse methods of analysis to be used without requiring the database to be changed or updated. Ideally, updates to the database itself would only involve addition of new data or correction of errors in old data.

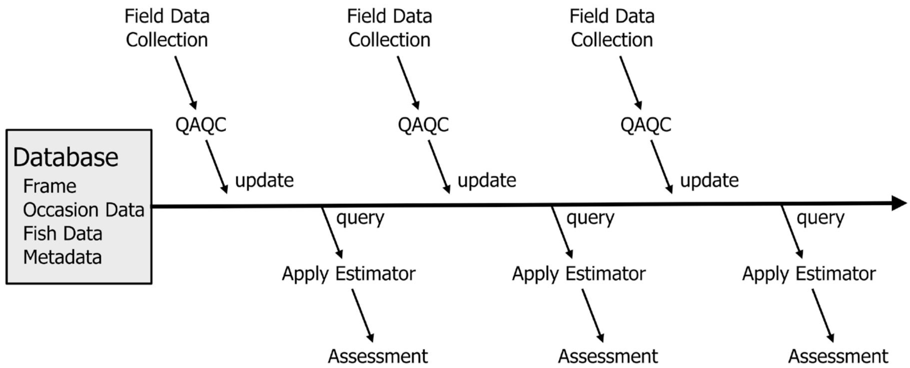

Figure 11. Schematic for data-management. A single database, with back-ups and standardized in third normal form, contains tables describing only the data: the sampling frame, individual sampling occasions, and individual fish observations. As new field data are generated, the database is updated; as new assessments are needed the database is queried and statistical estimators are applied. We do not recommend that estimates themselves be stored in the database. However, standard queries could be developed that automatically apply estimators to the most up-to-date dataset.

Only 3 basic types of data. For clarity, view the data as falling into one of three general types (Table 14, top): (1) numeric measurements or counts, (2) categorical data, including integer ID numbers such as PIT tag numbers, dates, and word codes that are descriptive yet standardized, and (3) descriptive data, such as field notes, which can have any kind of text. Only the first two types are typically used in analyses, but the third type contributes to interpretation of results and includes metadata that describes limitations or valid interpretations for a given dataset. Categorical data include word codes such as "anadromous adult" or "smolt" that can have ambiguous meaning if not carefully standardized and defined in metadata. Since categorical data are often used to group data into like types for analysis, they need to be standardized to match the existing database as a step during QA/QC.

TABLE 14. Fundamental data structures.

Entity	Description	Keys
KINDS OF DATA		
Numeric	Measurements or counts	Not suitable for keys
Categorical	ID numbers, dates, PIT-tag numbers, standardized word codes: taking one of a pre-specified and defined set of values	Suitable for keys
Descriptive	Freeform text, such as field notes	Not suitable for keys
KINDS OF TABLES		
Locations	Discrete places that can be sampled, usually defined in a finite sample frame	Unique Reach ID for each sample unit
Sampling Occasions	Individual sampling occasions where a particular location was sampled on a particular day	Reach ID + Date or Date-Time identifying time and place of occasion
Fish	Encounters with individual fish in which they have been sampled and measured in some way	Fish ID defining 1 sampling event for 1 fish
SOP	Standard operating procedures, defining the meaning of data such as units of measurement, definitions of categories, protocols for measuring, protocols for sampling	Linked to particular types of sampling occasions or field names for numeric, categorical, and descriptive data
KINDS OF RETRIEVAL		
Data Query	Extraction of a subset of rows from a table or set of linked tables	Keys may be used to extract linked records in other tables
Standard Report	Extraction of data and application of recommended or standard estimators	Keys may be used to extract linked records in other tables

Normalization. Standard practice is to establish a database structure that maintains data integrity during updates. Technical references such as Watt and Eng (2014) emphasize the importance of this principle, and provide a set of steps for achieving it, known as "database normalization" (see Watt and Eng 2014, chapter 12). The point of this process is to eliminate redundancies and inconsistencies in the database so that internal contradictions cannot develop over time. A database that meets these criteria is said to be in third normal form. Normalization typically converts a dataset into a set of distinct tables having no redundancies across them and linked by a precisely defined relational structure using keys. Keys are unique identifiers, such as ID numbers, that link together corresponding data in different tables. For example, each record of an individual fish would have a key linking it to a particular sample reach where the fish was captured.

Only 4 basic types of tables. In our experience, after normalization, fish and habitat data collected from river systems tend to sort into four general types of tables (Table 14, middle). Although one might think that only fish data are generated, these fish data are only interpretable in the context of three other types of tables describing the location of the fish, the sampling occasion in which it was observed, and a description of field methods used to collect data. In a normalized database, all three of these types of tables have a one-to-many relationship with data for fish, as well as with each other. Of key importance, but often overlooked, is what we refer to as occasion data, describing the exact sampling occasion where the fish was observed or captured. The occasion data, which include a description of when, where, and how a particular sampling event occurred, are key to determining which estimators can be validly applied to the data queried from the database.

Centralization, version control. To prevent confusion the data should be stored in a central database with version control. Version control is a comprehensive record of all changes made to the database through time, with ability to roll the data back to a previous version if it gets corrupted. Centralization allows version control to be uniquely defined to prevent internal conflicts from developing in different copies of the database. Centralization still allows for cloning—identical copies of the database distributed to various users—while maintaining overall database integrity.

Data retrieval. Finally, we recommend an online interface for the database that can serve up data in at least two general formats: as data queries and as standard reports (Table 14). A data query provides a way for a user to specify a subset of records to extract from the database, to which she applies analytical methods of her own choosing. A standard report provides a way for the user to extract records and apply a standard estimation method. In this case the analytical method is selected by the CMP Science Team for the coastal monitoring program, based on their best understanding of how the data were collected, what the underlying assumptions were, and the appropriate analytical framework. A set of standard reports for the various metrics and designs would thus serve up processed estimates representing a consensus of the Science Team on the best approach for that particular watershed and metric. Although such standard reports can be updated from time to time as methods evolve, at any point in time it will ensure a single "best current estimate."

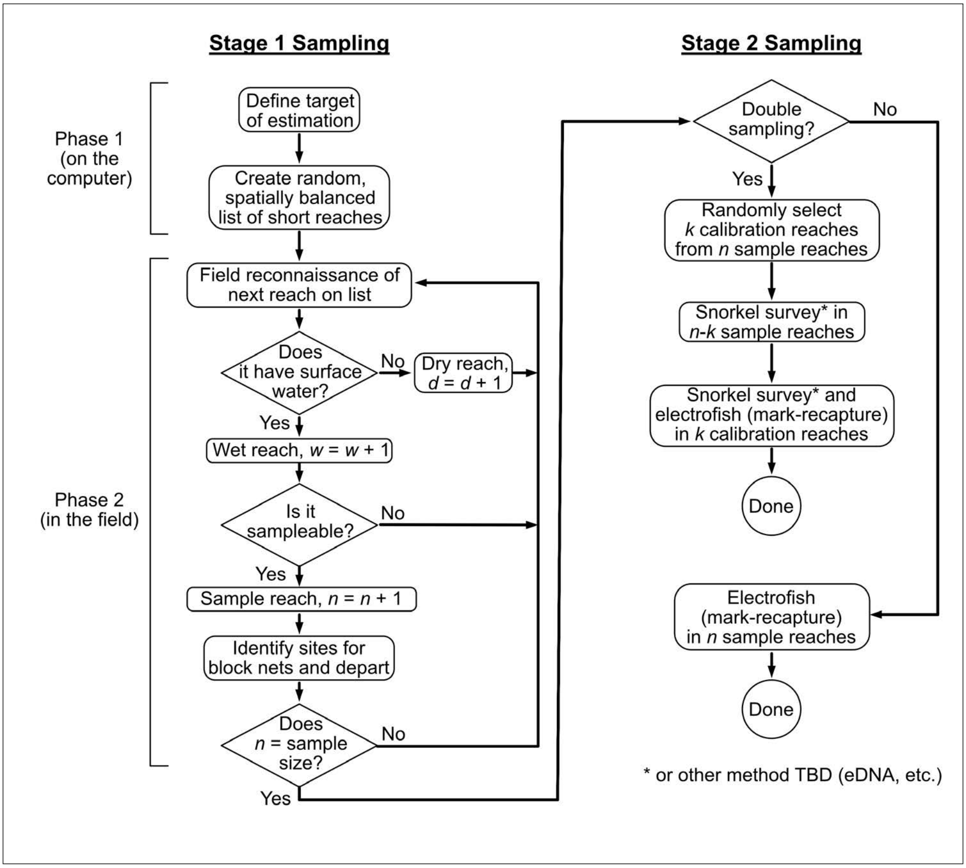

FIGURE 12. Workflow for low-flow surveys.

IMPLEMENTING LOW-FLOW SURVEYS

Low-flow surveys use well-established methods and can begin as soon as there is a sampling frame and a field crew to sample it. A workflow for the recommended approach is sketched in Figure 12, and estimators for the various indicators are given in Appendix B. Because the targets of estimation are the various biogeographic population groups (BPGs; Figure 6), the sampling efforts in different BPGs can ramp up independently of one another. In general, the number of sample reaches *n* needed to achieve a certain precision will be comparable across BPGs, though perhaps slightly larger in the BPGs with greater geographic extent, since these are likely to have more heterogeneous conditions. Below we discuss precision in terms of the coefficient of variation (CV), in which the standard error of an estimate is expressed as a percent of the estimate itself. We use some existing datasets to identify the approximate *n* needed to achieve a CV of 30%.

Number of reaches for Design 1 electrofishing. To get a sense of the number of sample reaches (wet, sampleable) necessary for each target, we considered data from two years of low-flow surveys in the Carmel River (Table 15). These used electrofishing to estimate fish density but differ slightly from what is being proposed here: reach-sampling was geographically stratified, and fish-sampling used depletion rather than mark-recapture. However, the underlying sources of variance should be similar to what would be obtained using our recommended design.

TABLE 15. Density estimates from Carmel River for 2017, 2018.

Est.[1]	Description	2017	2018
n	Number of wet reaches sampled	17	14
d	Number of dry reaches observed	6	9
$\hat{f}_w$	Fraction of wet reaches	0.73	0.59
$\hat{V}_{\hat{f}w}$	Estimator variance	0.022	0.033
	CV	20%	31%
$\hat{D}_w$	Fish density in wet reaches (m^{-2})	0.174	0.308
$\hat{V}_{\hat{D}w}$	Estimator variance	0.00731	0.00755
	CV	49%	28%

[1] From stratified-random sampling / depletion-electrofishing rather than SRS / mark-recapture. Sample reaches drawn using GRTS, stratified by geographic subregions and omitting Los Padres National Forest due to lack of samples. Estimates made using appropriate estimators, except optimal ***k*/*w*** was treated as simple random sample.

In one year, the target CV of 30% was achieved for fish density with just 14 wet reaches, though in 2017 the CV was about 50%, due more to a lower estimated density than a higher variance. These results suggest a sample of $n = 20$ or 30 wet reaches per BPG may be enough, at least as a starting point. Given nine BPGs in the southern monitoring area and $n = 20$, this translates to 180 wet reaches sampled annually during the low-flow season from August – October each year. Assuming one field crew can sample one reach per day, and a total of about 60 work days during this period, this translates to a need for three field crews kept busy for a three-month period annually. This is a rough estimate that only accounts for fish-sampling in the sample of wet reaches. Additional reconnaissance effort is required in July and August for the initial screening for wet reaches, and travel to remote reaches may require additional work days.

Number of reaches for Design 2 snorkel counts. In general, we expect that electrofishing is the preferred method, but snorkel counts may be preferred in certain BPGs for non-statistical reasons: They can be deployed even when water temperatures are greater than 18° C and have a smaller impact on fish and other stream organisms. In such cases there is still a need for double-sampling with an electrofisher at a subsample of reaches, to provide genetic samples and to calibrate the snorkel counts. Potentially this could be a very small fraction of sampled reaches but would need to be (1) randomly sampled from the set of snorkeled reaches,

(2) sufficient to capture the heterogeneity in visual conditions that affect snorkel counts (including outliers), and (3) sufficient to provide a genetic sample. In addition, the snorkel counts need to include procedures for identifying and counting adult Rainbow Trout, using fork length to identify adults as described on p. 30.

What is the optimal proportion of calibration reaches (*k*/*n*)? To answer this we considered six double-sampled short reaches in Arroyo Seco and its tributaries in Monterey County (Boughton et al. 2009 and unpublished data). In this case the calibration reaches were selected for convenient access from a larger stratified-random sample of reaches where only snorkeling was performed. These six reaches showed a very tight relationship between snorkel counts and an estimate from depletion-electrofishing (Figure 13). The reaches had high detection rates for the snorkel counts, and a ratio-variance that was small relative to the variance of the electrofishing estimates (Table 16), suggesting that the optimal ratio was to double-sample about a quarter to a third of sample reaches, depending on cost ratio (Table 16, bottom). Here, cost ratio is quantified as the number of short reaches that can be snorkeled per day by a crew that could double-sample one reach per day. One can imagine a plan in which a field crew works together to sample one calibration site per day, and then splits into smaller subgroups to snorkel two to five of the remaining

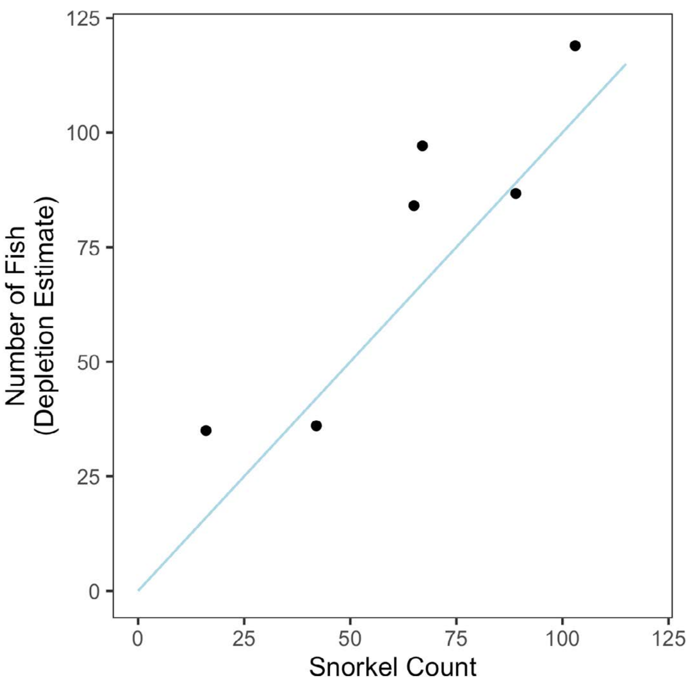

FIGURE 13. Association between snorkel-counts and estimates of fish abundance at six reaches from Arroyo Seco and its tributaries in Monterey County.

sites per day (Table 16, bottom. For example, if the field crew can snorkel-survey four short reaches per day (cost ratio 1:4), then the Arroyo Seco data suggest that the optimal $k = 0.26n$.

TABLE 16. Double-Sampling in Arroyo Seco with Snorkel Counts.

Parameter.[1]	Description	Estimate
k	Number of double-sampled reaches	6
r	Ratio	1.20
$(1/r)$	Detection probability	0.83
$\widehat{Var}(r)$	Ratio Variance	239
$\widehat{Var}(\hat{y}_k)$	Among-Reach Variance of Estimates	1150
	Double-sampling?[2] (optimal k/n)	
	Cost ratio 1:2	0.36
	Cost ratio 1:3	0.29
	Cost ratio 1:4	0.26
	Cost ratio 1:5	0.23

[1]From six double-sampled reaches in Arroyo Seco and tributaries, treated as simple random sample.

[2] Cost ratio = number of calibration reaches per day (snorkel + mark-recapture): number of snorkel-only reaches per day. k = number of calibration reaches, n = total number of reaches sampled.

However, we would expect this optimal ratio to be sensitive to outliers. Ours was a very small sample unlikely to pick up the occasional outlier with low visibility or complex habitat, either of which would weaken the tight relationship in Figure 13 and require a larger proportion of calibration reaches.

Genetic Sampling. The purpose of routine genetic monitoring is to track the proportion of Omy5 chromosomes that have the "A" haplotype, which we argue is a lagging indicator for the sustained participation of anadromous fish in the breeding population. During low-flow surveys, a subset of fish at each electrofishing reach has tissue samples taken, and these are analyzed to determine whether each fish has 0, 1, or 2 chromosomes with the A haplotype. These data are then used to estimate the proportion of A haplotypes across all fish in the target of estimation. An estimator for the proportion is described in Appendix B.

An important point is that for calculating the overall estimate, the proportion of the haplotype at any given sample reach needs to be weighted by the estimated number of fish at the sample reach. This estimate is made using the mark-recapture approach described previously, either at all reaches, or at the calibration reaches if the snorkel design is used. For the fish handled during the mark-recapture procedure, a subset is sampled for tissue to conduct the genetic analysis.

The size of the subset can vary from reach to reach—it can be all fish handled, or a random subsample. In addition, tissue samples do not need to be taken at every electrofishing reach for the estimator to be valid, but the subsample that is used should be selected randomly from the sample of wet reaches. In general, though, the omission of reaches should be avoided: we expect that genetic composition varies substantially among reaches and the variance of the estimator will benefit from collecting tissues across as many sample reaches as practical. If the limiting factor is the total number of genetic samples that can be analyzed, it is better to reduce the samples taken per reach than to reduce the number of reaches.

COMPREHENSIVE ABUNDANCE MONITORING

Three key issues affect the ability to monitor comprehensively: feasibility of redd surveys, the large number of discrete coastal populations in the southern area, and the need to monitor Rainbow Trout.

Although we reviewed both redd surveys and counting stations as potential methods, we feel that redd surveys are not feasible in many areas. Redd surveys are clearly not even remotely feasible in large inland watersheds with extensive wilderness that is inaccessible to field crews in the wet season. In more accessible areas, redd surveys may still be problematic if flow regimes typically preclude either access or effective sampling in substantial parts of these stream networks for substantial portions of the migration season.

The large number of steelhead populations in the southern area also poses a challenge: 81 populations total and 34 backbone populations. To improve operational efficiency, we recommend that a subset of populations be monitored annually—one per biogeographic area. The rest would be monitored on a rotating schedule of every two to four years. This scheme involves the concept of portable counting stations that would be deployed at a particular stream for a season at a time and then moved to the next stream in the rotation the following season. As such, it is not compatible with the PIT-tag approach to monitoring steelhead abundance, since smolts or juveniles tagged in a particular year would return as adults over a multi-year interval due to varying years spent in the ocean.

Finally, the implementation scheme must include a Rainbow Trout component. Wherever a counting station is deployed for a season to count steelhead abundance, a low-flow survey would also be deployed to estimate Rainbow Trout abundance. In practice this simply means additional reach sampling in the low-flow surveys already being conducted. That is, a sample of ~20 or 30 reaches is drawn annually for each biogeographic population group, as described earlier. Then, additional reaches are sampled for the specific populations being monitored that year. Due to the problem with distinguishing the small redds of Rainbow Trout from other sorts of bed disturbances, we also recommend that areas using redd surveys also use the low-flow surveys to estimate Rainbow Trout abundance.

Below we sketch an implementation scheme that accounts for these challenges (summarized in Table 17), though we recognize it to be preliminary and likely to benefit from further refinement.

Table 17. Proposed Backbone Populations and Scheme for Abundance-Monitoring.

Biogeographic Population Group	Target of Estimation	Steelhead Method	Rainbow Trout Method	Rotation
Interior Coast Range	Pajaro River	Counting Station	Low-Flow Survey	3 yr
	Salinas River pops.	Counting Station	Low-Flow Survey	Annual
Carmel	Carmel River	Counting Station	Low-Flow Survey	Annual
Big Sur Coast	San Jose Creek	Counting Station	Low-Flow Survey	3 yr
	Little Sur River	Counting Station	Low-Flow Survey	3 yr
	Big Sur River	Counting Station	Low-Flow Survey	Annual
SLO Terrace	San Carpoforo Cr.	Counting Station	Low-Flow Survey	4 yr
	San Simeon Creek	Counting Station	Low-Flow Survey	4 yr
	Santa Rosa Creek	Counting Station	Low-Flow Survey	4 yr
	San Luis Obispo Cr.	Counting Station	Low-Flow Survey	Annual
	Arroyo Grande Cr.	Counting Station	Low-Flow Survey	4 yr
Mt. Arido Highlands	Santa Maria River	Counting Station	Low-Flow Survey	3 yr
	Santa Ynez River	Counting Station	Low-Flow Survey	3 yr
	Ventura River	Counting Station	Low-Flow Survey	Annual
	Santa Clara River	Counting Station	Low-Flow Survey	3 yr
Conception Coast	Carpenteria Creek	Redd Survey	Low-Flow Survey	Annual
	Gaviota Creek	Redd Survey	Low-Flow Survey	Annual
	All other creeks	Redd Survey	Low-Flow Survey	3 yr
Santa Monica Mtns.	Topanga Canyon	Redd Survey	Low-Flow Survey	Annual
	All other creeks	Redd Survey	Low-Flow Survey	3 yr
Mojave Rim	San Gabriel River	Counting Station	Low-Flow Survey	4 yr
	Santa Ana River	Counting Station	Low-Flow Survey	4 yr
	Los Angeles River	Counting Station	Low-Flow Survey	4 yr

Biogeographic Population Group	Target of Estimation	Steelhead Method	Rainbow Trout Method	Rotation
St. Catalina Gulf Coast	San Juan Creek	Counting Station	Low-Flow Survey	4 yr
	San Mateo Creek	Counting Station	Low-Flow Survey	4 yr
	San Onofre Creek	Counting Station	Low-Flow Survey	4 yr
	Santa Margarita R.	Counting Station	Low-Flow Survey	Annual
	San Luis Rey River	Counting Station	Low-Flow Survey	4 yr
	San Dieguito River	Counting Station	Low-Flow Survey	4 yr
	San Diego River	Counting Station	Low-Flow Survey	4 yr
	Sweetwater River	Counting Station	Low-Flow Survey	4 yr
	Otay River	Counting Station	Low-Flow Survey	4 yr
	Tijuana River	Counting Station	Low-Flow Survey	4 yr

Interior Coast Range and Carmel. Redd surveys are not feasible in most of the Arroyo Seco drainage due to inaccessibility during the wet season. In more accessible areas, such as the Carmel River watershed, redd surveys are still problematic due to the flow regime (K. Urquhart, pers. comm.). We therefore recommend annual counting stations in the Salinas and Carmel Rivers. The Pajaro would also be monitored with a counting station, either annually or as part of a three-year rotation with two creeks in the Big Sur Coast (Table 17).

Big Sur Coast and SLO Terrace. Streams of these two biogeographic regions are generally more accessible in winter, but still have significant inaccessible areas in their upper reaches, such as the upper Big Sur River (M. Michie, pers. comm.). It is possible to conduct redd surveys in a GRTS sample while omitting inaccessible areas, but then the estimates formally apply only to the accessible reaches, and thus omit a potentially significant proportion of the population. It is possible to expand the estimate numerically to the inaccessible region, but this involves an untestable assumption that the accessible and inaccessible reaches have identical statistical properties. This would only be justified when the unsurveyed proportion is relatively small, and not disproportionately productive relative to the surveyed area. We conclude it would be best to deploy counting stations in backbone populations and propose a mix of annual sites and rotating sites in Table 17. Depending on site conditions a different set of streams selected as backbone populations might make sense.

Conception Coast and Santa Monica Mountains. These two biogeographic regions have numerous small creeks that are readily accessible throughout the winter.

The creeks are short enough that redds can generally be surveyed across an entire creek in single day (i.e. from tidewater to limits of anadromy). Steelhead abundances are expected to be extremely low in individual creeks—often in the single digits or zero (Dagit et al. 2018)—but across the entire set of creeks may add up to a demographically significant number.

These features suggest a modified redd survey would be feasible and more operationally efficient than a series of counting stations. Based on input from regional staff, we recommend two modifications to the basic design of the redd survey. First, rather than focus on backbone populations, the entire set of creeks will be surveyed in each biogeographic region. Second, redd surveys will be conducted in entire creeks, on a rotating schedule, rather than in samples of reaches as in a conventional redd survey. In short, some creeks will be entirely surveyed every year (annual census); the rest will be surveyed every three years (rotating census). The proportion of creeks with an annual census versus a rotating census can be set based on available funding and logistical considerations, but once determined should be retained over the long-term. Likewise, the identity of creeks censused annually should not be altered over time due to changing conditions, as this defeats the overall CMP goal of generating design-based, unbiased estimates. In each biogeographic area, one of the creeks receiving an annual census should also have a counting station, to calibrate the redd surveys.

To monitor Rainbow Trout, it makes sense to expand the low-flow surveys described previously to mirror the censusing scheme used for steelhead abundance. That is, the sampling plan first develops a standard GRTS sample of short reaches from the entire biogeographic region. Then, each year the plan is expanded to also census all reaches within the creeks being surveyed for redds.

In this hybrid approach, we recommend that creeks still be divided into short and long reaches with lengths consistent with other areas, and that raw data still be reported reach-by-reach rather than creek-by-creek. This maintains consistency with data collected in other regions.

Monte Arido Highlands. This biogeographic region has four relatively large river systems: Santa Maria, Santa Ynez, Ventura, and Santa Clara Rivers. We recommend that all four be surveyed annually for adult abundance using counting stations or redd surveys, whichever is most suitable for successful implementation. Abundance of Rainbow Trout is estimated from low-flow surveys annually in each system.

Redd surveys have been successful in the past in the Ventura system and can continue there, either using a spatially-balanced random sample of reaches (i.e., using a GRTS sample draw based on reaches) or as a complete annual census, with the choice to be determined by regional managers and staff. The selected approach should be consistently used over the long-term.

We expect that counting stations will prove more feasible in the other three systems, based on the inaccessibility of substantial portions of the Santa Maria and Santa Clara systems that cannot be surveyed biweekly in winter, and the past successful implementation of counting stations in the Santa Ynez system.

Mojave Rim and Santa Catalina Gulf Coast. The two southernmost biogeographic regions consist of large arid watersheds draining various mountain ranges set back from the coast. Rainbow Trout are sparsely distributed, and steelhead are apparently rarer still, though not much surveyed at this point. Accessibility for field crews may be a problem in some mountainous areas, but the larger issue with redd surveys is the likely sparsity of redds within quite extensive stream networks. Due to this sparsity, a large effort could go into redd surveys that produce zero redds in the sample, but with a low confidence that the true abundance is zero.

We therefore recommend counting stations be deployed, with Santa Margarita River monitored annually and the remaining 12 basins monitored on a rotating scheme (Table 17). As with the other biogeographic regions, once the assignment of basins to different sampling rotations is complete, it should be retained over the long term regardless of changing conditions. Rainbow Trout are monitored by augmenting the low-flow surveys in the same rotation, as described earlier.

Siting Annual Stations. For operational efficiency, the backbone populations selected for annual counting should be suitable for the site of a future life-cycle monitoring station (LCM). Key considerations for how to choose include:

- Siting of LCMs also requires good landowner and physical access, and security.
- LCMs require the typical abundances of adults be sufficient for good estimates of marine survival. Abundances less than 20-30 adults per year will likely give poor precision, although even these modest abundances may be difficult to find in many parts of the southern monitoring area. Such abundances are also needed where redd surveys are planned, to support estimation of spawner-per-redd ratios.
- Siting of LCMs in watersheds with extensive ongoing or planned habitat restoration is useful, because population response to these efforts can be monitored as a basis for productivity research and adaptive management.
- Suitable locations for infrastructure (Acoustic camera deployments, weirs, PIT-tag monitoring stations, depending on design).
- Site conditions that fit well with the pros and cons of one of the LCM designs in Table 11 and Table 13.
- Suitability for intensive low-flow surveys to estimate abundance of Rainbow Trout.

SMOLT PRODUCTION AND LCMS

The final implementation step is to begin monitoring smolt production at populations where annual counting stations have been established. How many and where depends on site conditions appropriate for each method; whether the number and proportion of outmigrants that can be captured is sufficient to estimate production; and how consistent marine survival is among different populations. If marine survival is relatively consistent among stations, fewer would need to be deployed to generate an overall estimate of marine survival for the region.

The two most promising methods for producing estimates of smolt production are outmigrant-trapping and juvenile-tagging methods. Most trapping methods do not seem very promising for the south-coast area due to the flashy, debris-laden hydrograph. The most promising is rotary screw-trapping, an established method with standard operating procedures, but even so its feasibility for use in the southern area has been questioned. According to Volkhardt et al. (2007), "streams and rivers exhibiting a flashy hydrograph are very difficult to trap due to high fluctuations in flow conditions and debris loads." During the low-flow portion between storms, or in smaller creeks or low-gradient channels, the water velocity may be insufficient to operate the trap, or to trap a strong swimmer such as a steelhead smolt. Clearly some trials need to be done, with careful attention to site selection and points raised by Volkhardt et al. (2007). It seems likely that rotary-screw trapping can be adapted to the southern situation to some degree, but also that site constraints may limit its flexibility.

The other promising method we identified is the integration of PIT-tagging into low-flow surveys. Table 18 summarises the results from a trial of this method in the Carmel River. Seventeen wet reaches were randomly selected in fall 2017, with six dry reaches identified in the process. Abundance estimation was not quite the same as recommended here, using depletion-elecrofishing and a stratified design for sampling reaches, but the results should be comparable. Fish captured at the sample reaches were implanted with PIT tags if they were of sufficient size to carry a tag (FL > 70mm); additional tags were implanted in fish captured at index reaches for a total of 1810 fish tagged.

At the PIT-tag monitoring station, the detection rate for tagged fish at each antenna was only about 50%, and the estimated smolting rate was fairly low (0.068 smolts per parr), but even so the PIT-tagging approach achieved a CV of only 8% for the estimate of smolting rate. The CVs for the estimates of juvenile abundance and smolt production were higher (26% and 27% respectively), but still below the target CV of 30%. This suggests that smolt production can be estimated with a relatively modest effort of about ~20 wet reaches sampled and 2000 fish tagged, even if smolting rate is low. In populations with higher smolting rates, a more modest tagging effort would probably be sufficient. However, this method will not be feasible in populations where fish are too sparse to capture and tag 2000 fish.

TABLE 18. Smolt production in Carmel River 2018, estimated from PIT-tagging.

Parameter	Value	Description
n	17	Number of wet reaches sampled in 2017
d	6	Number of dry reaches observed
$\hat{f}_w$	0.73	Fraction of wet reaches
N	356	Total number of reaches available for sampling
	1415	Number of fish handled during low-flow survey
T_J	59032	Estimate of total juvenile abundance
σ_{TJ}	15465	Standard error of estimate
CV	26%	Coefficient of variation
t	1810	Number of fish tagged in 2017 low-flow survey (incl. index reaches)
f	63	Number of tagged fish detected at first antenna in spring 2018
l	65	Number detected at second antenna
b	33	Number detected at both antennae
s	0.068	Estimated smolting rate in spring 2018 (smolts per parr)
σ_s	0.0055	Standard error
CV	8%	Coefficient of variation
S	4019	Estimated smolt production in 2018
σ_S	1105	Standard error of estimate
CV	27%	Coefficient of variation

METHODS REFINEMENT

The methods and implementation sketched out above would benefit from further refinement, in consultation with the leadership of the Science Team for the Coastal Monitoring Plan. Here we describe the key needs for refinement, roughly in order of priority.

Counting Stations. The optimal design for counting stations is still not certain. In our view, sonar cameras and PIT-tagging approaches have the most promise. However,

- Sonar cameras require a practical solution to the species-apportionment problem. Promising avenues of development are use of weekly netting sessions to assign species proportions (Denton et al. 2014, Denton et al. 2015), and pairing sonar cameras with visual-light cameras (Killam 2012). However, existing designs need to be adapted to and tested within the typical field conditions of the southern monitoring area, as well as validated for very small run sizes and potential confusion with nonnative species. Repeated, quantitative sampling of eDNA at the monitoring site is also a promising method for species apportionment but is completely undeveloped at this point.

- Field deployments of sonar cameras must be able to effectively record fish movements as flows rise and fall during episodic flow events. Infrastructure such as bankside tracks or ramps may typically be necessary. An A-frame sled design has been effective in Ventura and Topanga, and is easily assembled, easily anchored, and can be moved and adjusted effectively. We recommend that this and other promising designs be further developed and described, ideally leading to a universal design that can be adapted to a broad array of site conditions. The design should support easy assembly and disassembly of counting stations, so that individual stations are portable and can be seasonally rotated to new sites as needed.
- As currently implemented, designs based on sonar or visual-light cameras require personnel to review imagery in a tedious and labor-intensive process. In this age of facial recognition software and other technological advances, it should be feasible to develop automated post-processing methods (for example, see Tarayama et al. 2019, Yin et al. 2020). The engineering of such methods may require nontrivial investment of time and money to develop but would likely greatly increase operational efficiency.
- Rotary screw traps may have very low capture efficiency for steelhead smolts (D. Baldwin, pers. comm.) or may be difficult to successfully operate in the episodic flow regime of the southern area (C. Hamilton, pers. comm.). The method needs to be either adapted to meet these challenges or abandoned in favor of the PIT-tagging approach to monitoring smolt production.
- Practical, effective operating procedures need to be developed for PIT-tag designs. In particular, tagged smolts are a challenge to detect, but either floating antennae (Barlaup et al. 2018) or the slack-loop design developed by ODFW appear to be suitable solutions. Pass-over antenna may work in areas where streams stay shallow during migration season. However, details of successful deployment and operation still need to be worked out and synthesized into standard operating procedures. Generally, PIT-tag monitoring stations will need to be deployed on river channels upstream of the estuary, and so there is also a need to develop data-screening conventions for the tag detections to distinguish between fish moving into the estuary versus expressing true anadromy of one or more years at sea.
- In some settings, weirs may be the preferred solution for counting adult steelhead but will likely provide poor data unless an engineering solution can be found to the "weir problem" (Figure 9).

In the near term, sonar cameras can be deployed in most river systems but will have problems with interpretation until the species-apportionment problem is resolved. We recommend the CMP Science Team give high priority to the systematic development of effective sonar and PIT-tag designs, including publication of a report for practitioners, describing how to design, site, deploy and operate the two designs.

Field Data Refinements. For the various fish-sampling methods, a number of specific field methods need to be refined:

- For low-flow surveys, we proposed first-order field criteria to distinguish adult Rainbow Trout from *O. mykiss* juveniles (p. 29). However, we recommend criteria be further validated or refined to develop a better-supported scheme for quantifying adult Rainbow Trout. Three potential schemes would be (1) to establish numbers of mature Rainbow Trout by dissecting gonads (Ohms et al. 2014) from a subsample of fish to develop a rating curve (maturity as a function of size); (2) to use handheld ultrasound devices in the field to identify mature females (Bangs and Nagler 2014); or (3) to use genetics to identify the characteristic body size at which sex-ratios diverge from 50:50, which is an indicator that fish above this size have foregone smolting and will likely mature into Rainbow Trout (Kelson et al. 2019, Pearse et al. 2019). Any one of these methods could potentially be used to develop locally-tailored rating curves for maturity as a function of body length and perhaps other field criteria such as condition factor. Samples for developing rating curves would need to be drawn from anadromous waters because residual populations above impassable dams are likely under selection for distinctly different maturation schedules. It is not clear how much variation to expect in maturation schedules among different populations of steelhead, but ideally the study design used to develop rating curves would account for this possibility as well. The first method—dissection of gonads—is the most definitive but also requires lethal sampling and so can only be applied sparingly; the others could potentially be integrated into routine sampling depending on time and expense.
- Downstream migrants in the spring are typically a mixture of smolts heading for the ocean, and parr heading for downstream freshwater or estuarine habitats (e.g., Hayes et al. 2011). For estimates of smolt production, field criteria (e.g., IEP 1998) are needed to distinguish the smolts from the parr. We recommend validation of appearance criteria using assays of gill Na+, K+-ATPase which is associated with saltwater tolerance (see Appendix C for details). We recommend that methods used for estimating smolt production be modified to allow for estimates of size-at-ocean-entry as well. This would allow estimates of marine survival to be standardized by size-at-ocean-entry, providing a more powerful index of ocean conditions.
- For low-flow surveys, a specific method needs to be clarified for estimating fish density and adult abundance in pools too deep to electrofish, most likely by adding a snorkeling component for individual habitat units within an

encompassing sample reach that gets electrofished elsewhere. Another need is to fine-tune the mark-recapture field procedures to ensure that they support the equal-catchability and random-mixing assumptions of the estimator. Key issues are the method used to redistribute fish from the first pass into the sample reach at extremely low flows, and the wait time necessary to allow them to mix with the uncaptured fish.

Lagoon Sampling. Our recommendations for diversity monitoring include sampling lagoons to determine density of rearing juvenile steelhead. Normally this would use a two-stage sampling scheme, similar to that used for low-flow surveys in streams. In stage one, habitat units are randomly sampled from the coastal portion of the sampling frame (estuaries and coastal channels), stratified by biogeographic area. During stage 2, fish are sampled from suitable units to determine occurrence of steelhead. A specific field method and statistic framework needs development. The field method would likely be based on seining or snorkeling, though estuaries are typically too murky for the latter technique. A starting point is the combination of seining and mark-recapture used by Jankovitz and Diller (2019) and Frechette et al. (2016).

Hydraulic Egg Sampling and Microchemistry. Earlier we recommended redds be assigned to steelhead versus Rainbow Trout using redd area, but this first-order scheme leaves a category of enigmatic redds whose size could be due to either life-history type. Depending on the how often redd surveys are used and how often these enigmatic redds turn up in datasets, it might prove desirable to develop a less ambiguous method. At this writing the most promising method is to sample individual eggs or fry from redds and analyze microchemistry as described earlier, but this is potentially very invasive. The method of hydraulic sampling (Collins et al. 2000, Berejikian et al. 2011) could potentially be modified to minimize incidental mortality of non-sampled eggs and fry. Is there a way to sample a small number of eggs, without trampling the redd, introducing disease, dislodging non-target eggs, altering internal hydrology and oxygen supply of the redd, or causing other unwanted impacts? One way to minimize risk associated with hydraulic sampling would be to develop a carefully crafted low-risk field method, and have a specific team that is highly trained in this type of sampling and have that team responsible for sampling eggs throughout the southern area.

Ongoing Methods Development. We expect that new field methods will be invented in the coming years, and hope that practitioners will accommodate advances in data collection while retaining the overall monitoring framework laid out here. In our view this involves retaining the framework of VSP parameters, spatially balanced sampling of long and short reaches, and stratification by targets of estimation (BPGs, backbone populations), while allowing field methods and statistical models to evolve as needed. Promising field methods include sampling eDNA for estimating species occurrence and density; in the latter case the concentration of eDNA in the water is an indicator of fish density and would be calibrated using a double-sampling design with electrofishing at the calibration reaches. Other promising methods are drone-based remote sensing platforms for surveying redds and habitat; other advances in remote sensing that may allow for identification of wet versus dry reaches from satellite imagery; and ongoing improvement in the performance and minimum sizes of various electronic tagging technologies. Many of these promising methods could lead to greater operational efficiency by reducing the labor needs or environmental constraints for field sampling. As such field methods are developed and integrated into the existing sampling scheme, they would be applied concurrently for a number of years with the method they are meant to replace, so that any difference in bias between the old and new methods could be estimated and incorporated into estimation of trends.

BROADER INTEGRATION

Although our focus in this bulletin has been on broad-scale viability indicators, we would like to touch on how the core components of the monitoring framework described can support a much broader integration for different types of biological assessment of steelhead and Rainbow Trout.

POPULATION RESTORATION

According to CDFW (2018), restoration monitoring comprises one or more of three activities: implementation monitoring, effectiveness monitoring, and validation monitoring (Table 19). These concepts can be broadened beyond restoration to management actions more generally, asking if a given management action was implemented as intended; how effectively it produced the desired habitat conditions or watershed processes; and whether the biota responded as originally intended.

Table 19. Components of restoration monitoring (Duffy 2006).

Component	Purpose
Implementation Monitoring	Documents the fulfillment of restoration contracts or compliance with regulations or laws
Effectiveness Monitoring	Estimates trends in resource condition following a management action, usually associated with physical or chemical processes and habitat conditions
Validation Monitoring	Estimates the response of biota to restoration actions, ideally establishing cause-and-effect relationships between restoration actions and biota

Effectiveness and validation monitoring can be readily integrated with the reach-sampling framework outlined in this report, by following these steps:

1. Select response variables (habitat metrics, fish life-stages, etc.) appropriate to the goals of the restoration project or management action.
2. Select a sampling frame, either long reaches or short reaches, depending on the appropriate field method for the response variables.
3. Define targets of estimation, typically a restoration or management area, and a broader control area that is comparable ecologically but will not be affected.
4. Randomly sample reaches from each of the targets and collect data on response variables before and after the restoration.
5. Effect size can be estimated using the asymmetrical Before-After/Control Impact statistical framework (Underwood 1994).

In general, the statistical power of such efforts to detect response is increased by sampling multiple restoration sites (Baldigo and Warren 2008) and choosing an appropriate scheme for revisiting each sample reach (Underwood and Chapman 2003). A useful modification is to define the control area as a "reverse control," a relatively pristine habitat supporting good fish numbers, which differs from the restoration site initially, but toward which it is predicted to converge after the action.

The benefits of this integration are that, if carefully done, the effect size can be validly estimated and the results validly generalized to the entire control area. Thus, a key step is the selection of the control area, including the judgment of other stream reaches that are comparable ecologically. Typically such judgments would be based on broad-scale physical and biotic controls (Melles et al. 2012), such as climate, process-based channel types (Montgomery and Buffington 1997) or presence of species that act as ecological engineers (Rosell et al. 2005, Francis et al. 2009). Another key step is the selection of response variables. Ideally the validation response is measured via an unbiased estimate of a biologically meaningful quantity—typically abundance, survival, or meaningful traits (e.g., size distribution, genetic composition) of some life stage of interest. If so, then the response can be validly incorporated into data-driven life-cycle models that link management activities to overall population and ESU response.

One of the most attractive response variables for freshwater management actions is smolt production. Smolt production is meaningful in the sense that it summarizes the signal of how freshwater management actions influence steelhead production, whereas if adult steelhead production is used as an indicator, the signal can be obscured by the "noise" of variable ocean survival. Thus smolt production is a useful summary indicator of the effects of freshwater management. Smolt production, especially if implemented by tagging juveniles and monitoring outmigration, can be used to compare production between management and control areas using the Before-After/Control-Impact framework. Smolt production can be estimated at time scales useful for management actions, typically giving a response with 1-2 year time lag rather than 3-5 year lag to see a response in adult numbers. Finally, statistical power of tagging studies is typically greater for smolt production estimates than for adult abundance estimates because much larger numbers of tagged fish will be detected going out than coming in from the ocean several years later.

More generally, validation comparisons can be made of abundance or survival at any specific life stage by deploying appropriate field methods within the reach-sampling framework given above. Some useful methodologies to develop include validation designs for assessing upstream or downstream migration success, redd survival via revisits of redds, and evaluation of survival and growth in lagoons using mark-recapture approaches or scale analysis.

FISHERIES MANAGEMENT

Although both steelhead Distinct Population Segments in the southern monitoring area are currently on the Federal Endangered Species List, the ultimate goal of the listings and recovery plans is to restore them to viability so that they can be delisted. Although such delisting is likely still some ways in the future, if indeed it can ever be justified, we note that the viability-monitoring scheme outlined here can also be used to manage the recreational fishery both before and after delisting. The key goals of such a scheme would be to (1) continue demonstrating that a DPS is neither threatened nor endangered, and (2) provide data to support management of the recreational steelhead fisheries. Two of the five factors described in the Endangered Species Act for a listing determination (ESA; 16 U.S.C. §1533 (a)(1)) are "overutilization for commercial, recreational, scientific, or educational purposes," and "inadequacy of existing regulatory mechanisms," and the transition to a fisheries-monitoring scheme would directly address both these factors.

A fisheries-monitoring scheme would likely realign adult abundance monitoring, to emphasize fish populations as a valuable resource as well as a vulnerable species. Perhaps the rotation scheme for counting stations could be broadened to include fishable populations, or all populations for broad-scale monitoring. On the Rainbow Trout side, particular populations beyond the backbone could receive increased sampling to inform fisheries management decisions.

With respect to abundance monitoring, our tighter integration with recovery plans and our focus on backbone populations departs from the original goal of Adams et al. (2011) to provide data for broad-scale trend analysis. Trends within individual backbone populations are not necessarily the same as broader trends across the entire ESU. However, the low-flow surveys do provide trends at the level of biogeographic areas for two key indicators: Rainbow Trout abundance and the genetic Omy5 indicator of anadromy. If the plan is implemented as recommended in this bulletin, these can be aggregated for trend analysis at the ESU scale or for the entire southern monitoring area.

In addition, individual backbone populations can produce estimates for these two indicators, as well as for anadromous abundance and anadromous fraction. For these populations, the recommended monitoring scheme will support estimation of the means and covariances for the three-way relationship between Rainbow Trout abundance, steelhead abundance, and genetic haplotype proportions. Moreover, these estimates will be replicated across multiple backbone populations, allowing for a second-level estimate of interpopulation variation in this relationship. From such data and analysis, it should be feasible to develop a model-based approach for estimating trends in steelhead abundance at the broad scale originally envisioned by Adams et al. (2011).

INTEGRATED LIFE-CYCLE MODELS

The data collected at life-cycle monitoring stations imply a simple population model that separates the steelhead life-cycle into a freshwater and marine phase, with explicit estimation of abundance at each phase transition (smolt production, adult returns). The ratios of the abundances are then conceptualized as demographic rates during the phases, namely marine survival and, on the freshwater side, the product of spawner survival, spawner fecundity, and the resulting progeny's freshwater survival and smolting rate. This implicit life-cycle model can be formalized as an explicit quantitative model using the techniques of integrated population models (Schaub and Abadi 2011). Integrated population models provide a statistically and conceptually valid way to integrate (combine) different datasets to analyze responses of populations and groups of populations to external drivers. When carefully constructed and supported by adequate datasets, such models can:

- Compare effects of various past management actions on different life stages,
- Add up the cumulative effects of various past management actions on different life stages,
- Characterize "effect" in terms of meaningful indicators such as adult abundance or anadromous fraction,
- Ask "what if" questions about future management scenarios or environmental scenarios,
- Accurately and flexibly characterize the uncertainty embodied by such analyses, especially if developed using the Bayesian statistical framework,
- Be restrained and data-driven or, if useful, incorporate additional complexity such as growth patterns or age structure via additional model assumptions.

Typically, an integrated population model is composed of a quantitative population model linked to one or more observation models for particular life stages. The population model has estimated parameters describing abundance, survival, reproduction, and transitions of different life stages, whereas each observation model defines a statistical model that links abundance of a specific life-stage to a specific type of observation made in the real world. In such models, the focus is on statistically valid inference about population growth and vital rates, such as stage-specific survival or maturation rates. This approach can provide deep insight into population dynamics, as it allows for the systematic assessment of links between limiting factors, demography, and population growth (Schaub and Abadi 2011).

The flexibility of the CMP sampling framework and the flexibility of integrated population models are complementary. The flexibility of the sampling framework is that additional life-stages and targets of estimation can be developed as needed for specific management questions, as described in the previous section. The flexibility of integrated population models is two-fold: first, these new data can be incorporated into existing quantitative models by defining additional observation models and parsing out "lumped" life-stages; and second, inference can be extended laterally to other populations by incorporating multilevel statistical models (Gelman and Hill 2007). Thus, an integrated population model can naturally grow and provide more insight as the supporting datasets grow and provide more information.

In a sense, some of these techniques for integrated population modeling are already embedded in the core monitoring strategy. For example, in our Design 2, redd abundance is estimated in all backbone populations but adult abundance and smolt production are only monitored at life-cycle monitoring stations, similar to the northern monitoring area. Clarity about the interpretation of redd abundances is thus obtained by estimating additional key life stages at a subset of populations (LCMs) and extending inference laterally to other populations.

The conceptual and analytical tools for creating integrated population models already exist (Kery and Schaub 2012). Implementation of CMP monitoring would provide the necessary datasets for the basic approach. Additional data collected as needed for specific questions would support assessments to inform specific management decisions. All the pieces would be in place for a quantitative learning institution supporting long-term recovery and management of the steelhead fisheries resource in the southern area.

PRODUCTIVITY AND ADAPTIVE MANAGEMENT

Productivity research is basic research focused on identifying the ecological factors that limit population productivity and estimating their effect size. Adams et al. (2011) envisioned LCMs as the sites for most productivity research, in part because the monitoring data collected at LCMs provides a useful context for interpreting the results of productivity research (see previous section). But the flexible framework for validation monitoring and life-cycle modeling described above means that productivity research can be pursued more widely, including in intensively managed watersheds. When it is used to inform management decisions, it is commonly framed as a form of adaptive management. Adaptive management comprises actions that are implemented as experiments with testable outcomes, leading to a structured process of learning by doing (Walters 1997). We regard productivity research and adaptive management as similar activities that in practice often grade into each other, and recommend a flexible approach to implementing productivity studies in a variety of populations to address specific management challenges as needed.

Monitoring the status and trends of VSP parameters, combined with targeted productivity research and adaptive management, provide a powerful framework for restoring habitat and recovering steelhead in the southern area because they comprise a learning institution. Bernhardt et al. (2005) warned that river restoration practitioners would not be able to learn what practices are effective and thus advance restoration science, without suitable learning systems. A learning institution in this sense is a set of social conventions that provide an informational feedback loop, like the data-collection/ peer review/ publication cycle pursued by most scientists, or indeed any system of trial and error that tracks successes and failures and learns from them.

To implement these concepts, Palmer et al. (2005) proposed a general, open-ended standard for judging the ecological success of river restoration efforts. They distinguished ecological success as distinct from other types of success such as cost-effectiveness or social acceptability, noting that ecological restoration or rehabilitation is the explicit goal of many projects and that progress in the practice of river rehabilitation is hampered by a lack of agreed criteria for judging ecological success. They proposed five general criteria for ecological success:

- A guiding image exists: A dynamic ecological endpoint is identified a priori and used to guide the restoration. Here the overall guiding image is the set of VSP criteria for steelhead, reviewed every 5 years as part of NMFS status review updates.
- The ecosystem is demonstrably improved: Ecologically successful restoration will induce measurable changes in the ecosystem that move it towards the agreed upon guiding image. The updated monitoring scheme provides the basis for demonstrating this movement by monitoring the set of VSP indicators.
- Resilience is increased: Steelhead populations are more self-sustaining than prior to the restoration. Though outside the scope of this document, resilience is demonstrated when VSP indicators display good status and trends without the need for intrusive management activities such as hatcheries or rearing facilities.
- No lasting harm is done: Implementing the recovery effort does not inflict irreparable harm to the resource. The framework for productivity research and adaptive management provides a basis for evaluating harm as well as benefit.
- Ecological assessment is completed: Some level of both pre- and post-project assessment is conducted and the information made available. The framework for adaptive management experiments and validation monitoring, integrated with viability indicators via life-cycle models, provides a state-of-the-art basis for such assessment. In addition, the generation of CMP data by CDFW and partners, combined with status review updates conducted every 5 years by NMFS under the

Endangered Species Act, comprise a regular cycle of legally mandated ecological assessment of each DPS as a whole.

With the rapidly changing climate we are now experiencing, we believe the need for such a learning institution will be ongoing if we hope to prevent steelhead from going extinct, restore them to viability, and reestablish them as a fisheries resource. The most certain thing is that the future will be different from the past, and adaptability will be an ongoing process in managing salmonid populations in California. Thus, although this bulletin is an update of the original formulation laid out by Adams et al. (2011), we neither hope nor expect that it will be the last.

ACKNOWLEDGEMENTS

This update was originally based on a series of discussions with Dana McCanne (CDFW), Ann-Marie Osterback (NMFS, UC Santa Cruz), Joel Casagrande (NMFS), and Rick Bush (NMFS). It incorporated numerous ideas and comments from K. McLaughlin (CDFW), M. Michie (CDFW), D. Baldwin (CDFW), D. Michniuk (CDFW), M. Paul (CDFW) and H. Ohms (UCSC). The update was greatly helped by additional review and comment from K. Atkinson (CDFW), J. Ferreira (CDFW), S. Ricker (CDFW), T. MacDonald (West Inc.), P. Adams (NMFS retired), M. Hereford (ODFW), S. Jacobson (California Trout), E. Bell (Stillwater Sciences), T. Nickelson, an anonymous reviewer known as "kisutch," and discussions on topics too numerous to mention with M. Capelli (NMFS), K. Urquhart (MPWMD), and C. Hamilton (MPWMD). M. Mohr (NMFS) provided helpful pen-and-ink renditions of statistical matters. B. Spence (NMFS) gave the final draft a close reading and provided feedback that greatly aided its clarity.

LITERATURE CITED

Abadia-Cardoso, A., D. E. Pearse, S. Jacobson, J. Marshall, D. Dalrymple, F. Kawasaki, G. Ruiz-Campos, and J. C. Garza. 2016. Population genetic structure and ancestry of steelhead/Rainbow Trout (*Oncorhynchus mykiss*) at the extreme southern edge of their range in North America. Conservation Genetics **17**:675-689.

Adams, P. B., L. B. Boydstun, S. P. Gallagher, M. K. Lacy, T. McDonald, and K. E. Shaffer. 2011. California coastal salmonid population monitoring: strategy, design, and methods. Fish Bulletin. State of California, Department of Fish and Game **180**:4-82.

Ainslie, B. J., J. R. Post, and A. J. Paul. 1998. Effects of pulsed and continuous DC electrofishing on juvenile Rainbow Trout. North American Journal of Fisheries Management **18**:905-918.

Allen, M. 2014. Steelhead population and habitat assessment in the Ventura River / Matilija Creek Basin 2006–2012. Normandeau Associates, Inc., Arcata, California.

Alley, D. W. 2017. Fishery and water quality monitoring of Pajaro River Lagoon in 2017. Santa Cruz County Environmental Health Department, Santa Cruz, CA, USA.

Apgar, T. M., D. E. Pearse, and E. P. Palkovacs. 2017. Evolutionary restoration potential evaluated through the use of a trait-linked genetic marker. Evolutionary Applications **10**:485-497.

Arriaza, J. L., D. A. Boughton, K. Urquhart, and M. Mangel. 2017. Size-conditional smolting and the response of Carmel River steelhead to two decades of conservation efforts. Plos One **12**.

Ashbrook, C. E., K. W. Hassel, J. F. Dixon, and A. Hoffmann. 2007. Tangle nets. Pages 341-361 *in* D. H. Johnson, B. M. Shrier, J. S. O'Neil, J. A. Knutzen, X. Augerot, T. A. O'Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Socieity, Bethesda, Maryland.

Atkinson, K. A., M. K. Lacy, and R. J. Bellmer. 2016. Dual frequency identification sonar (DIDSON) deployment and preliminary performance as part of the California coastal salmonid monitoring plan. Administrative Report 2016-01. California Department of Fish and Wildlife, Sacramento, CA.

Baker, C. F., K. Reeve, D. Baars, D. Jellyman, and P. Franklin. 2017. Efficacy of 12-mm half-duplex passive integrated transponder tags in monitoring fish movements through stationary antenna systems. North American Journal of Fisheries Management **37**:1289-1298.

Baldigo, B. P., and D. R. Warren. 2008. Detecting the response of fish assemblages to stream restoration: effects of different sampling designs. North American Journal of Fisheries Management **28**:919-934.

Bangs, P. D., and J. J. Nagler. 2014. A comparison of nonlethal methods for evaluating the reproductive status of female coastal cutthroat trout. Journal of Fish and Wildlife Management **5**:183-190.

Barlaup, B. T., H. Rund, E. S. Normann, S. Stranzl, S. Mahlum, and K. W. Vollset. 2018. Out of sync: monitoring the time of sea entry of wild and hatchery salmon *Salmo salar* smolt using floating passive-integrated transponder antennae. Journal of Fish Biology **93**:455-464.

Berejikian, B. A., J. T. Gable, and D. T. Vidergar. 2011. Effectiveness and trade-offs associated with hydraulic egg collections from natural salmon and steelhead redds for conservation hatchery programs. Transactions of the American Fisheries Society **140**:549-556.

Bernhardt, E. S., M. A. Palmer, J. D. Allan, G. Alexander, K. Barnas, S. Brooks, J. Carr, S. Clayton, C. Dahm, J. Follstad-Shah, D. Galat, S. Gloss, P. Goodwin, D. Hart, B. Hassett, R. Jenkinson, S. Katz, G. M. Kondolf, P. S. Lake, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano, B. Powell, and E. Sudduth. 2005. Ecology—Synthesizing US river restoration efforts. Science **308**:636-637.

Bertrand, K. N., K. Gido, and C. S. Guy. 2006. An evaluation of single-pass versus multiple-pass backpack electrofishing to estimate trends in species abundance and richness in prairie streams. Transactions of the Kansas Academy of Science **109**:131-138.

Bjorkstedt, E. P. 2005. DARR 2.0: Updated software for estimating abundance from stratified mark-recapture data. NOAA Technical Memorandum NMFS-SWFSC **368**:1-13.

Bond, M. H. 2006. Importance of estuarine rearing to central California steelhead (*Oncorhynchus mykiss*) growth and marine survival. Master's thesis. University of California Santa Cruz, Santa Cruz, California.

Bond, M. H., S. A. Hayes, C. V. Hanson, and R. B. MacFarlane. 2008. Marine survival of steelhead (*Oncorhynchus mykiss*) enhanced by a seasonally closed estuary. Canadian Journal of Fisheries and Aquatic Sciences **65**:2242-2252.

Booth, D. B., Y. T. Cui, Z. E. Diggory, D. Pedersen, J. Kear, and M. Bowen. 2013. Determining appropriate instream flows for anadromous fish passage on an intermittent mainstem river, coastal southern California, USA. Ecohydrology **7**:745-759.

Boughton, D. A. 2010. Estimating the size of steelhead runs by tagging juveniles and monitoring migrants. North American Journal of Fisheries Management **30**:89-101.

Boughton, D. A., P. Adams, E. Anderson, C. Fusaro, E. A. Keller, E. Kelley, L. Lentsch, J. Nielsen, K. Perry, H. M. Regan, J. J. Smith, C. Swift, L. Thompson, and F. Watson. 2006. Steelhead of the south-central/southern California coast: population characterization for recovery planning. NOAA Technical Memorandum NMFS-SWFSC **394**.

Boughton, D. A., P. Adams, E. Anderson, C. Fusaro, E. A. Keller, E. Kelley, L. Lentsch, J. Nielsen, K. Perry, H. M. Regan, J. J. Smith, C. Swift, L. Thompson, and F. Watson. 2007. Viability criteria for steelhead of the south-central and southern California coast. NOAA Technical Memorandum NMFS-SWFSC **407**.

Boughton, D. A., H. Fish, K. Pipal, J. Goin, F. Watson, J. Casagrande, J. Casagrande, and M. Stoecker. 2005. Contraction of the southern range limit for anadromous *Oncorhynchus mykiss*. NOAA Technical Memorandum NMFS-SWFSC **380**.

Boughton, D. A., H. Fish, J. Pope, and G. Holt. 2009. Spatial patterning of habitat for *Oncorhynchus mykiss* in a system of intermittent and perennial streams. Ecology of Freshwater Fish **18**:92-105.

Boughton, D., J. Fuller, G. Horton, E. Larson, W. Matsubu, and C. Simestad. 2017. Spatial structure of water-quality impacts and foraging opportunities for Steelhead in the Russian River Estuary: An energetics perspective. NOAA Technical Memorandum SWFSC-**569**.

Brown, R. J., and K. P. Severin. 2009. Otolith chemistry analyses indicate that water Sr:Ca is the primary factor influencing otolith Sr:Ca for freshwater and diadromous fish but not for marine fish. Canadian Journal of Fisheries and Aquatic Sciences **66**:1790-1808.

Buffington, J. M., D. R. Montgomery, and H. M. Greenberg. 2004. Basin-scale availability of salmonid spawning gravel as influenced by channel type and hydraulic roughness in mountain catchments. Canadian Journal of Fisheries and Aquatic Sciences **61**:2085-2096.

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

CDFW. 2018. Fisheries Habitat Restoration 2018 Proposal Solicitation Notice. Pages 1-44. California Department of Fish and Wildlife, Wildlife and Fisheries Division, Fisheries Branch and Watershed Restoration Grants Branch, https://www.wildlife.ca.gov/Grants/FRGP/Solicitation, Sacramento, California.

Clemento, A. J., E. C. Anderson, D. Boughton, D. Girman, and J. C. Garza. 2009. Population genetic structure and ancestry of *Oncorhynchus mykiss* populations above and below dams in south-central California. Conservation Genetics **10**:1321-1336.

Collins, K. M., E. L. Brannon, L. L. Moulton, M. A. Cronin, and K. R. Parker. 2000. Hydraulic sampling protocol to estimate natural embryo mortality of pink salmon. Transactions of the American Fisheries Society **129**:827-834.

Constable, R. J., jr., and E. Suring. 2015. Juvenile salmonid monitoring in coastal Oregon and lower Columbia streams, 2014. Monitoring program report number OPSW-ODFW-2015-1. Oregon Department of Fish and Wildlife, Salem, Oregon.

Crisp, D. T., and P. A. Carling. 1989. Observations on siting, dimensions and structure of salmonid redds. Journal of Fish Biology **34**:119-134.

Cuthbert, R., P. Cuthbert, A. Fuller, and M. Hellmair. 2014. Salinas River basin adult steelhead escapement monitoring, 2014 Annual Report. FishBio, Oakdale, CA.

Dagit, R., K. Adamek, D. Alvarez, S. Contreras, R. Dauksis, B. Demercie, D. Hofflander, J. Mongolot, and E. Montgomery. 2018. Santa Monica Bay anadromous adult and juvenile steelhead monitoring 2013-2018. Prepared for CDFW contracts P1250013 and P1450013 by the Resource Conservation District fo the Santa Monica Mountains, Topanga, CA.

Dagit, R., M. T. Booth, M. Gomez, T. Hovey, S. Howard, S. D. Lewis, S. Jacobson, M. Larson, D. McCanne, and T. H. Robinson. 2020. Occurrences of steelhead trout (*Oncorhynchus mykiss*) in southern California, 1994-2018. California Fish and Wildlife **106**:39-58.

Dauphin, G. J. R., G. Chaput, C. Breau, and R. A. Cunjak. 2019. Hierarchical model detects decadal changes in calibration relationships of single-pass electrofishing indices of abundance of Atlantic salmon in two large Canadian catchments. Canadian Journal of Fisheries and Aquatic Sciences **76**:523-542.

Dennis, E. B., B. J. T. Morgan, and M. S. Ridout. 2015. Computational aspects of N-mixture models. Biometrics **71**:237-246.

Denton, K. P., M. McHenry, R. Moses, E. Ward, M. Liermann, O. Stefankiv, W. Wells, and G. Pess. 2014. 2014 Elwha River steelhead escapement estimate based on DIDSON/ARIS multi-beam SONAR data. Prepared for the Lower Elwha Tribe. Provided to Olympic National Park by the Lower Elwha Klallam Tribe in partial fulfillment of task number 6-1 of contract number P13PC00296.

Denton, K. P., M. McHenry, R. Moses, E. Ward, O. Stefankiv, W. Wells, and G. Pess. 2015. 2015 Elwha River steelhead escapement estimate based on DIDSON/ARIS multi-beam SONAR data. Provided to Olympic National Park by the Lower Elwha Klallam Tribe in partial fulfillment of contract number P16PC00005.

Donohoe, C. J., P. B. Adams, and C. F. Royer. 2008. Influence of water chemistry and migratory distance on ability to distinguish progeny of sympatric resident and anadromous Rainbow Trout (*Oncorhynchus mykiss*). Canadian Journal of Fisheries and Aquatic Sciences **65**:1060-1075.

Duffy, W. G. 2006. Protocols for monitoring the response of anadromous salmon and steelhead to watershed restoration in California. Pages 1-85. California Department of Fish and Game, Salmon and Steelhead Trout Restoration Account Agreement No. P0210565, https://www.wildlife.ca.gov/Grants/FRGP/Guidance, Arcata, California.

Elsdon, T. S., and B. M. Gillanders. 2004. Fish otolith chemistry influenced by exposure to multiple environmental variables. Journal of Experimental Marine Biology and Ecology **313**:269-284.

Elsdon, T. S., and B. M. Gillanders. 2005. Alternative life-history patterns of estuarine fish: barium in otoliths elucidates freshwater residency. Canadian Journal of Fisheries and Aquatic Sciences **62**:1143-1152.

Emmett, R., R. Llanso, J. Newton, R. Thom, M. Hornberger, C. Morgan, C. Levings, A. Copping, and P. Fishman. 2000. Geographic signatures of North American West Coast estuaries. Estuaries **23**:765-792.

Fierstine, H. L., K. F. Kline, and G. R. Garman. 1973. Fishes collected in Morro Bay, California between January, 1968 and December, 1970. California Fish and Game **59**:73-88.

Foley, K., A. Rosenberger, and F. Mueter. 2015. Effectiveness of single-pass backpack electrofishing to estimate juvenile coho salmon abundance in Alaskan headwater streams. Fisheries Science **81**:601-610.

Francis, R. A., D. Corenblit, and P. J. Edwards. 2009. Perspectives on biogeomorphology, ecosystem engineering and self-organisation in island-braided fluvial ecosystems. Aquatic Sciences **71**:290-304.

Frechette, D. M., W. H. Satterthwaite, A. M. K. Osterback, and S. A. Hayes. 2016. Steelhead abundance in seasonally closed estuaries estimated using mark recapture methods. NOAA Techical Memorandum NMFS SWFSC **555**.

Funk, W. C., J. K. McKay, P. A. Hohenlohe, and F. W. Allendorf. 2012. Harnessing genomics for delineating conservation units. Trends in Ecology & Evolution **27**:489-496.

Gabrielsson, R. M., J. Kim, M. R. Reid, C. H. Stirling, M. Numata, and G. P. Closs. 2012. Does the trace element composition of brown trout *Salmo trutta* eggs remain unchanged in spawning redds? Journal of Fish Biology **81**:1871-1879.

Gallagher, S. P., and C. M. Gallagher. 2005. Discrimination of chinook salmon, coho salmon, and steelhead redds and evaluation of the use of redd data for estimating escapement in several unregulated streams in northern California. North American Journal of Fisheries Management **25**:284-300.

Gallagher, S. P., P. K. Hahn, and D. H. Johnson. 2007. Redd counts. Pages 197-234 *in* D. H. Johnson, B. M. Shrier, J. S. O'Neil, J. A. Knutzen, X. Augerot, T. A. O'Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Socieity, Bethesda, MD, USA.

Gallagher, S. P., D. W. Wright, B. W. Collins, and P. B. Adams. 2010. A regional approach for monitoring salmonid status and trends: results from a pilot study in coastal Mendocino County, California. North American Journal of Fisheries Management **30**:1075-1085.

Garwood, J., and S. Ricker. 2011. Spawner survey sample frame development for monitoring adult salmonid populatons in California. California Department of Fish and Game, Anadromous Fisheries Resource and Monitoring Program, Arcata, CA, USA.

Gelman, A., and J. Hill. 2007. Data analysis using regression and multilevel/hierarchical models. Cambridge University Press, Cambridge.

Gibson-Reinemer, D. K., B. M. Johnson, P. J. Martinez, D. L. Winkelman, A. E. Koenig, and J. D. Woodhead. 2009. Elemental signatures in otoliths of hatchery Rainbow Trout (*Oncorhynchus mykiss*): distinctiveness and utility for detecting origins and movement. Canadian Journal of Fisheries and Aquatic Sciences **66**:513-524.

Gilbert, E. I., J. C. Dean, and M. R. Maglothin. 2017. Responses of American Bullfrog, *Lithobates catesbeianus*, and Southern Leopard Frog, *Lithobates sphenocephalus*, to low voltages in uniform aquatic electrical fields. Southwestern Naturalist **62**:148-154.

Gilchrist, J., D. Suddjian, B. Mori, K. Reeves, K. Lyons, J. Smith, L. Herter, S. Nielson, J. Haltiner, L. Sklar, J. Buchholz, and D. Spicher. 1992. Salinas River lagoon management and enhancement plan, volume 2: technical appendices., Monterey County Water Resources Agency, Salinas, CA, USA.

Hagar Environmental Science (HES). 2017. Technical memorandum, City of Santa Cruz habitat conservation plan, lagoon fish population sampling 2016. City of Santa Cruz, Santa Cruz, CA, USA.

Hankin, D. G., and G. H. Reeves. 1988. Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences **45**:834-844.

Harrison, L. R., A. E. East, D. P. Smith, J. B. Logan, R. M. Bond, C. L. Nicol, T. H. Williams, D. A. Boughton, K. Chow, and L. Luna. 2018. River response to large-dam removal in a mediterranean hydroclimatic setting: Carmel River, California, USA. Earth Surface Processes and Landforms **DOI: 10.1002/esp.4464**.

Harrison, L. R., A. Pike, and D. A. Boughton. 2017. Coupled geomorphic and habitat response to a flood pulse revealed by remote sensing. Ecohydrology **DOI: 10.1002/eco1845**.

Hayes, S. A., M. H. Bond, C. V. Hanson, E. V. Freund, J. J. Smith, E. C. Anderson, A. J. Ammann, and R. B. Macfarlane. 2008. Steelhead growth in a small central California watershed: upstream and estuarine rearing patterns. Transactions of the American Fisheries Society **137**:114-128.

Hayes, S. A., M. H. Bond, C. V. Hanson, A. W. Jones, A. J. Ammann, J. A. Harding, A. L. Collins, J. Perez, and R. B. MacFarlane. 2011. Down, up, down and "smolting" twice? Seasonal movement patterns by juvenile steelhead (*Oncorhynchus mykiss*) in a coastal watershed with a bar closing estuary. Canadian Journal of Fisheries and Aquatic Sciences **68**:1341-1350.

Horn, M. H. 1980. Diel and seasonal variation in abundance and diversity of shallow-water fish populations in Morro Bay, California. Fishery Bulletin **78**:759-770.

Huso, M. M. P., D. Dalthorp, D. Dail, and L. Madsen. 2015. Estimating wind-turbine-caused bird and bat fatality when zero carcasses are observed. Ecological Applications **25**:1213-1225.

Hwan, J. L., and S. M. Carlson. 2016. Fragmentation of an intermittent stream during seasonal drought: intra-annual and interannual patterns and biological consequences. River Research and Applications **32**:856-870.

Ibbotson, A. T., W. R. C. Beaumont, D. Collinson, A. Wilkinson, and A. C. Pinder. 2004. A cross-river antenna array for the detection of miniature passive integrated transponder tags in deep, fast flowing rivers. Journal of Fish Biology **65**:1441-1443.

IEP. 1998. Steelhead life-stage assessment protocol. Interagency Ecological Program Steelhead Project Work Team. December, 1998.

Jankovitz, J., and R. Diller. 2019. Summary of annual fish sampling, water quality monitoring, and active management. Pescadero Creek lagoon complex, 2018., California Department of Fish and Wildlife, Sacramento, California.

Jones, M. L., and J. D. Stockwell. 1995. A rapid assessment procedure for the enumeration of salmonine populations in streams. North American Journal of Fisheries Management **15**:551-562.

Kelson, S. J., M. R. Miller, T. Q. Thompson, S. M. O'Rourke, and S. M. Carlson. 2019. Do genomics and sex predict migration in a partially migratory salmonid fish, *Oncorhynchus mykiss*? Canadian Journal of Fisheries and Aquatic Sciences **76**:2080-2088.

Kendall, N. W., J. R. McMillan, M. R. Sloat, T. W. Buehren, T. P. Quinn, G. R. Pess, K. V. Kuzishchin, M. M. McClure, and R. W. Zabel. 2015. Anadromy and residency in steelhead and Rainbow Trout (*Oncorhynchus mykiss*): a review of the processes and patterns. Canadian Journal of Fisheries and Aquatic Sciences **72**:319-342.

Kery, M., and M. Schaub. 2012. Bayesian population analysis using WinBUGS: A hierarchical perspective. Academic Press, Waltham, MA.

Killam, D. 2012. Chinook salmon populations for the upper Sacramento River Basin in 2011. RBFO Technical Report No. 03-2012. California Department of Fish and Game - Northern Region, Red Bluff, CA.

Kincaid, T. M., A. R. Olsen, and M. H. Weber. 2019. spsurvey: Spatial survey design and analysis. R package version 4.1.0. R Foundation, https://CRAN.R-project.org/package=spsurvey.

Kondolf, G. M., K. Podolak, and T. E. Grantham. 2013. Restoring mediterranean-climate rivers. Hydrobiologia **719**:527-545.

Korman, J., J. Schick, and B. Mossop. 2016. Estimating riverwide abundance of juvenile fish populations: how much sampling is enough? North American Journal of Fisheries Management **36**:213-229.

Kristensen, E. A., G. P. Closs, R. Olley, J. Kim, M. Reid, and C. Stirling. 2011. Determining the spatial distribution of spawning by anadromous and resident brown trout *Salmo trutta* L using strontium content of eggs collected from redds. Ecology of Freshwater Fish **20**:377-383.

Leitwein, M., J. C. Garza, and D. E. Pearse. 2017. Ancestry and adaptive evolution of anadromous, resident, and adfluvial Rainbow Trout (*Oncorhynchus mykiss*) in the San Francisco bay area: application of adaptive genomic variation to conservation in a highly impacted landscape. Evolutionary Applications **10**:56-67.

Magie, R. J., M. S. Morris, J. P. Bender, B. F. Jonasson, B. P. Sandford, and R. D. Ledgerwood. 2015. Development of a passive integrated transponder (PIT) detection system for adult salmonids in the lower Columbia River, 2013. Report of research by Fish Ecology Division, Northwest Fisheries Science Center, National Marine Fisheries Service. National Oceanographic and Atmospheric Administration, Seattle, WA.

May, C. L., and D. C. Lee. 2004. The relationships among in-channel sediment storage, pool depth, and summer survival of juvenile salmonids in Oregon Coast Range streams. North American Journal of Fisheries Management **24**:761-774.

McCormick, S. D. 1993. Methods for nonlethal gill biopsy and measurement of Na+, K+-ATPase activity. Canadian Journal of Fisheries and Aquatic Sciences **50**(3):656-658.

McElhany, P., M. H. Ruckelshous, M. J. Ford, T. C. Wainwright, and E. P. Bjorkstedt. 2000. Viable salmonid populations and the recovery of evolutionarily significant units. NOAA Techical Memorandum NMFS SWFSC **42**.

McMillan, J. R., G. R. Pess, M. Liermann, S. A. Morley, M. L. McHenry, L. A. Campbell, and T. P. Quinn. 2015. Using redd attributes, fry density, and otolith microchemistry to distinguish the presence of steelhead and Rainbow Trout in the Elwha River dam removal project. North American Journal of Fisheries Management **35**:1019-1033.

Melles, S. J., N. E. Jones, and B. Schmidt. 2012. Review of theoretical developments in stream ecology and their influence on stream classification and conservation planning. Freshwater Biology **57**:415-434.

Mills, J. S., J. B. Dunham, G. H. Reeves, J. R. McMillan, C. E. Zimmerman, and C. E. Jordan. 2012. Variability in expression of anadromy by female *Oncorhynchus mykiss* within a river network. Environmental Biology of Fishes **93**:505-517.

Mohammadi, M. 2018. A variance bound for unbiased estimation in inverse sampling without replacement. Statistical Papers **59**:1649-1655.

Montgomery, D. R., and J. M. Buffington. 1997. Channel-reach morphology in mountain drainage basins. Geological Society of America Bulletin **109**:596-611.

MPWMD. 2015. 2013-2014 Annual Report for the MPWMD Mitigation Program. Monterey Peninsula Water Management District. 178 pp. Available at https://www.mpwmd.net/environmental-stewardship/mitigation-program/.

Nadeau, T. L., and M. C. Rains. 2007. Hydrological connectivity between headwater streams and downstream waters: How science can inform policy. Journal of the American Water Resources Association **43**:118-133.

Nagahama, Y. 1983. The functional-morphology of teleost gonads. Fish Physiology **9**:223-275.

Nichols, K. M., A. F. Edo, P. A. Wheeler, and G. H. Thorgaard. 2008. The genetic basis of smoltification-related traits in *Oncorhynchus mykiss*. Genetics **179**:1559-1575.

NMFS. 2012. Southern California Steelhead Recovery Plan. National Marine Fisheries Service, Southwest Region, Protected Resources Division, Long Beach, California.

NMFS. 2013. South-Central California Coast Steelhead Recovery Plan. National Marine Fisheries Service, West Coast Region, California Coastal Office, Long Beach, California.

O'Malley, K. G., T. Sakamoto, R. G. Danzmann, and M. M. Ferguson. 2003. Quantitative trait loci for spawning date and body weight in Rainbow Trout: Testing for conserved effects across ancestrally duplicated chromosomes. Journal of Heredity **94**:273-284.

O' Neal, J. S. 2007. Snorkel Surveys. Pages 325-340 *in* D. H. Johnson, B. M. Shrier, J. S. O'Neil, J. A. Knutzen, X. Augerot, T. A. O'Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Socieity, Bethesda, MD, USA.

Ohms, H. A., A. I. Gitelman, C. E. Jordan, and D. A. Lytle. 2019. Quantifying partial migration with sex-ratio balancing. Canadian Journal of Zoology **97**:352-361.

Ohms, H. A., M. R. Sloat, G. H. Reeves, C. E. Jordan, and J. B. Dunham. 2014. Influence of sex, migration distance, and latitude on life history expression in steelhead and Rainbow Trout (*Oncorhynchus mykiss*). Canadian Journal of Fisheries and Aquatic Sciences **71**:70-80.

Page, L. M., H. Espinosa-Perez, L. T. Findley, C. R. Gilbert, R. N. Lea, N. E. Mandrak, R. L. Mayden, and J. S. Nelson. 2013. Common and scientific names of fishes from the United States, Canada, and Mexico, 7th Edition. American Fisheries Society, Bethesda, MD, USA.

Palmer, M. A., E. S. Bernhardt, J. D. Allan, P. S. Lake, G. Alexander, S. Brooks, J. Carr, S. Clayton, C. N. Dahm, J. F. Shah, D. L. Galat, S. G. Loss, P. Goodwin, D. D. Hart, B. Hassett, R. Jenkinson, G. M. Kondolf, R. Lave, J. L. Meyer, T. K. O'Donnell, L. Pagano, and E. Sudduth. 2005. Standards for ecologically successful river restoration. Journal of Applied Ecology **42**:208-217.

Pearse, D. E. 2016. Saving the spandrels? Adaptive genomic variation in conservation and fisheries management. Journal of Fish Biology **89**:2697-2716.

Pearse, D. E., N. J. Barson, T. Nome, G. T. Gao, M. A. Campbell, A. Abadia-Cardoso, E. C. Anderson, D. E. Rundio, T. H. Williams, K. A. Naish, T. Moen, S. X. Liu, M. Kent, M. Moser, D. R. Minkley, E. B. Rondeau, M. S. O. Brieuc, S. Rod Sandve, M. R. Miller, L. Cedillo, K. Baruch, A. G. Hernandez, G. Ben-Zvi, D. Shem-Tov, O. Barad, K. Kuzishchin, J. C. Garza, S. T. Lindley, B. Koop, G. H. Thorgaard, Y. Palti, and S. Lien. 2019. Sex-dependent dominance maintains migration supergene in Rainbow Trout. Nature Ecology & Evolution **3**:1731-1743.

Pearse, D. E., M. R. Miller, A. Abadia-Cardoso, and J. C. Garza. 2014. Rapid parallel evolution of standing variation in a single, complex, genomic region is associated with life history in steelhead/Rainbow Trout. Proceedings of the Royal Society B-Biological Sciences [online serial] **281**:article 20140012.

Peterson, J. T., R. F. Thurow, and J. W. Guzevich. 2004. An evaluation of multipass electrofishing for estimating the abundance of stream-dwelling salmonids. Transactions of the American Fisheries Society **133**:462-475.

Phillis, C. C., J. W. Moore, M. Buoro, S. A. Hayes, J. C. Garza, and D. E. Pearse. 2016. Shifting thresholds: rapid evolution of migratory life histories in steelhead/Rainbow Trout, *Oncorhynchus mykiss*. Journal of Heredity **107**:51-60.

Pipal, K., M. Jessop, D. Boughton, and P. Adams. 2010a. Using dual-frequency identification sonar (DIDSON) to estimate adult steelhead escapement in the San Lorenzo River, California. California Fish and Game **96**:90-95.

Pipal, K., M. Jessop, G. Holt, and P. Adams. 2010b. Operation of dual-frequency identification sonar (DIDSON) to monitor adult steelhead (*Oncorhynchus mykiss*) in the central California coast. NOAA Technical Memorandum NMFS-SWFSC **454**.

Pipal, K. A., J. J. Notch, S. A. Hayes, and P. B. Adams. 2012. Estimating Escapement for a Low-Abundance Steelhead Population Using Dual-Frequency Identification Sonar (DIDSON). North American Journal of Fisheries Management **32**:880-893.

Pollock, K. H. 1982. A capture-recapture design robust to unequal probability of capture. Journal of Wildlife Management **46**:752-757.

Ramirez, B., S. Starcevich, and P. Barry. 2013. PIT tag antenna design and construction. Report based on PIT tag antenna workshop presented by Kurt Steinke (USFWS), September 16-19, 2013. Native Fish Investigations Program, Oregon Department of Fish and Wildlife, Corvallis, OR, USA.

Riebe, C. S., L. S. Sklar, B. T. Overstreet, and J. K. Wooster. 2014. Optimal reproduction in salmon spawning substrates linked to grain size and fish length. Water Resources Research **50**:898-918.

Rosell, F., O. Bozser, P. Collen, and H. Parker. 2005. Ecological impact of beavers *Castor fiber* and *Castor canadensis* and their ability to modify ecosystems. Mammal Review **35**:248-276.

Roth, J. C., M. H. Fawcett, D. W. Smith, J. Martini, J. Mortenson, and J. Hall. 2000. Biological and water quality monitoring in the Russian River estuary, 1999. Fourth annual report. Merritt Smith Consulting, Lafayette, CA, USA.

Royle, J. A. 2004. N-mixture models for estimating population size from spatially replicated counts. Biometrics **60**:108-115.

Satterthwaite, W. H., M. P. Beakes, E. M. Collins, D. R. Swank, J. E. Merz, R. G. Titus, S. M. Sogard, and M. Mangel. 2009. Steelhead life history on California's central coast: insights from a state-dependent model. Transactions of the American Fisheries Society **138**:532-548.

Satterthwaite, W. H., M. P. Beakes, E. M. Collins, D. R. Swank, J. E. Merz, R. G. Titus, S. M. Sogard, and M. Mangel. 2010. State-dependent life history models in a changing (and regulated) environment: steelhead in the California Central Valley. Evolutionary Applications **3**:221-243.

Satterthwaite, W. H., S. A. Hayes, J. E. Merz, S. M. Sogard, D. M. Frechette, and M. Mangel. 2012. State-dependent migration timing and use of multiple habitat types in anadromous salmonids. Transactions of the American Fisheries Society **141**:781-794.

Schaub, M., and F. Abadi. 2011. Integrated population models: a novel analysis framework for deeper insights into population dynamics. Journal of Ornithology **152**:227-237.

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. MacMillan, New York.

Shaffer, M., L. H. Watchman, W. J. I. Snape, and I. K. Latchis. 2002. Population viability analysis and conservation policy. Pages 123-142 *in* S. R. Beissinger and D. R. McCullough, editors. Population viability analysis. The University of Chicago Press, Chicago.

Skiles, T. D., R. M. Yoshiyama, and P. B. Moyle. 2013. Pink salmon (*Oncorhynchus gorbuscha*) in the Salinas River, California: new record and historical perspectives. California Fish and Game **99**:55-59.

Sparkman, M. D., M. P. Griffin, C. M. Boone, D. A. Vitali, S. Sanches, P. K. Bairrington, and M. Wheetley. 2017. Sonar estimation of California Coastal (CC) Chinook Salmon (*Oncorhynchus tshawytscha*) abundance and migration patterns in the Mad River, Humboldt County, California, 2013/14 and 2014/15. California Department of Fish and Wildlife, Anadromous Fisheries Resource Assessment and Monitoring Program.

Spence, B. C., E. P. Bjorkstedt, J. C. Garza, J. J. Smith, D. G. Hankin, D. Fuller, W. E. Jones, R. Macedo, T. H. Williams, and E. Mora. 2008. A framework for assessing the viability of threatened and endangered salmon and steelhead in the North-Central California Coast Recovery Domain. NOAA Technical Memorandum NMFS-SWFSC **423**.

Tarayama, K., K. Shin, K. Mizuno, and K. Tsuda. 2019. Integration of sonar and optical camera images using deep neural network for fish monitoring. Aquacultural Engineering **86**:102000.

Temple, G. M., and T. N. Pearsons. 2006. Evaluation of the recovery period in mark-recapture population estimates of Rainbow Trout in small streams. North American Journal of Fisheries Management **26**:941-948.

Temple, G. M., and T. N. Pearsons. 2007. Electrofishing: backpack and drift boat. Pages 95-132 *in* D. H. Johnson, B. M. Shrier, J. S. O'Neil, J. A. Knutzen, X. Augerot, T. A. O'Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Socieity, Bethesda, MD, USA.

Thompson, S. K. 2012. Sampling, third edition. John Wiley & Sons, Inc., Hoboken, NJ, USA.

Tranquilli, V. 2010. Guide to half-duplex RFID: Lessons learned tracking bull trout in the upper Willamette Basin since 2001. Oregon Department of Fish and Wildlife, https://nrimp.dfw.state.or.us/crl/default.aspx?pn=MWBTRPPITTag.

Underwood, A. J. 1994. On beyond BACI—sampling designs that might reliably detect environmental disturbances. Ecological Applications **4**:3-15.

Underwood, A. J., and M. G. Chapman. 2003. Power, precaution, Type II error and sampling design in assessment of environmental impacts. Journal of Experimental Marine Biology and Ecology **296**:49-70.

Volk, E. C., A. Blakley, S. L. Schroder, and S. M. Kuehner. 2000. Otolith chemistry reflects migratory characteristics of Pacific salmonids: Using otolith core chemistry to distinguish maternal associations with sea and freshwaters. Fisheries Research **46**:251-266.

Volkhardt, G. C., S. L. Johnson, B. A. Miller, T. E. Nickelson, and D. E. Seiler. 2007. Rotary screw traps and inclined plane screen traps. Pages 235-266 *in* D. H. Johnson, B. M. Shrier, J. S. O'Neil, J. A. Knutzen, X. Augerot, T. A. O'Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Society, Bethesda, Maryland.

Waite, E. M., G. P. Closs, J. Kim, B. Barry, A. Markwitz, and R. Fitzpatrick. 2008. The strontium content of roe collected from spawning brown trout *Salmo trutta* L. reflects recent otolith microchemistry. Journal of Fish Biology **72**:1847-1854.

Walters, C. J. 1997. Challenges in adaptive management of riparian and coastal ecosystems. Conservation Ecology **1(2)**.

Waples, R. S. 1991. Pacific salmon, *Oncorhynchus* spp., and the definition of "species" under the Endangered Species Act. Marine Fisheries Review **53**:11-22.

Watt, A., and N. Eng. 2014. Database design, second edition. Available from https://opentextbc.ca/dbdesign01/. BCcampus, Victoria, Canada.

Williams, T. H., B. C. Spence, D. A. Boughton, R. C. Johnson, L. G. Crozier, N. J. Mantua, M. R. O'Farrell, and S. T. Lindley. 2016. Viability assessment for Pacific salmon and steelhead listed under the Endangered Species Act: Southwest. NOAA Technical Memorandum NMFS-SWFSC **564**.

Williams, T. H., B. C. Spence, W. Duffy, D. Hillemeier, G. Kautsky, T. E. Lisle, M. McCain, T. E. Nickelson, E. Mora, and T. Pearson. 2008. Framework for assessing viability of threatened coho salmon in the Southern Oregon/Northern California Coast Evolutionarily Significant Unit. NOAA Technical Memorandum NMFS-SWFSC **432**:96 pp.

Yin, T., X. Zang, Z. Hou, P. T. Jacobson, R. P. Mueller, and Z. Deng. 2020. Bridging the gap between laboratory and field experiments in American Eel detection using transfer learning and convolutional neural network. Pages 932-939 *in* Proceedings of the 53rd Hawaii International Conference on System Sciences. Shidler College of Business, University of Hawaii at Manoa, Grand Wailea, Maui, Hawaii, USA.

Zimmerman, C. E., and G. H. Reeves. 2000. Population structure of sympatric anadromous and nonanadromous *Oncorhynchus mykiss*: evidence from spawning surveys and otolith microchemistry. Canadian Journal of Fisheries and Aquatic Sciences **57**:2152-2162.

Zimmerman, C. E., and L. M. Zabkar. 2007. Wiers. Pages 385-398 *in* D. H. Johnson, B. M. Shrier, J. S. O'Neil, J. A. Knutzen, X. Augerot, T. A. O'Neil, and T. N. Pearsons, editors. Salmonid field protocols handbook: techniques for assessing status and trends in salmon and trout populations. American Fisheries Socieity, Bethesda, MD, USA.

APPENDIX A. CORE POPULATIONS

Table A-1. Core 1/2/3 rating scheme for populations of the south-central California coast DPS of steelhead.

BPG	Core 1	Core 2	Core 3
Interior Coast Range	Pajaro R. Salinas River Populations		
Carmel Basin	Carmel R.		
Big Sur Coast	San Jose Cr. Little Sur R. Big Sur R.	Garrapata Cr. Bixby Cr.	Rocky Cr. Big Cr. Limekiln Cr. Prewitt Cr. Willow Cr. Salmon Cr.
San Luis Obispo Terrace	San Simeon Cr. Santa Rosa Cr. San Luis Obispo Cr. Pismo Cr. Arroyo Grande Cr.	San Carpoforo Cr. Arroyo de la Cruz Little Pico Cr. Pico Cr. Morro Bay & tribs	Villa Cr. Cayucos Cr. Toro Cr. Old Cr. Morro Cr.

Table A-2. Core 1/2/3 rating scheme for populations of the southern California DPS of steelhead.

BPG	Core 1	Core 2	Core 3
Monte Arido Highlands	Santa Maria R. Santa Ynez R. Ventura R. Santa Clara R.		
Conception Coast	Mission Cr. Carpinteria Cr. Rincon Cr. Goleta Slough Complex	C. Gaviota	Jalama Cr.; C. Santa Anita; Agua Caliente; C. San Onofre; Arroyo Hondo; Arroyo Quemado; Tajiguas Cr.; C. Refugio; C. Venadito; C. Corral; C. Capitan; Gato Canyon; Dos Pueblos Cyn; Eagle Cyn; Tecolote Cyn; Bell Cyn; Arroyo Burro; Montecito Cr.; Oak Cr.; S. Ysidro Cr.; Romero Cr.; Arroyo Paredon; Carpenteria Salt Marsh Complex

BPG	Core 1	Core 2	Core 3
Santa Monica Mountains	Malibu Cr. Topanga Cyn.	Arroyo Sequit*	Big Sycamore Cyn; Solstice Cr
Mojave Rim	S. Gabriel R.	S. Ana R.*	Los Angeles R.*
Santa Catalina Gulf Coast	S. Juan Cr. S. Mateo Cr. S. Margarita R. S. Luis Rey R.	S. Onofre Cr.* S. Dieguito R.*	S. Diego R.*; Sweetwater R.* Otay R.; Tijuana R.

* Core 2 & 3 populations that require abundance monitoring to meet viability criteria.

TABLE A-3. Core 1 populations designated by Recovery Plans for recovering to viability, and existing monitoring activities as of the last status review update (Williams et al. 2016). Note that monitoring activities are not necessarily standardized and differ from the recommendations of this plan because they pre-date it.

Population	Adult Abundance?	Spatial Structure?	Smolt Counts?
South-Central California Coast DPS			
Interior Coast Range populations			
Pajaro River	N	I	N
Salinas River pops.	Y	I	B
Carmel River population			
Carmel River	B	I	N
Big Sur Coast populations			
San Jose Creek	N	N	N
Little Sur River	N	N	N
Big Sur River	B*	N	N
San Luis Obispo Terrace			
San Simeon Creek	N	N	N
Santa Rosa Creek	N	N	N
San Luis Obispo Creek	B*	N	N
Pismo Creek	N	N	N
Arroyo Grande Creek	N	N	N

Population	Adult Abundance?	Spatial Structure?	Smolt Counts?
Southern California Coast DPS			
Monte Arido Highlands populations			
Santa Maria River	N	N	N
Santa Ynez River	B	Y(I)	B
Ventura River	B	Y(I)	B
Santa Clara River	B	N	B
Conception Coast populations			
Canada de la Gaviota	N	N	N
Goleta Slough complex	N	N	N
Mission Creek	N	N	N
Carpenteria Creek	Y	N	N
Rincon Creek	N	N	N
Santa Monica Mountains populations			
Arroyo Sequit	B	Y	Y
Malibu Creek	B	Y	Y
Topanga Canyon	B	Y	Y
Mojave Rim populations			
San Gabriel River	N	N	N
Santa Ana River	N	N	N
Santa Catalina Gulf Coast populations			
San Juan Creek	N	N	N
San Mateo Creek	N	N	N
San Onofre Creek	N	N	N
Santa Margarita River	N	N	N
San Luis Rey River	N	N	N
San Dieguito River	N	N	N

Y = Yes, N = No, B = estimates are likely biased (B* = redd counts, which can be bias-corrected with data from life-cycle monitoring stations), I = index reaches rather than randomly sampled reaches or complete census of anadromous habitat.

APPENDIX B. ESTIMATORS

Estimators summarized below are mostly adapted from Thompson (2012), in keeping with the original vision of Adams et al. (2011), with additional information drawn from the statistical literature as needed. For simplicity the estimators focus on status rather than trend, omitting the tighter variances that can sometimes be obtained from estimators tailored for rotating panels, spatially balanced sampling, hierarchical nuisance parameters, covariates, and the like.

LOW-FLOW SURVEYS

Here we assume that prior to sampling, a number n of wet reaches was established as the desired sample size. Field surveys sample reaches randomly, recording which reaches are dry, until w wet reaches and n sampleable wet reaches have been identified (see Table 7). Electrofishing is then used to conduct closed-population mark-recapture at each sampleable reach. Below we outline the various estimates to make from low-flow surveys, and make extensive use of the ratio estimator described in Thompson (2012) and many others.

Wet fraction of reaches. To produce a simple random sample, a large sample (>> n) of reaches from the target of estimation are listed in a random order, and surveyors work down the list to the n^{th} wet, sampleable reach and then stop, so that the number of dry reaches found in the process is a random variable, and the number of wet reaches is $w \geq n$ (see Figure 12). The strategy of sampling until a prescribed number pass an initial screening is sometimes called inverse sampling. A minimum-variance unbiased estimator for the wet fraction under inverse sampling (Mohammadi 2018) is:

Eq. 3 $$\hat{f}_w = \frac{w-1}{w+d-1}$$

where w is the number of wet reaches (both sampleable and unsampleable), and d is the number of dry reaches counted at the time the n^{th} sampleable reach is identified (see Figure 12). An approximate estimate of variance for the wet fraction is given in Table B-1.

TABLE B-1. Approximations for the variance of wet fraction under inverse sampling.[1]

	Condition[2]	Approximate Estimator of Variance
Eq. 4	$\hat{f}_w < \frac{(w-2)\sqrt{N}+w-1}{N}$	$\hat{V}_{\hat{f}w} \approx \frac{\hat{f}_w(1-\hat{f}_w)}{1+(N(w-2))/(N\hat{f}_w-w+1)}$
Eq. 5	$\hat{f}_w > \frac{(w-2)\sqrt{N}+w-1}{N}$	$\hat{V}_{\hat{f}w} \approx \frac{2\hat{f}_w(1-\hat{f}_w)}{w-2(1-\hat{f}_w)+\sqrt{\left(w-2(1-\hat{f}_w)\right)^2+4\hat{f}_w(1-\hat{f}_w)}}$

[1] Mohammadi (2018).

[2] Two approximations are given for variance, with the best one for a given situation depending on the estimated value. ***N*** is the total number of reaches in the sampling frame for the target of estimation, and the top or bottom approximation is closer to the true unbiased estimator depending on the condition on the left-hand side.

Density. Here, every sample reach is electrofished with two passes, mark-recapture, and closed-population assumptions (i.e., block nets are set). In addition to fish sampling, the wetted area is estimated. For reaches with continuous surface flow, wetted area is product of the length of sampled channel (channel distance between block nets) and its average wetted width. For reaches with isolated pools, it is the product of the total length of the pools and their average width.

First compute reach-level estimates of fish abundance. Use the mark-recapture estimator (Thompson 2012, eq. 18.5) for the abundance of fish at each reach *i* in the set of *n* sample reaches,

Eq. 6 $$\widehat{M}_i = \frac{(f_i+1)(l_i+1)}{b_i+1} - 1$$

where f_i is the number of fish captured on the first pass, l_i is the number captured on the last pass, and b_i is the number captured on both passes. An approximately unbiased estimator for variance of each abundance is the "Chapman estimator,"

Eq. 7 $$\hat{V}_{\widehat{M}[i]} = \frac{(f_i+1)(l_i+1)(f_i-b_i)(l_i-b_i)}{(b_i+1)^2(b_i+2)} - 1$$

(Thompson 2012, eq. 18.6). The estimate for fish density in the wet portion of the target of estimation is

Eq. 8 $$\widehat{D}_w = \frac{\sum_{i=1}^{n} \widehat{M}_i}{\sum_{i=1}^{n} A_i}$$

where A_i is an estimate of wetted area from each sample reach (distance between block nets × mean wetted width). The next steps are to calculate some variances that will be combined to obtain the sample variance for the above estimate.

The variance for the above ratio itself is estimated as

Eq. 9 $$\hat{V}_r = \frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{M}_i - \hat{D}_w A_i\right)^2$$

The variance for the measurement error from the mark recapture events is calculated as the average over all events,

Eq. 10 $$\hat{V}_m = \frac{1}{n}\sum_{i=1}^{n}\hat{V}_{\hat{M}[i]}$$

Finally, the approximate sampling variance for the estimated density (Eq. 8) is calculated as

Eq. 11 $$\hat{V}_{\hat{D}w} = \frac{1}{\bar{A}^2}\left[\frac{\left(N\hat{f}_w - n\right)}{nN\hat{f}_w}\hat{V}_r + \frac{1}{N\hat{f}_w}\hat{V}_m\right]$$

where Ā is the average wetted stream area for the n sampled reaches, and $N\hat{f}_w$ is an estimate of the total number of wet stream reaches. The X% confidence limits for the estimate can then be estimated as

Eq. 12 $$\hat{D}_w + \left(\hat{V}_{\hat{D}w}\right)^{1/2} t\left(\frac{1}{2} \pm \frac{X}{200}, n-1\right)$$

where $t(p, df)$ is a Student-t distribution with probabily p and degrees of freedom df and X is the desired confidence interval (e.g., X=95 for 95% confidence interval).

Density with Double Sampling. In double-sampling, the sample of wet reaches is surveyed using snorkel surveys. These biased data, known as auxiliary data, are calibrated by randomly selecting a subset of k reaches from the sample, and conducting mark-recapture estimates of abundance using two passes of an electrofisher. Snorkel surveys should be conducted *before* the mark-recapture at calibration reaches, so that the snorkel counts are identically distributed (i.e., unshocked, naïve fish are uniformly surveyed) at both types of reaches.

With double-sampling, we recommend estimating density using the ratio-estimator (Thompson 2012, §14). In general we assume measurement error of the auxiliary variable and of reach area to be negligible compared to measurement error from the mark-recapture estimate.

First calculate the estimates from Eq. 6 and Eq. 7 described above, for the mark-recapture data from the *k* calibration reaches rather than the full set of *n* sample reaches. Then calculate the ratios

Eq. 13 $$r_1 = \frac{\sum_{i=1}^{n} S_i}{\sum_{i=1}^{n} A_i}$$

Eq. 14 $$r_2 = \frac{\sum_{i=1}^{k} \widehat{M}_i}{\sum_{i=1}^{k} S_i}$$

where S_i is a snorkel count at each reach, and A_i is an estimate of wetted area of the sampled portion of the reach (between the block nets). Note that the first ratio is calculated over the larger sample of *n* wet reaches, and the second is only calculated over the *k* calibration reaches.

Then, estimate the density for the entire target of estimation as

Eq. 15 $$\widehat{D}_w = r_1 r_2$$

To get the sampling variance for this estimate, first calculate three components of that variance. The variances for r_1 and r_2 are estimated as

Eq. 16 $$\widehat{V}_{r1} = \frac{1}{n-1}\sum_{i=1}^{n}(S_i - r_1 A_i)^2$$

Eq. 17 $$\widehat{V}_{r2} = \frac{1}{k-1}\sum_{i=1}^{k}\left(\widehat{M}_i - r_2 S_i\right)^2$$

and the variance for the measurement error from the mark recapture events is calculated as the average over all events,

Eq. 18 $$\hat{V}_m = \frac{1}{k}\sum_{i=1}^{k} \hat{V}_{\hat{M}[i]}$$

Finally, the approximate sampling variance for the estimated density (Eq. 15) is calculated as

Eq. 19 $$\hat{V}_{\hat{D}w} = \frac{\bar{S}^2}{\bar{S}_k^2\ \bar{A}^2}\left[\frac{\left(N\hat{f}_w - n\right)}{N\hat{f}_w n}\hat{V}_{r1} + \frac{(n-k)}{nk}\hat{V}_{r2} + \frac{1}{N\hat{f}_w}\hat{V}_m\right]$$

where $\bar{A}$ and $\bar{S}$ are the average wetted stream area and snorkel count for the *n* sampled reaches, respectively, and $\bar{S}_k$ is the average snorkel count for only the *k* calibration reaches. The term $N\hat{f}_w$ is an estimate of the total number of wet stream reaches, with *N* being the total number of reaches available for sampling in the target of estimation, and $\hat{f}_w$ estimated from Eq. 3.

The X% confidence limits for the estimate can then be estimated as

Eq. 20 $$\hat{D}_w + \left(\hat{V}_{\hat{D}w}\right)^{1/2} t\left(\frac{1}{2} \pm \frac{X}{200}, n-1\right)$$

where $t(p, df)$ is a Student-t distribution with probability *p* and degrees of freedom *df* and *X* is the desired confidence interval (e.g., *X*=95 for 95% confidence interval).

ADULT ABUNDANCE

Redd Surveys. For redd surveys, abundance of anadromous adults can be estimated as described in Adams et al. (2011), to which we refer the reader. One potential complication stems from extreme low abundance of redds. Due to rarity, samples may sometimes turn up zero reaches with redds (all sampled reaches have zero redds). In such cases the best estimate of abundance is zero, but it is still possible to estimate an upper bound (non-zero abundance) that is consistent with the data, using Bayesian statistical models (e.g., Huso et al. 2015).

Counting Stations. In principle, counting stations produce a census—a complete count of anadromous adults—and do not require estimation methods. In practice, the count consists of fish moving downstream as well as upstream. Some of the downstream movers may be post-spawning kelts, in which case they can be omitted from the final count, but other downstream movers may be fish moving up and downstream prior to spawning, potentially producing double-counts of an unknown fraction of the run. Pipal et al. (2010b) discuss this issue and provide some conventions for accounting for the ambiguity in the final estimate. Ultimately there is a need for additional research on how to estimate double-counts from such datasets.

Pit-Tagging Smolts. This approach can be added on to efforts to estimate smolt production using outmigrant traps, such as rotary-screw traps. All or a portion of the captured smolts are implanted with PIT tags prior to release back to the river, and their return as adults one to four years later are monitored using PIT-tag detection stations in the lower river. Typically, the detection station will have two independent antennas for detecting tagged fish, as this allows the detection efficiency to be estimated.

Estimating anadromous abundance is a four-step procedure. First, smolt production is estimated for each year that contributed tagged adults to the current run being estimated (see next section). Second, marine survival is estimated for each of the smolt-years from the tagging data. Adult abundance for each smolt-year is then estimated as smolt production times marine survival. In the final step, total adult abundance is summed from the estimates for each different smolt-year contributing the current run.

Smolt production is estimated as in the next section. Marine survival for smolt year *y* is estimated using the mark-recapture estimator (Thompson 2012, eq. 18.5), modified for the total number of tagged fish that were released in year *y*,

Eq. 21 $$\hat{m}_y = \frac{1}{t_y}\left[\frac{(f_y+1)(l_y+1)}{b_y+1} - 1\right]$$

where t_y is the number of smolts tagged in year *y* and the remaining variables (f_y, l_y and b_y) are counts of tagged fish from that year that were detected returning in the current year for which the run is being estimated: f_y is the count of tagged fish from year *y* that were detected at the first antenna of the monitoring station, l_y is the number detected at the second antenna, and b_y is the number detected at both. An approximately unbiased estimator for variance of marine survival of this smolt group is

Eq. 22 $$\hat{V}_{\hat{m}_y} = \frac{1}{{t_y}^2}\left[\frac{(f_y+1)\left(l_y+1\right)(f_y-b_y)(l_y-b_y)}{(b_y+1)^2(b_y+2)} - 1\right]$$

(Thompson 2012, eq. 18.6).

The number of anadromous adults from smolt year y is then estimated as

Eq. 23 $$\hat{T}_y = \widehat{m}_y \hat{S}_y$$

where $\widehat{S}_y$ is the estimate of smolt production from that year (Eq. 50 for a particular year). Variance of the estimator is

Eq. 24 $$\hat{V}_{\hat{T}_y} = \hat{V}_{\hat{S}_y} \hat{V}_{\widehat{m}_y} + \hat{S}_y^2 \hat{V}_{\widehat{m}_y} + \widehat{m}_y^2 \hat{V}_{\hat{S}_y}$$

To obtain total abundance of anadromous adults, simply add up the abundance for each smolt year,

Eq. 25 $$\hat{T} = \sum_y \hat{T}_y$$

Treating the estimates for each smolt-year as independent, the variances sum as well,

Eq. 26 $$\hat{V}_{\hat{T}} = \sum_y \hat{V}_{\hat{T}_y}$$

The independence assumption is valid if the counts in Eq. 22 are compiled separately for each smolt year. If too few detections make it infeasible to estimate separate marine survivals for each smolt year, an alternative is to sum up the smolt production and tag counts across all smolt years y. Development of estimators for this approach would need to be developed.

Abundance of Rainbow Trout. Every wet reach is electrofished, with two passes, mark-recapture, and closed-population assumptions (i.e., block nets are set). In addition to fish sampling, the length (channel distance between block nets) is measured for each sample reach, and captured fish are categorized as adult Rainbow Trout or not (parr, smolts), using predefined criteria.

In general, we do not expect enough captures of adults to provide well-constrained mark-recapture estimates, so instead we take a different tack. As with density, mark-recapture is used to estimate the total number of *O. mykiss* at each reach, which is then multiplied by the fraction of adults in the sample of captured fish. The underlying assumption is that adult fish are captured at the same rate as parr.

First compute reach-level estimates of fish abundance, using the mark-recapture estimators in Eq. 6 and Eq. 7. Also sum up the number of captured fish for each sample as

Eq. 27 $$m_i = f_i + l_i - b_i$$

where f_i, l_i, and b_i are counts of fish caught on the first pass, second pass, or both passes of the mark-recapture session. Calculate similar counts for the number of fish meeting the criteria for adult Rainbow Trout,

Eq. 28 $$a_i = f_i^* + l_i^* - b_i^*$$

where in this case the f_i^*, l_i^* and b_i^* are the counts of fish meeting the criteria for adult Rainbow Trout on the first, last and both passes.

Then estimate the linear density of adults (per meter of stream reach),

Eq. 29 $$\widehat{D}_R = \frac{\sum_{i=1}^{n} a_i \sum_{i=1}^{n} \widehat{M}_i}{\sum_{i=1}^{n} m_i \sum_{i=1}^{n} L_{2,i}}$$

where $L_{2,i}$ is the sampled length of each sample reach, as in Figure 8, and $\hat{M}_i$ is from Eq. 6.

The estimator for the total abundance of adult Rainbow Trout is

Eq. 30 $$\widehat{T}_R = N \hat{f}_w \bar{L}_1 \widehat{D}_R$$

where $\hat{f}_w$ is the estimate of wet fraction (Eq. 3) and $\bar{L}_1$ is the mean GIS length of all reaches in the sample frame (the *N* reaches in the target of estimation), as in in Figure 8. To get the approximate sampling variance, first compute the ratio variance,

Eq. 31 $$\widehat{V}_{r3} = \frac{1}{n-1} \sum_{i=1}^{n} \left(\widehat{M}_i a_i - \widehat{D}_R L_{2,i} m_i \right)^2$$

Then compute the approximate sampling variance for the adult linear density,

Eq. 32 $$\widehat{V}_{\widehat{D}R} = \frac{1}{\bar{L}_2^2 \bar{m}^2} \left[\frac{\left(N\hat{f}_w - n \right)}{N\hat{f}_w n} \widehat{V}_{r3} + \frac{1}{N\hat{f}_w} \widehat{V}_m \right]$$

where $\bar{L}_2$ and $\bar{m}$ are the means for sampled reach lengths and captured fish ($L_{2,i}$ and m_i respectively) in the *n* sampled reaches, and $\hat{V}_m$ is from Eq. 10. Finally, the approximate sampling variance for the estimated adult abundance (Eq. 30) is calculated as

Eq. 33 $$\hat{V}_{\hat{T}R} = \hat{V}_{\hat{D}R}\left(N\hat{f}_w\bar{L}_1\right)^2$$

with *N* being the total number of reaches available for sampling in the target of estimation, and $\hat{f}_w$ estimated from Eq. 3. The X% confidence limits for the estimate can then be estimated as

Eq. 34 $$\hat{T}_R + \left(\hat{V}_{\hat{T}R}\right)^{1/2} t\left(\frac{1}{2} \pm \frac{X}{200}, n-1\right)$$

where *t*(*p*, *df*) is a Student-t distribution with probability *p* and degrees of freedom *df* and *X* is the desired confidence interval (e.g., *X*=95 for 95% confidence interval).

Rainbow Trout Abundance with Double Sampling. A double-sampling approach can also be used to estimate adult abundance if snorkel surveys can provide separate counts for adults (above the size threshold as visually estimated by survey teams). A subset of *k* calibration reaches are electrofished after snorkeling, to produce mark-recapture estimates.

To estimate abundance, first, at the *k* calibration reaches, estimate abundance $\hat{M}_i$ at each reach using Eq. 6. The full set of *n* reaches (including calibration reaches) will have snorkel counts S_i as well as counts R_i of the subset of fish that met the criteria for adult Rainbow Trout. From these data, estimate the ratios

Eq. 35 $$r_4 = \frac{\sum_{i=1}^{n} R_i}{\sum_{i=1}^{n} L_{2,i}}$$

Eq. 36 $$r_5 = \frac{\sum_{i=1}^{k} \hat{M}_i}{\sum_{i=1}^{k} S_i}$$

where $L_{2,i}$ is the surveyed length of each sample reach, as in Figure 8.

Then, estimate the adult density (per unit length of stream channel) for the entire target of estimation as

Eq. 37 $$\hat{D}_R = r_4 r_5$$

The estimator for the total abundance of adult Rainbow Trout is

Eq. 38 $$\hat{T}_R = N\hat{f}_w\bar{L}_1\hat{D}_R$$

where $\hat{f}_w$ is the estimate of wet fraction (Eq. 3) and $\bar{L}_1$ is the mean GIS length of all reaches in the sample frame (the *N* reaches in the target of estimation), as in in Figure 8. To get the approximate sampling variance, first compute the ratio variances for r_4 and r_5,

Eq. 39 $$\hat{V}_{r4} = \frac{1}{n-1}\sum_{i=1}^{n}\left(R_i - r_1L_{2,i}\right)^2$$

Eq. 40 $$\hat{V}_{r5} = \frac{1}{k-1}\sum_{i=1}^{k}\left(\hat{M}_i - r_2S_i\right)^2$$

and the variance for the measurement error from the mark-recapture events is calculated as the average over all calibration reaches,

Eq. 41 $$\hat{V}_m = \frac{1}{k}\sum_{i=1}^{k}\hat{V}_{\hat{M}[i]}$$

Then compute the approximate sampling variance for the adult linear density (Eq. 37),

Eq. 42 $$\hat{V}_{\hat{D}R} = \frac{1}{\bar{L}_2^2\bar{S}_k^2}\left[\frac{\left(N\hat{f}_w - n\right)}{N\hat{f}_w n}\hat{V}_{r4} + \frac{(n-k)}{nk}\hat{V}_{r5} + \frac{1}{N\hat{f}_w}\hat{V}_m\right]$$

where $\bar{L}_2$ and $\bar{S}_k$ are the means for sampled reach lengths and snorkel counts ($L_{2,i}$ and S_i respectively), the first from the *n* sampled reaches, the second from the *k* calibration reaches. Finally, the approximate sampling variance for the estimated adult abundance (Eq. 38) is calculated as

Eq. 43 $$\hat{V}_{\hat{T}R} = \hat{V}_{\hat{D}R}\left(N\hat{f}_w\bar{L}_1\right)^2$$

with *N* being the total number of reaches available for sampling in the target of estimation, and $\hat{f}_w$ estimated from Eq. 3. Confidence intervals can be calculated using the *t*-distribution as in Eq. 34.

SMOLT PRODUCTION

Outmigrant Trapping. Volkhardt et al. (2007) provide an introduction to various designs in which outmigrants are trapped, and a portion are tagged or marked prior to release to estimate trap efficiency. They provide estimators of smolt production for the various designs, to which we refer the reader. Bjorkstedt (2005) provides estimation procedures for stratified designs, which can be useful when trap efficiency and/or outmigration varies greatly within a season, which is typically the case.

PIT-Tagging Juveniles. Here, a proportion of fish is tagged during the low-flow survey, and a PIT-tag monitoring station is operated the following winter and spring. Smolt production is simply the product of juvenile abundance, from the low-flow survey; and smolting rate, from the tagged outmigrants detected at the monitoring station. For simplicity we assume that the site of capture of a fish is independent of its smolting probability.

For juvenile abundance, first compute reach-level estimates of fish abundance, using the mark-recapture estimators in Eq. 6 and Eq. 7. Then estimate the linear density of juvenile fish per wetted channel length,

Eq. 44 $$\hat{D}_J = \frac{\sum_{i=1}^{n}\hat{M}_i\sum_{i=1}^{n}(m_i - a_i)}{\sum_{i=1}^{n}L_{2,i}\sum_{i=1}^{n}(m_i)}$$

not to be confused with $\hat{D}_w$, which estimates density per wetted area rather than channel length. The counts m_i and a_i are from Eq. 27 and Eq. 28 and discount for the presence of adults that are unlikely to smolt. $L_{2,i}$ is the sampled length of each sample reach, as in Figure 8. The estimator for the total abundance of juveniles that may potentially smolt is

Eq. 45 $$\hat{T}_J = N\hat{f}_w\hat{D}_J\bar{L}_1$$

where $\hat{f}_w$ is the estimate of wet fraction (Eq. 3), and $\bar{L}_1$ is the mean GIS length of all reaches in the sample frame (the *N* reaches in the target of estimation), as in in Figure 8. To get the approximate sampling variance, first compute the ratio variance,

Eq. 46 $$\hat{V}_{r6} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{M}_i(m_i - a_i) - \hat{D}_J L_{2,i} m_i\right)^2$$

The approximate sampling variance for the estimated abundance (Eq. 45) is calculated as

Eq. 47 $$\hat{V}_{\hat{T}J} = \left(\frac{N\hat{f}_w\bar{L}_1}{\bar{L}_2\bar{m}}\right)^2\left[\frac{\left(N\hat{f}_w - w\right)}{N\hat{f}_w w}\hat{V}_{r6} + \frac{1}{N\hat{f}_w}\hat{V}_m\right]$$

with *N* being the total number of reaches available for sampling in the target of estimation, and $\hat{f}_w$ estimated from Eq. 3.

Smolting rate is estimated using the mark-recapture estimator (Thompson 2012, eq. 18.5), modified for the total number of tagged fish,

Eq. 48 $$\hat{s} = \frac{1}{t}\left[\frac{(f+1)(l+1)}{b+1} - 1\right]$$

where *t* is the number of fish tagged the previous fall, *f* is the number of those tagged fish that were detected at the first antenna of the monitoring station, *l* is the number detected at the second antenna, and *b* is the number detected at both. An approximately unbiased estimator for variance of this rate is

Eq. 49 $$\hat{V}_{\hat{s}} = \frac{1}{t^2}\left[\frac{(f+1)(l+1)(f-b)(l-b)}{(b+1)^2(b+2)} - 1\right]$$

(Thompson 2012, eq. 18.6).

Smolt production is then estimated as

Eq. 50 $$\hat{S} = \hat{s}\hat{T}_J$$

with sampling variance of

Eq. 51 $$\hat{V}_{\hat{S}} = \hat{V}_{\hat{T}J}\hat{V}_{\hat{s}} + \hat{T}_J^2\hat{V}_{\hat{s}} + \hat{s}^2\hat{V}_{\hat{T}J}$$

Note that Eq. 51 assumes covariance between the estimates of juvenile abundance and smolting rate is zero.

ANADROMOUS HAPLOTYPE FRACTION

Here we want the proportion of anadromy-associated haplotypes. This estimate is for all the fish in the target, not the average proportion for all the reaches in the target, and is estimated similarly to the way adult abundance is estimated as a proportion of the total catch.

First, estimate the number of anadromous haplotypes in each sample reach,

Eq. 52 $$\hat{H}_i = \hat{M}_i(2f_i + g_i)$$

where f_i is the number of fish that were homozygous for the anadromous haplotype, g_i is the number of fish that were heterozygous, and $\hat{M}_i$ is from Eq. 6. The sample consists of all m_i fish (Eq. 27) that were captured across all sampled reaches and subsequently analyzed for genetic composition, including those that tested homozygous for the resident haplotype. Although m_i can be the total number of fish captured, it does not need to be; it can be a random subset. If the sample is a double-sampling design then in all the calculations in this section the values are drawn from the *k* calibration reaches rather than the *n* sample reaches.

Next, estimate the proportion of anadromy-associated haplotypes using the ratio estimator

Eq. 53 $$\hat{P} = \frac{\sum_{i=1}^{n}\hat{H}_i}{2\sum_{i=1}^{n} m_i\hat{M}_i}$$

where the denominator is the total number of Omy5 chromosomes in the sample, two per fish. The next step is to calculate some variances that will be combined to obtain the variance for the above estimate. First, estimate the variance for the ratio itself as:

Eq. 54 $$\hat{V}_{r7} = \frac{1}{n-1}\sum_{i=1}^{n}\left(\hat{H}_i - 2\hat{P}m_i\hat{M}_i\right)^2$$

The sampling variance for the proportion of anadromous haplotypes is approximately

Eq. 55 $$\hat{V}_{\hat{P}} = \frac{1}{4\bar{m}^2\bar{M}^2}\left[\frac{\left(N\hat{f}_w - n\right)}{nN\hat{f}_w}\hat{V}_{r7} + \frac{4}{N\hat{f}_w}\hat{V}_m\right]$$

where $\hat{V}_m$ is from Eq. 10, multiplied by 4 to account for the change in units from fish^2 to Omy5 chromosomes^2. N is the total number of reaches available for sampling in the target of estimation, $N\hat{f}_w$ is an estimate of the total number of wet reaches, and $\bar{M}$ is the mean of the $\hat{M}_i$.

PERFORMANCE OF ESTIMATORS

We tested selected estimators by simulating datasets with different levels of wet fraction, number of reaches sampled, number of calibration samples, variation in fish density among reaches, smolting rate, and variation in various sampling parameters such as detection rates (Table B-2). In all we examined 2560 different combinations of parameters, with each combination simulated 1000 times. For each combination, the mean percent bias of the 1000 estimates was computed, which should be close to 0 if the estimators are performing well. Similarly, the mean of the empirical coverage value (ECV) was computed for the estimated 95% confidence intervals. Values should be close to 95%; ECV larger than 95% indicate that estimated confidence limits are too wide on average (i.e., conservative with respect to uncertainty), whereas values less than 95% indicate confidence intervals too narrow on average (underestimating the level of uncertainty). In general, bias tended to be small, within a few percent of the true value (Table B-3, top). Estimates of 95% confidence intervals also performed well (Table B-3, bottom), although those of population density tended to be slightly too narrow, especially for the double-sampling strategy.

Figure B-1 summarizes the realized precision of estimates as a function of the number of wet reaches sampled. Precision is summarized as coefficient of variation (square root of variance divided by the estimate, presented as a percentage). The estimators expressing a ratio (wet fraction of reaches, population density and haplotype proportion) achieved the target precision (CV < 30%) with just 20 sampled reaches on average, regardless of the parameter combination. Note that this is for the average of 1000 datasets; individual datasets would achieve the target at a lower rate. Estimators that estimated a total (adult abundance, smolt production) were much more uncertain, and would likely require much larger samples than 40 reaches to achieve CV < 30%. Adult abundance could not be reliably estimated for samples of 10 or 20 reaches; many datasets ended up with zero observations of adults (omitted from the boxplots of Figure B-1). This last result is no doubt sensitive to our assumption in the simulations that reaches have proportions of adults uniformly distributed between 0% and 1%; higher values would produce better results.

TABLE B-2. Values simulated to test selected estimators.[1]

Variable	Value(s) Simulated
Number of reaches in frame	1000
Wet fraction of reaches	0.25, 0.75
Number of wet reaches sampled	10, 20, 30, 40
Number of calibration reaches	5, 10, 15
Mean density of fish across all reaches (per m^2)	0.3
SD of density among reaches	0.1, 0.3
Smolting rate	0.1, 0.5
Maximum proportion of adults in each reach	0.01
Mean probability of detecting fish while snorkeling	0.8
SD of probability of detecting fish while snorkeling	0.05, 0.2
Mean probability of catching fish while electrofishing	0.8
SD of probability of catching fish while electrofishing	0.05, 0.2
Mean probability of detecting tagged outmigrant	0.3, 0.7
SD of probability of detecting tagged outmigrant	0.05
Mean proportion of anadromous haplotypes in reaches	0.5, 0.8
SD of proportion of anadromous haplotypes among reaches	0.05, 0.2

[1] The number of unique combinations of these variables is 2560. For each unique combination, 1000 datasets were simulated and used to make 1000 estimates of each of the following parameters: wet fraction, population density, population density with double-sampling, adult abundance, smolt production, and haplotype proportion.

TABLE B-3. Bias and Empirical Coverage Rates for Estimators.

Variable	Quantiles for 2560 Combinations of Parameters				
	2.5%	25%	Median	75%	97.5%
Mean Bias (%)					
f_w (wet fraction)	-0.9	-0.3	0	0.3	0.9
D_w (population density)	-1.1	-0.3	0	0.2	1
D_w*	-0.9	0	0.3	0.9	2.9
T_R (adult abundance)	-2.2	-0.7	0	0.7	2.3
S (smolt production)	-4.2	-3	-1.9	-1	0.2
P (haplotype proportion)	-0.7	-0.1	0	0.1	0.7
Mean ECR (%; true value = 95%)					
f_w (wet fraction)	92	94	95	96	97
D_w (population density)	89	92	93	94	95
D_w*	83	86	88	90	95
T_R (adult abundance)	88	94	95	97	99
S (smolt production)	84	90	95	98	100
P (haplotype proportion)	89	95	96	98	99

* Double-sampled

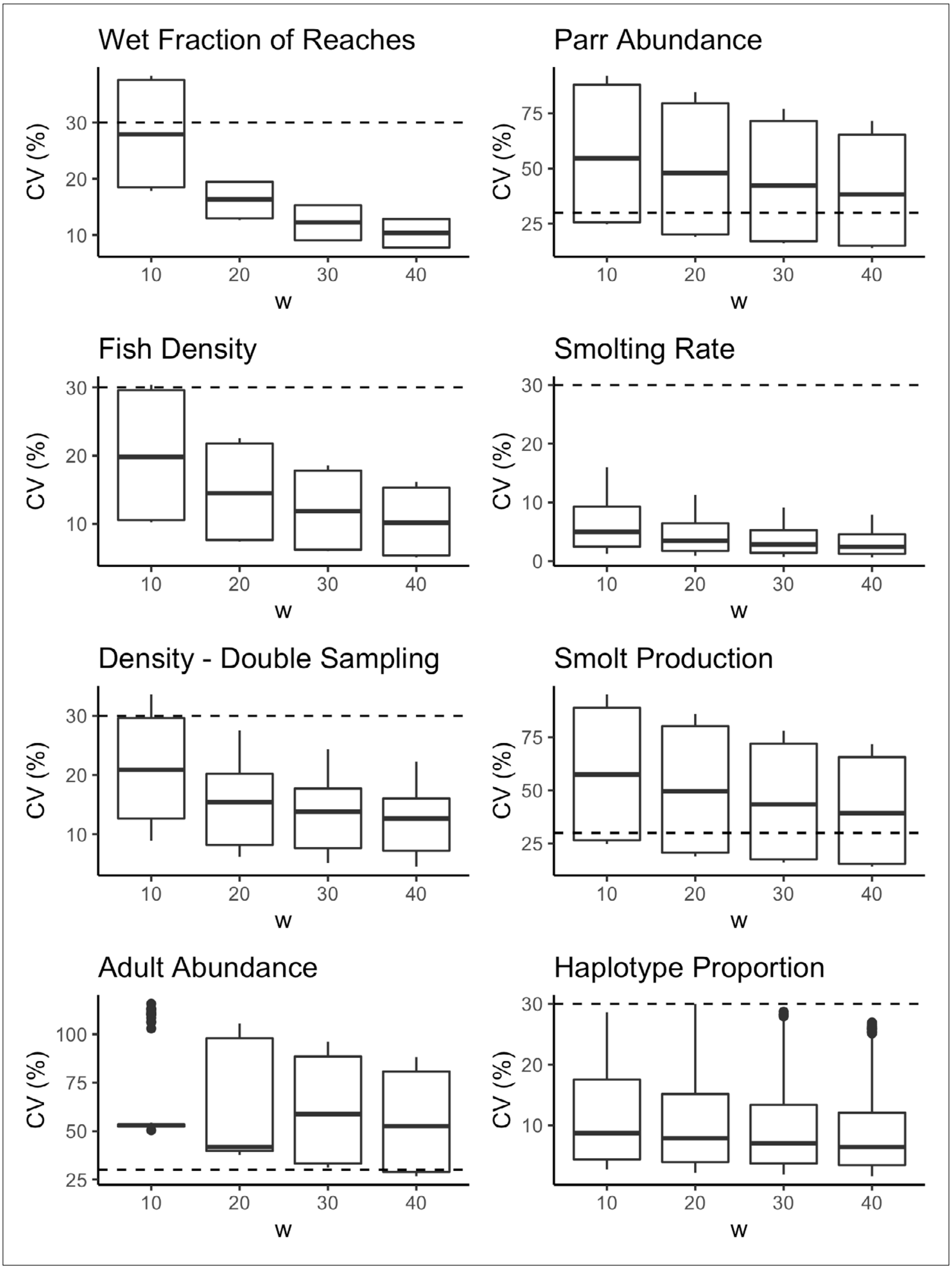

FIGURE B-1. Uncertainty (coefficient of variation) for selected estimators from the test simulations, as a function of number of reaches sampled (*w*). Dashed line marks the target CV of 30%.

APPENDIX C. INTERAGENCY ECOLOGICAL PROGRAM STEELHEAD LIFE-STAGE CRITERIA

A definitive way to determine smoltification is through a biopsy of gill tissue collected from live downstream migrants (Hayes et al. 2011) because the gills contain proteins (Na+, K+-ATPase) that regulate the movement of ions across the water/blood interface in the gills. Different isoforms, and different concentrations, of the protein are present in fish that are physiologically adapted to freshwater versus seawater, and the "switch" from the freshwater to the seawater isoform is triggered by hormones, though it can also be induced directly by exposure to seawater (see Boughton et al. 2017 for a fuller discussion). McCormick (1993) developed a non-lethal gill biopsy procedure to assay the protein and this assay is now widely used as a direct indicator of seawater tolerance, which is the central functional feature of smoltification.

However, gill biopsies are intrusive to the fish and somewhat involved to implement as a routine practice. An alternative is to use the 5-stage visual criteria for smoltification developed by IEP (1998) and reproduced in Table C-1 below:

TABLE C-1. Life-stage rating protocol for juvenile steelhead. (Table 1 in IEP 1998)

Code	Life-stage	Criteria
1	Yolk-sac fry	Newly emerged with visible yolk-sac
2	Fry	Recently emerged with yolk-sac absorbed ("button-up fry") Pigmentation undeveloped
3	Parr	Darkly pigmented with distinct parr marks No silvery coloration Scales firmly set
4	Silvery parr	Parr marks visible but faded Intermediate degree of silvering
5	Smolt	Parr marks highly faded or absent Bright silver or nearly white coloration Scales easily shed (deciduous) Black trailing edge of the caudal fin More slender body (i.e., lower condition)

Unfortunately, these criteria have not been validated against the "gold standard" of gill ATPase, so we recommend a one-time study to validate them against the ATPase assay and revise as necessary. This would involve the following steps:

1. Sample wild live juvenile fish across the range of visual stages described in the Table C-1;

2. Take photographs (under standard lighting and standard view) of each sampled fish to develop a visual record.

3. Sample gill tissue from each fish and determine smoltification using the ATPase assay.

4. Use the samples to analyze the relationship between seawater tolerance and visual stage and adjust visual criteria if necessary to improve the rating. This refinement of visual criteria would draw on the photographic record of each fish to develop novel field criteria if necessary.

Due to differences in environmental conditions (e.g., size of streams, habitat conditions, migration distance, lagoons, or lack thereof) among populations we recommend that at least two streams per BPG, each having different environmental characteristics, be sampled to validate visual criteria. For example, in the Salinas River watershed samples could be taken in Arroyo Seco River which is relatively close to the ocean and Atascadero Creek which is a substantial distance upstream. On the Big Sur coast, samples from the Big Sur River which has a large watershed and lagoon could be compared to samples from San Jose Creek which is much smaller in size and does not have a lagoon.