Independent Review Panel (IRP) Report for the 2017 Long-term Operations Biological Opinions (LOBO) Biennial Science Review

**A report to the
Delta Science Program**

Prepared by:

Dr. James A. Gore – University of Tampa (Emeritus)
Dr. Brian P. Kennedy – University of Idaho
Dr. Ronald T. Kneib – RTK Consulting Services & University of Georgia (Emeritus)
Dr. Nancy E. Monsen – Civil/Environmental Engineering Consultant
Dr. John Van Sickle – U.S. Environmental Protection Agency, Western Ecology Division (Retired)
Dr. Desirée D. Tullos – Oregon State University

January 2018

**Delta Stewardship Council
Delta Science Program**

Scope and Intent of Review: This report presents findings and opinions of the LOBO Independent Review Panel (IRP) assembled by the Delta Science Program in 2017. The intent is to provide objective feedback to the U.S. Bureau of Reclamation (Reclamation), the National Marine Fisheries Service (NMFS) and the U.S. Fish & Wildlife Service (USFWS) regarding the efficacy of regulatory actions prescribed by the agencies' Long-term Operations Biological Opinions' (LOBO) Reasonable and Prudent Alternative (RPA) actions for Central Valley water operations. The objective feedback and recommendations to the agencies are intended to inform rapid decision-making regarding system-wide water operations and effects on threatened/endangered species, evolutionarily significant units, or distinct population segments.

The last annual review was in 2015 and this is the first of two or more biennial reviews. This review primarily focuses on: (1) Stanislaus River Watershed/Eastside Division RPA actions across Water Years (WYs) 2011-2017, (2) evaluation of the Enhanced Delta Smelt Monitoring Program study plan and implementation, (3) Old and Middle Rivers (OMR) Index Demonstration Project, and (4) draft proposed Shasta RPA amendment.

After reviewing a required set of written documents (Appendix 1), the IRP convened at a public workshop in Sacramento, CA on 4-7 December 2017. The first day of the 4-day workshop included a field trip to the Stanislaus River to provide an opportunity for the IRP to observe floodplain habitat rehabilitation and gravel augmentation efforts in the Stanislaus River between Goodwin Dam and Buttonbush Park.

The second and third days included agency presentations and provided a forum for the IRP to interact and consider information presented on water operations and RPA Actions as implemented in past years or proposed for subsequent years. The IRP heard public comments near the end of each day's session. On the fourth day, the IRP deliberated in a private session beginning at 9:00 a.m. to prepare and present their initial findings at the public meeting at 2:00 p.m.

Following the IRP presentation of preliminary findings, there was an opportunity for agency representatives, members of the public, and the IRP members to comment and exchange impressions and information. Subsequent IRP communication and deliberations were conducted via email in the course of drafting this final report.

EXECUTIVE SUMMARY

Water Year (WY) 2017 was the first wet year following five consecutive years of drought conditions in California's Central Valley. It was also the first year the Long-term Operations Biological Opinions independent review panel was convened on a biennial rather than an annual schedule. The 2017 Independent Review Panel (IRP) remains positive about progress toward the incorporation of more direct links between the biological and physical components of the approaches used to guide water operations. The development of methods that explicitly link the success or failure of achieving desired temperatures, flows and other physical targets to the biological/ecological responses of the listed species is the only way that the intended goals of the RPA actions can be assessed in a scientific context. However, after nearly a decade of implemented RPA actions across water years ranging from critically dry to wet, there continues to be little evidence that declines in populations of listed species have been arrested or reversed.

The current IRP report focused on four major topics including: (1) the Stanislaus River, (2) the Enhanced Delta Smelt Monitoring (EDSM) program, (3) the Old and Middle River (OMR) index demonstration project, and (4) the draft proposed Shasta RPA amendment.

This year's LOBO review included a tour of the **Stanislaus River** that allowed the IRP to better understand the challenges faced in managing this system. It also provided a first-hand perspective on the efforts aimed at habitat enhancement for salmonids, including gravel augmentation and restructuring secondary channels in the floodplain. Land access and availability of funding are key limitations to the habitat enhancement efforts. Most importantly, it may be difficult – perhaps impossible – to demonstrate measurable positive effects of such local habitat enhancements on salmonid populations. However, the IRP discussed a number of ways that general ecosystem benefits and improvements in survival of salmonid early life history stages could be evaluated.

The Stanislaus Operations Group (SOG) appears to be functioning with a high level of coordination among the agencies and reflects a positive approach to adaptive management. With respect to the management of pulse flows and temperatures, the IRP recognized the challenges faced by SOG in meeting the co-equal goals of providing a reliable water supply and support for ecosystem functions. Water in the Stanislaus is extremely over-allocated and supplies available to Reclamation may be insufficient to meet all of the demands. In addition, the source of inflow to the reservoirs is derived from snowmelt and options for releasing water from mixed depth strata is constrained by the existing infrastructure. Given current climate change predictions, snowpack in the mountains may be severely reduced by the end of the century. In the long-term, infrastructure limitations, together with predicted reductions in cold-water inflows into the reservoirs, will only compound the challenges faced by water operators in the Stanislaus system.

Pulse flows from the reservoirs on the Stanislaus are intended to reduce the risk of dewatering redds and to trigger adult and juvenile salmonid migrations at the appropriate times. Although the shape and timing of pulse flows intended to trigger salmonid migrations could be important, little direct evidence was provided to support the contention that the pulses, as designed, are actually triggering migrations. The IRP suggested that some simple experimentation with the shape and timing of pulse flows could refine understanding of any underlying relationship between flows and migration events.

The **EDSM program** is intended to provide much needed refined estimates of Delta Smelt abundance and distribution. This was the first full year of the EDSM and the statistical design was viewed by the IRP as a positive step toward the overall goal. However, despite the rigorous statistical approach, the IRP was surprised by how much the collection of a single fish could affect the abundance estimates. Confidence intervals around the abundance estimates were very large and potential sources of that variation were considered by the IRP. The IRP identified a number of issues with the implementation of the sampling design, abundance estimates, sample volume calculations, sampling gear effectiveness and other issues (e.g., stopping rules) that may present challenges to achieving the EDSM intended objectives. It remains unclear if EDSM can fulfill its seven stated objectives and the IRP encourages the USFWS to proceed as rapidly as possible to evaluate the complete suite of objectives for this program, especially the entrainment estimation goal. A reliable estimate of Delta Smelt population size is essential for determining a jeopardy level and allowable take for this species.

The **OMR Index demonstration project** was presented to the IRP in a manner that seemed to imply that the OMR Index and the USGS Gage Data Method were separate methodologies being equally evaluated for their success in representing OMR flow. The IRP considered the advantages and disadvantages of each approach from the perspective of the agencies. However, the bottom line is that the RPA action uses the USGS Gage Data Method, while the water operators would prefer to use the OMR Index Method. The correlation between the OMR Index and the USGS Gage Data, as presented at the workshop, was inadequate for assessing how well the Index predicts the USGS Gage Data. The presented assessment also failed to focus on the critical conditions of interest, which are the tidally averaged OMR flows in the range of

-2500 cfs to -5000 cfs. There was no apparent attempt to understand the reasons underlying differences between the USGS Gage Data values and OMR Index values. The IRP provided both a detailed critique of the evaluation presented, as well as an algorithm that could provide a path to improved predictions of gage-based measurements.

IRP discussions on the draft proposed **Shasta RPA amendment** focused primarily in two areas: (1) critical temperature thresholds for survival of salmonid early life history stages, and (2) use of a 7-day average daily maximum temperature (7DADM) versus the daily average temperature (DAT) for meeting temperature compliance targets in the Sacramento River.

Critical temperatures for survival of Chinook Salmon embryos differ between laboratory and field observations. Martin et al. (2017) proposed a model to show that the discrepancy could be attributed to differences in water flow velocities between the lab and field. They suggested that water flow and temperature mediated oxygen limitation for embryos in redds and was the explanation underlying observed differences in thermal tolerance of embryos. The proposed Shasta RPA amendment uses this information to adjust temperature compliance points in the Sacramento River. The model presents a convincing case that oxygen deprivation in redds explains the discrepancy between lab and field observations, but the IRP expects that model predictions of embryo survival will still contain considerable uncertainty for a number of reasons that were discussed. Key among them is that temperature-related mortality should be distinguished from all other sources of mortality through the fry stage. Another interesting idea to explore is that temperature within the redds may be higher than that in the overlying water column due to direct radiant heating of the gravel coupled with reduced water flow velocity immediately above and within the gravel.

A broader consideration of temperature management for the Sacramento River involved the potential applicability of a conceptual model proposed by Mount et al. (2016). This model is based on experience of water managers in drought-prone regions of Australia. There are substantial differences between California’s Central Valley and the Australian system from which the model was developed, but perhaps the most important recommendation of the Mount et al. (2016) model is the essential nature of planning for drought rather than reacting to drought.

The IRP was unable to evaluate the tradeoffs between the use of 7-DADM and DAT for meeting temperature compliance targets without further analysis. Discussion of this topic centered on limitations of the 7DADM approach, the need to evaluate operational feasibility of the criterion, and the need to identify the biological relevance of the two alternatives. The IRP offered some alternative averaging approaches that could be considered in addition to the 7DADM. Included among the suggestions was a weighted moving average that represented fluctuating temperatures downstream at Jelly's Ferry better than the arithmetic mean over 7 days.

The rationale in the draft proposed RPA amendment for changing the temperature metric from DAT to 7DADM and moving the temperature compliance point upriver seems to rely on a combination of information sources; including the U.S. EPA (2003) report, consideration of predictions from the Martin et al. (2017) model, and recommendations of previous LOBO panels (2014, 2015), which offered an opinion for conserving cold-water resources by moving the temperature compliance point upriver where spawning was actually occurring. Lacking any scientific analysis to the contrary, moving the temperature compliance point upriver still seems reasonable.

Table of Contents

INTRODUCTION

Surface water resources of California's Central Valley are managed through a highly-engineered storage and delivery system to meet the needs of farms, industry, and millions of people who depend on these interconnected watersheds. A suite of rules and water rights govern the distribution of water, affecting flows and water quality of riverine and deltaic ecosystems associated with California's Central Valley. These and other anthropogenic alterations over time have been accompanied by substantive changes in aquatic flora and fauna, including a persistent decline in native fishes. With the passage of the Central Valley Project Improvement Act (CVPIA) in 1992, the U.S. Congress recognized the need for water management to consider the requirements of fish and wildlife. In 2009, California's state legislature adopted the coequal goals of improving the reliability of the water supply and protecting ecosystem health, including native fishes of the Central Valley (Delta Reform Act). Some of these fish species, or distinct populations, have been afforded federal protection under the Endangered Species Act (ESA). As a result, government agencies have been charged with developing ways of protecting these populations from further jeopardy associated, directly or indirectly, with water operation projects in the region.

Five consecutive years (2012-2016) of persistent drought recently presented a major obstacle to achieving the coequal goals of maintaining both a reliable water supply and a healthy ecosystem capable of supporting viable populations of threatened and endangered native fish species. While near-record precipitation in 2017 broke the most recent drought, much of it fell as rain and the runoff filled reservoirs more rapidly than anticipated. This type of variable oscillation in dry and wet years thus far has evaded long-term prediction and remains a persistent challenge to water operations.

Most of the natural water storage capacity in the form of historical riverine wetlands in this highly-engineered system has been lost over the last century (Cloern et al. 2016). This limits options for water management to the regulation of flow volumes and water temperatures by dams and pumping facilities. When water supplies are adequate, properly adjusting the "knobs" to control water temperature and flow volumes released into the rivers can provide acceptable results for some purposes. However, a multi-decadal warming trend in California's climate is expected to continue through the balance of this century, increasing the frequency and severity of extreme droughts and floods. Water operations in years characterized by extreme conditions will have limited options for meeting RPA actions, particularly those intended to provide cold water resources to support viable salmonid populations below dams. For this and a number of other reasons, the candid opinion of many senior salmon scientists and government policy experts familiar with conditions in the Central Valley is not optimistic with respect to the future fate of wild salmon in the system (Franks and Lackey 2015).

The IRP reiterates the point expressed by previous panels that expectations for both water supply and extant ecosystem components in California's Central Valley will be forced to adapt to a "new normal" driven in large part by climate change.

***Background on the LOBO RPA action review process*:** NOAA's National Marine Fisheries Service (NMFS) and the U.S. Fish and Wildlife Service (USFWS) have issued Biological Opinions (BiOps) on state and federal long-term water operations affecting freshwater input to the Sacramento-San Joaquin Delta ecosystems. The BiOps include Reasonable and Prudent Alternatives (RPA) designed to alleviate jeopardy to listed species and adverse modification of critical habitat. NMFS' BiOp required the U.S. Bureau of Reclamation (Reclamation) and NMFS to host a workshop each year to review the prior water year's operations and to determine if any measures prescribed in the RPA actions should be altered in light of new information (NMFS' OCAP Opinion, section 11.2.1.2 of the 2009 RPA with 2011 amendments, starting on page 9). Amendments to the RPA actions must be consistent with the underlying analysis and conclusions of the BiOps, and must not limit the effectiveness of the RPAs in avoiding jeopardy to the listed species, or result in adverse modification of critical habitat. In April 2016, NMFS and Reclamation agreed to temporarily change the frequency of the science review from annual to biennial from 2016 through 2020. Consequently, there was no annual review in 2016 and this year (2017) is the first of the biennial reviews.

The intent of the biennial review of the Long-term Operations Biological Opinions (LOBO) is to provide all of the involved agencies with an independent perspective that is useful in informing management as to the effectiveness of operations and regulatory actions prescribed by the RPAs. In addition, the LOBO aims to provide recommendations and assessments that may assist the agencies in making timely and scientifically-justified adjustments to implementation of RPA actions when necessary for future water operations.

Since the BiOps were issued, NMFS, USFWS, Reclamation, U.S. Geological Survey (USGS), California Department of Fish and Wildlife (CDFW) and the Department of Water Resources (DWR) have been performing scientific research and monitoring in concordance with the implementation of the RPAs. Technical teams and/or working groups, including the geographic divisions specified in the NMFS' BiOp, have summarized their data and results following implementation of the RPA actions within technical reports. The data and summary of findings related to the implementation of the RPAs provide the primary context for scientific review regarding the effectiveness of the RPA actions for minimizing the effects of water operations on listed species and critical habitat in the Sacramento-San Joaquin watersheds. A subset of these technical reports was provided for review by the 2017 LOBO IRP (see Appendix 1).

General charge and scope for the 2017 LOBO IRP: The first annual review of long-term water operations and the BiOps considered all of the RPA actions. In subsequent years, the panel's charge has focused on a subset of RPA actions.

This year's (2017) first biennial review included a consideration of:

(1) Stanislaus River Watershed/Eastside Division RPA actions across water years 2011-2017;

(2) Evaluation of the Enhanced Delta Smelt Monitoring Program's study plan, implementation, lessons learned, and opportunities for improvement;

(3) Results from the Old and Middle Rivers (OMR) index demonstration project; and

(4) Draft proposed Shasta RPA amendments.

As in previous years, the specific scope of the 2017 LOBO review was defined by questions posed to the IRP by the agencies and technical teams/task groups that presented materials for review. This IRP report addresses each of the 14 questions posed from a scientific perspective. In addition, the report provides observations, opinions and recommendations where, in the panel's opinion, they seemed related to the research being conducted and potentially useful to agency staff for consideration.

Acknowledgments**:** The IRP appreciates and acknowledges the efforts of the agency and technical team representatives and contractors who responded to questions and suggestions made by previous panels, prepared the written materials, organized and led the Stanislaus River tour, and delivered the workshop presentations. The IRP members are cognizant that much of the material has to be compiled, analyzed, and organized in a relatively short time. Even though the last review was two years ago, the IRP recognizes that government agency personnel faced substantial challenges in contending with effects of dry to wet annual climatic conditions, as well as a dynamic regulatory environment. Despite the many competing demands on the workshop participants, the materials were largely presented with the usual professionalism. The IRP wishes to express a special thanks to John Callaway (Lead Scientist), Lindsay Correa (Program Manager), and the staff of the Delta Science Program for providing the organization and logistical support to facilitate our task. In particular, Dylan Stern attended to a variety of technical and provisional details in support of the IRP's efforts before, during, and following the workshop. Title page photo credit: R.T. Kneib (Goodwin Dam, December 4, 2017).

COMMENTS ON CONDITIONS & RPA ACTIONS IN WATER YEARS 2011-2017

General comments and observations

As we approach nearly a decade since RPAs have been implemented, there is scant evidence of stabilization or reversal in the decline of protected fish populations in the Central Valley and Delta. While not all stressors can be controlled and mitigated by the managing agencies, several efforts could improve the ability of the agencies to document and communicate any benefits produced by RPA actions, and the ability to respond rapidly and appropriately to changing conditions. With this aim, the IRP offers the following general observations to provide additional context for the findings and suggestions in this first biennial LOBO report:

Continued need for linking RPA actions and biological responses.

It was encouraging to note the continuing effort to link physical criteria in RPA actions to biological responses, but there continues to be substantial capacity for improvement in this area. For example, as discussed in further detail on page 14, the effects of gravel augmentation may not immediately produce measurable impacts to the number of spawners or survival. However, the benefits of these projects in producing food resources for salmon could be enumerated through documentation of increases in primary and secondary productivity using benthic macroinvertebrate surveys. While increased productivity is only one benefit of the gravel augmentation, the data could be used in bioenergetics models to demonstrate how the gravel projects contribute to growth and survival, and how much more gravel is needed to support a viable population.

Expecting immediate positive population responses to RPA actions in any given year would be overly optimistic, but evaluating impacts to individual life stages is possible through the use of field observations and numerical models.

Need for distinguishing habitats from RPA action targets.

The IRP encourages agencies to better define their targets as something other than "habitat," unless they are targeting the sum of physico-chemical and biological conditions. Many reports and presenters used the term habitat to mean the number of river miles maintained at a given target temperature range. However, habitat is defined as the place where a species normally lives (Calow 1998) and includes a complete suite of physico-chemical characteristics required by that species to survive, grow and reproduce. Habitats often comprise microhabitats (e.g., conditions within salmonid redds) or sub-habitats (e.g., areas used for salmonid spawning or rearing) that are essential for the survival and growth of a particular life stage, and also include biological interactions that may restrict access to physical and chemical conditions that are suitable for survival and recovery. A certain water temperature is not the only defining characteristic of suitable habitat for salmonids. As such, individual actions alone, such as meeting temperature targets or adding gravel, do not necessary reflect creation of suitable habitats. In addition, some RPA actions (e.g., floodplain reconnection) lack identification of measurable targets for enhancement of habitat components. The result is an inability to identify biologically-relevant outcomes for success of projects and an inability to assess progress towards that success. Thus, the IRP encourages the use of caution and explicit definition when discussing habitat rehabilitation and enhancement activities.

Need for communicating lessons learned in all water years.

Given the biennial cadence of this review, and the importance of WY 2016 as a transition year, the IRP expected more information on WY 2016 in some of the report and/or presentations. Since this is now a biennial review, the IRP expects content that reflects the previous two years. There was a tendency to present comparisons between WY 2015 (the last year of an extreme drought) and WY 2017 (an extreme wet year) to demonstrate management and operational opportunities, challenges, and decisions. While the IRP appreciates the challenges in managing these difficult periods, it would have been helpful to hear the agencies synthesize the lessons that were learned from a more moderate year (WY 2016). An analysis of recovery during the first post-drought WY may have helped to show how water operations and fish species responded in a transition year within a longer period defined by precipitation extremes. The information from WY 2016 is expected to be useful in planning recovery responses after the next drought event.

Coming to terms with the new normal.

Increased frequency, severity, and duration of extreme droughts, punctuated by wet weather conditions, is characteristic of the changes projected by global circulation models of climate change. The recent occurrence and severity of extreme events suggest that weather patterns in California may already be transitioning to a “new normal.”

Depending on the assumptions made regarding carbon emissions over the coming decade, as well as the potential changes in the impacts of various climatic oscillations (e.g., Stenseth et al. 2003, Kelly and Gore 2008, Mantua et al. 1997), the projected potential loss of snowmelt water over the coming century may have serious consequences for both water operations and salmonid populations. Large-scale predictions of impending climate change suggest the likelihood of declining snow water availability over the next 100 years (Table 1).

By some estimates, snowpack in the southern mountains of the West, including the northern Sierra Nevada range, is projected to virtually disappear by the end of the 21st century (Easterling et al. 2017). More of the region's precipitation is expected to enter western watersheds as rainfall instead of snowmelt, potentially reducing the storage of available cold-water resources in reservoirs to support viable populations of endangered native fishes downstream. Precipitation events also are expected to become less evenly distributed through time, with a projected increase in both the intensity and frequency of storms, which will provide additional challenges for water operations in terms of balancing the needs for flood control and drought preparations. Therefore, some of the management options in the CVP system should be reconsidered in light of these extreme changes. As an example, forecasted components of the New Melones Index (Index) will be more uncertain, and the Index thresholds for water year type may require adjustment as a result.

Table 1. Projected April 1 deviations from "average" Snow Water Equivalents for the combined San Joaquin, Sacramento, and Trinity Rivers (derived from Cayan et al. 2008).

Elevation	Years 2005-2034	Years 2035-2064	Years 2070-2099
1000-2000 m	-13% to -48%	-26% to -68%	-60% to -93%
2000-3000 m	+12% to -33%	-8% to -36%	-25% to -79%
3000-4000 m	+19% to -13%	-2% to -16%	-2% to -55%
All Elevations	+6% to -29%	+12% to -42%	-32% to -79%

Agencies are encouraged to create and test various climate oscillation and climate change scenarios as a means of anticipating new conditions that historical records might not predict. Whether or not impacted by loss of snowmelt contributions to the system, many water resource agencies are facing the prospect of altering management strategies to anticipate changes in water availability with climatic oscillations and, more importantly, to begin to plan for climate change. While it is recognized that real-time management is occurring in the Sacramento-San Joaquin system, long-range planning for climate change is essential for setting expectations of water users and for viability of threatened and endangered species. A number of approaches exist within the literature for long-range climate change planning. As an example, the Southwest Florida Water Management District applies a wet season and dry season (each being 30 years in duration) management strategies that reflect significant changes in community structure as climatic oscillations occur (e.g., see Munson, et al. 2005, Munson and Delfino 2007, and Gore et al. 2016). Wilby (2016) and Brown et al. (2012) also suggests some strategies for managing rivers and water resources in a changing climate.

Stanislaus River Watershed/Eastside Division RPA Actions Across Water Years 2011-2017

Stanislaus River tour for the 2017 LOBO Review.

The IRP appreciated the opportunity to visit the Stanislaus catchment and the gravel bed augmentation and floodplain rehabilitation work being performed. Although the panel members are familiar with various components of the Central Valley Project (CVP), none have had the opportunity to visit all areas. Such field trips allow the IRP to better understand the topography, physical conditions, and challenges faced by operators and managers.

Ongoing efforts include creating new gravel-bed spawning areas and mitigating for loss of natural floodplain habitats in the Stanislaus River below Goodwin Dam. These actions follow a general assumption that gravel augmentation mitigates against the loss of access to upstream salmonid spawning habitat. However, the gravel volumes that have been put in place so far on the Stanislaus are only a fraction of that which was initially prescribed, and it is not clear that the RPA prescriptions are adequate to produce desired ecological outcomes. In addition, both gravel augmentation and floodplain rehabilitation are relatively expensive undertakings that, in order to meet desired outcomes, would require monitoring and maintenance. In order to justify the continued effort, it will be necessary to demonstrate a measurable benefit from these activities in terms of improvements in salmonid survival and production.

It may be difficult, if not impossible, to directly demonstrate system-wide increases in salmonid production as a result of these habitat enhancements. However, measurable improvements in local conditions known to support survival and growth of salmonid early life stages would demonstrate general ecosystem benefits. For example, one measure of success could include changes in primary production, secondary production and, perhaps forage fish production, in river reaches adjacent to the mitigation efforts.

There are at least two possible methods to demonstrate the establishment of these communities. Assuming that most of the colonization by periphyton and macroinvertebrates will be derived from upstream sources, sampling both source areas and the new sub-habitats could create an index of success. In some rivers, macroinvertebrate communities were virtually identical to upstream sources within 6 to 8 weeks of placement of a new gravel bed, while forage fish communities were identical in about a year from initiation (Gore 1982). Examination of colonization and establishment of functioning communities is an important tool in examining the success of restoring river ecosystems (Gore and Milner 1990). An alternative method that does not require comparison to sources of colonizers (if none exist upstream, for example), is to sample similar gravel bed communities along the length of the river, tributaries, and adjacent catchments to create a "reference" condition as a target for the new gravel beds (Hughes et al. 2010). Either method can be used to demonstrate success without having to rely upon production of top carnivores, which may take years to create. In addition, using a modeling system such as PHABSIM, a simple mapping technique within the model which can compare pre- and post-augmentation spawning habitat, to map or assess the increases in spawning habitat could be used to demonstrate the benefit of gravel augmentation.

Stanislaus Operations Group (SOG).

Strengths of the program include the extensive monitoring being conducted in the basin, and the integration of the field data (e.g., weir counts, water and air temperatures) with weather forecasts to focus SOG operational recommendations in real time, with direct impacts on the timing and duration of pulse flows. In addition, for the most part, the pulse flows are well justified based on biological processes. For example, the transition to a three-peak pulse and the October implementation (prior to peak spawning) is logical given the desire to avoid construction of redds at higher elevations that will be dewatered later in the year. Furthermore, there appears to be a high level of coordination between NOAA Fisheries, Reclamation, and other SOG members. This reflects a strong attempt at adaptive management.

The program could be improved in a couple of key ways. Some of these improvements are related to variability in water years, and thus are addressed under Question 3. However, specific to the SOG shaping of releases, the committee identified one key uncertainty regarding evaluating the effects of the pulses. Despite the strong qualitative logic model linking biological processes to the design of flow pulses, quantitative analysis to demonstrate that the pulse flows are indeed achieving the intended biological outcomes seems lacking. For example, the primary objective of the fall pulse flows is to trigger migration up into the Stanislaus while operators are drawing the reservoir down. While it is clear the SOG is designing the fall pulses to avoid redds from being dewatered, it is not clear that the pulses, as designed, actually trigger migration. In addition, it is not clear if, or how, the shape of the pulse impacts the timing of fish arriving in the basin. The pulse design for migration cueing could be important. However, it should be based on a more refined understanding and justification that is ideally numerical (e.g., Sykes et al. 2009) and/or determined experimentally. Similarly, field observations could be examined to investigate the effectiveness of spring pulses in achieving their outcomes of reducing temperatures, inundating shallow habitats, and flushing smolts out of the river and through the Delta. Such an analysis would be beneficial both for shaping pulses and for prioritizing areas for restoration actions.

Management of pulse flows, temperature and the challenges of interannual hydrologic extremes.

Operating the series of dams on the Stanislaus River to provide for a reliable water supply and supporting ecosystem functioning is particularly difficult for many reasons. Perhaps most fundamentally, Stanislaus River water is extremely over-allocated and the supply available to Reclamation may be insufficient to meet the demands of all flow requirements (State Water Resources Control Board 2016).

Furthermore, much of the inflow to reservoirs on the Stanislaus is from snowmelt and options for releasing water that is from mixed depth strata is limited. For example, adjusting the temperature releases from New Melones dam is not feasible under most circumstances. The cooler hypolimnion water can only be released through a tunnel at the base of the New Melones Dam at very low reservoir elevations. In addition, the majority of New Melones cool water pool is held back by an additional barrier, the submerged original Melones dam, directly upstream of the New Melones dam.

There is no universally accepted approach to maintaining species subject to hydrologic extremes, especially when much of a species' historical habitat is unavailable and the population of interest is near the limit of the species' historical range (e.g., Chinook Salmon in California). However, some key concepts can be applied in the future as a general framework for thinking about managing extreme events in modified systems. These concepts emphasize robustness (Herman et al. 2015), a condition where the species' sensitivity to the disturbances is reduced, and resiliency, where species are capable of recovering following a disturbance (Walker et al. 2004).

First, a synthesis of lessons learned would be a useful exercise for the agencies. Such a synthesis should summarize the hydrologic year at a weekly to monthly time scale, outline operational decisions in response to data and associated releases, examine measures of how operations and extreme conditions impacted fish, and identify what additional information and/or alternative actions would be needed in a future event.

Second, there is a growing body of literature on managing novel ecosystems. Even absent the effects of extreme conditions, structure and function of the Stanislaus River ecosystem was sufficiently altered over the last century to classify it as novel. Key species (e.g., Spring-run Chinook) have been lost, invasive predators have been introduced, and spawning and rearing habitats of salmonids have been relocated to historically unsuitable reaches below the dams. Hobbs et al. (2006) argues for examining the persistence and value of the new ecosystem, and for managing these types of systems in a way that is fundamentally different from current approaches. For example, Seastedt et al. (2008) suggest that management should emphasize maintaining genetic and species diversity, as well as the biogeochemical characteristics that favor the desirable species, rather than attempting to recreate historical conditions. If the system is stable, what are the costs and risks of attempting to guide the system to a more desirable state? If the system is changing, what are the costs and benefits of maintaining current species compositions? Identifying where sites fall along a range of wild to intensively modified will help managers prioritize conservation of the least impacted systems and identify systems in which it is not feasible to maintain or restore historical species communities.

Third, the increasing frequency of extreme events raises the need to revisit the guiding vision for the watershed. In unaltered river systems subject to extreme events, critical refuges from droughts and floods are provided by side channels, floodplain habitats, deep pools, and other complex habitats that are largely missing from the modern Stanislaus River. Resource managers may want to reconsider what outcomes are feasible for this system. Gravel augmentation and other habitat enhancements are likely to become increasingly important for species' survival, but there is no evidence that the agencies will be able to create and maintain enough of these habitats to produce a measurable benefit at the population level. The RPA action requiring the addition of 50,000 yd^3 of gravel by 2014, and 8,000 yd^3 annually thereafter, has been stalled by lack of funding, land access, and other issues.

Other RPA actions also lack measurable, time bound objectives (e.g., floodplain restoration, predation management) that link to biologically-relevant outcomes. Effects of RPA actions can be measured in a variety of ways (e.g., primary and secondary production, diversity, etc.), but ultimately habitat rehabilitation projects need to demonstrate meaningful connections to the viability of targeted salmonid populations. What proportional contribution to new spawning habitat is expected from 50,000 yd^3 of gravel, if that could be achieved? How many redds can that area support? Similarly, are small side channel projects at an adequate scale to produce a biological response, or are landscape-scale projects needed to provide adequate habitat for protecting fish during dry and/or wet years? Will predators consume most or all of the expected increased production of salmonids from the floodplain and gravel projects?

Finally, SOG requested the IRP to comment on the approach they introduced to update water year classifications for the New Melones Index as forecasts become available. The IRP recognizes the management challenge SOG faces in the highly variable hydrologic conditions, with water years transitioning from critically dry to extremely wet within a few months. SOG's attempts at adaptive management are laudable, and the general approach to updating the New Melones Index seems reasonable. It might seem logical to assume that, since the 90% forecasted inflow has the highest likelihood of being exceeded, the 90% exceedance will likely be most protective of fish, particularly in the driest years and for meeting late summer temperature targets. However, lacking data, the impacts of 50% or 90% exceedance criteria on extant salmonid populations remains unclear. In a 2010 report (NOAA 2010), SOG summarized a preliminary analysis on the effects of 50% and 90% exceedances. However, the data were unavailable for the IRP to review and analyses were too vaguely described. It was thus impossible for the IRP to determine if the results are conclusive or logical. A re-analysis that includes data from both drought and flood years would help clarify the biological effects of the 50% and 90% exceedance thresholds on storage for late summer temperatures. A scenario that shifts between 50% and 90% exceedances, based on initial assessments of water year conditions, could be appropriate for conserving water resources needed for later in the water year.

Enhanced Delta Smelt Monitoring (EDSM) Program

An improved method of surveying Delta Smelt in a manner that can produce refined abundance estimates has been greatly needed. The first year's (WY 2017) implementation of EDSM was a positive step in that direction. However, the IRP identified a number of issues that may present challenges to achieving the program's intended objectives.

In spite of EDSM's substantial increase in sampling effort relative to previous surveys, its catches of Delta Smelt remain quite low, with many zero-catch sampling events. The resulting abundance and distribution estimates are highly uncertain. In considering the challenges associated with monitoring the Delta Smelt population, the "Lessons Learned" and "Concerns" sections of the EDSM report were very helpful to the IRP regarding the following topics:

Sampling design.

The generalized random tessellation stratified (GRTS) design seems an appropriate approach for obtaining unbiased estimates of total Delta Smelt abundance across the entire Delta. GRTS provides randomized site selection, uniformly-dense spatial coverage, and oversample properties needed to eliminate bias. However, practical limitations on sample size and collection methods may prevent this design from adequately representing potential "hot spots" of abundance, particularly in shallower waters, and other microhabitats (e.g., spawning areas) in which Delta Smelt may congregate at certain times. Such "hot spots" may be quite small in terms of area and/or volume relative to the entire Delta while containing much higher Delta Smelt densities at certain times (e.g., see Bennett and Burau 2014). The EDSM's first-year experiments included attempts to sample some shallow habitats, but it may be advantageous to sample such potential hot spots more thoroughly by either adding sampling effort or redistributing the existing effort.

A greater-than-proportional number of sampling locations could be selected in hot spot areas via GRTS, either by defining habitat-based strata (Section 5.1.2, EDSM report) or by using unequal probability sampling. The stratification option would simplify estimation of the abundance model parameters. However, model parameters can also be estimated from unequal-probability data by using Thomas Lumley's "survey" package in R. The "survey" package will also give design-based standard errors, which may partly address the EDMS report's concerns about "wrong" variances (Section 5.2). Also, the neighborhood variance estimator of Stevens and Olsen (2003), assuming it could be applied here, would reduce variance estimates only to the extent that Delta Smelt densities changed slowly and smoothly over space relative to the scale of separation between sampling locations. Given the mobility of Delta Smelt, and the extreme patchiness of actual densities, there may be little value in pursuing neighborhood variance estimation (EDSM Section 5.2).

The current sampling design specifies some strata boundaries that violate the assumption of density of fish varying slowly and smoothly between sampling locations. For example, the Suisun Bay strata contains multiple habitat types (shallow bays and energetic deep channels). Water velocities within the Suisun Bay strata are highly variable and are tidally influenced.

Abundance estimation.

There are several aspects of the abundance estimation that require further development including volume calculations and the fish sampling methods. The zero-inflated negative binomial (ZINB) seems appropriate for modeling the large number of zero-catch samples and the apparent over-dispersion (relative to Poisson) of fish counts in positive-catch samples. This approach is well-supported by the literature (e.g., Wenger and Freeman 2008).

Given the statistical rigor involved in developing the EDSM abundance estimation approach, it was surprising to note how abundance estimates were influenced by the collection of a single fish. In the weekly reports (EDSM, Draft preliminary abundance analysis, Phase 2/Phase 3), there were estimates of zero abundance when no fish were captured, but enormous abundances (hundreds of thousands of fish) were estimated when only one fish was captured. Even though the confidence intervals on the nonzero-catch estimates were also very large, it is still difficult to find credibility in such large differences in the point estimates of abundance. In addition, the EDSM report also noted wide swings in sequential estimates of occurrence probability (π_0) and unrealistic changes over time in the total abundance estimates. These unstable, very spiky estimation results are likely caused by a combination of factors including: many zero-catch samples, very low counts when fish are caught, and the very large expansion factor of the ratio of the total water volume to the volume sampled by the tows.

The EDSM team's proposal to address these estimation problems by aggregating the tow data over space and time seems reasonable. The approach should reduce the number of zero-catch cases and increase the cumulative counts when fish are caught (EDSM, Section 5.1.1). Exploring such aggregated estimates would help to better understand how estimation uncertainty decreases with increasing degrees of aggregation.

The IRP cannot recommend pursuing the more complex alternative models mentioned in EDSM Section 5.1.1 at this time. Such models would have even more parameters, thus leading to greater estimation difficulties and greater estimation uncertainty. Instead, it may be more useful to evaluate the robustness of the current estimates by comparing them to estimates from alternative, credible models having complexity similar to the ZINB. For example, although the zero-inflated Poisson had an inferior fit quality for the available data, it was also not ruled out by the data. Its estimates would be interesting to compare with the ZINB. Another option is to specify the ZINB as a conditional model rather than a mixture model, to see how this option affects estimates (Cunningham and Lindenmayer 2005). Finally, repeated tows at the same or nearby locations could be treated as replicates, with zero catches regarded as non-detections of a rare, but present, species rather than as true absences (McKenzie et al. 2005). If abundance estimates from these different approaches are similar, then one would have more confidence in the robustness of the ZINB estimates.

Volume calculations.

A high priority should be assigned to reducing uncertainties related to the volume sampled during multiple tows (EDSM Section 5.3) and the mismatch between sampled volumes and the volumes extrapolated over (EDSM Section 5.4). The sample volume and the extrapolated-over volume are both direct multipliers for calculating abundance, so their errors would have a substantial impact on the final estimates. In moving forward with the EDSM it is important to explore and increase the robustness of the current abundance estimation method.

Perhaps fish densities expressed as surface area rather than volume would help reduce variation associated with multipliers based on volume. It is already assumed that the Delta Smelt population is restricted to the top 4 meters of the water column, which is the only depth stratum sampled (i.e., the top 0-4 meters of waters > 2 m depth) in the EDSM. Using a two-dimensional surface area as an expression of density (fish m^{-2}) versus three-dimensional volume (fish m^{-3}) reduces the multiplier effect of volume. Density based on surface area may allow for direct comparison of densities in mid-channel stations with those in shallow water using different gear types. For example, if a mid-channel tow filtered a volume of 3000 m^3 including a depth of 4 m, the surface area of the tow would be 750 m^2 (3000 m^3 / 4 m). Assuming use of a smaller gear type that was 1/10 the size in shallow water (e.g., 1 m depth) and that filtered 75 m^3 per tow, it would have collected from a surface area of 75 m^2 (75 m^3 / 1 m). Assuming that 10 fish were captured in the larger mid-channel gear and 1 fish was captured in the shallow water, the density in terms of surface water area would be the same (13.3 fish/1000 m^2) in both the mid-channel and shallow stations. However, expressing density in terms of volume yields a much lower density (3.3 fish/1000 m^3) in the mid-channel than in shallow water (13.3 fish/1000 m^3).

The bathymetry of the Delta and the connection between open water habitats and the adjacent channels also are important to consider. For instance, in the south Delta around Mildred Island, there is large open region separated from adjacent channels by levees. Lopez et al. 2006 showed through field sampling studies of primary production in this region that the connectivity and transport between Delta habitats, controlled by bathymetry and levee breaks, is important for defining habitat quality throughout the Delta. Therefore, the abundance of fish sampled in a channel cannot directly be extrapolated to an abundance within the open water region.

Fish sampling methodology.

There are a number of other methodological considerations that may contribute to the uncertainties in the abundance estimates as well. These include efficiency of the collecting gear, the effect of the stopping rules, and variation in efficiency of sampling crews. In addition, interacting behavioral and hydrologic considerations may contribute to uncertainties. These would include tidal surfing behavior of Delta Smelt, particularly during upstream spawning migrations (Bennett and Burau 2015), and the water sources that are actually being sampled in each stratum (e.g., Sacramento and San Joaquin sourced water are combined in some strata but differ in water quality attributes). The complex hydrodynamics of the Delta may result in sampling water originating from multiple sources within a given stratum.

The EDSM assumes that the Kodiak trawl is 100% efficient at collecting adult and juvenile Delta Smelt. However, no data were presented to support this assumption. Trawls are notoriously inefficient because of the way they function, essentially herding fish in front of them until the fish tire and fall back into the cod end of the net. In addition, there is usually considerable escapement due to net avoidance involved in trawling. Trawls might be better characterized as collecting gear than sampling devices. When multiple gear types are compared in the same area, one type may consistently capture more fish but this is simply a relative measure of effectiveness. Rarely are trawl hauls compared to known densities of fish in an area, but when acoustic data have been used in conjunction with trawls, trawl efficiencies are substantially lower than 100% (e.g., Hylen et al. 1995).

In order to tire fish herded in front of a trawl, the net must be towed at a speed that is close to, or greater than, the critical swimming capacity of the species of interest.

According to Swanson et al. (1998), Delta Smelt exhibit swimming performance that is comparable to other species of the same size and are capable of sustaining moderately high velocities averaging 27.6 cm sec^{-1}. This is equivalent to traveling a distance of over 165 m during a standard 10-minute tow. Swanson et al. (1998) also found that endurance was highly variable and not normally distributed, so a subset of the Delta Smelt population may be more or less susceptible to capture at trawl tow speeds approaching the maximum sustainable swimming speed of the smelt. This also suggests that the location of fish along a tow path may contribute to variation in catch. The farther along the tow path fish are first encountered, the less likely they are to be captured. This concept is illustrated in Figure 1.

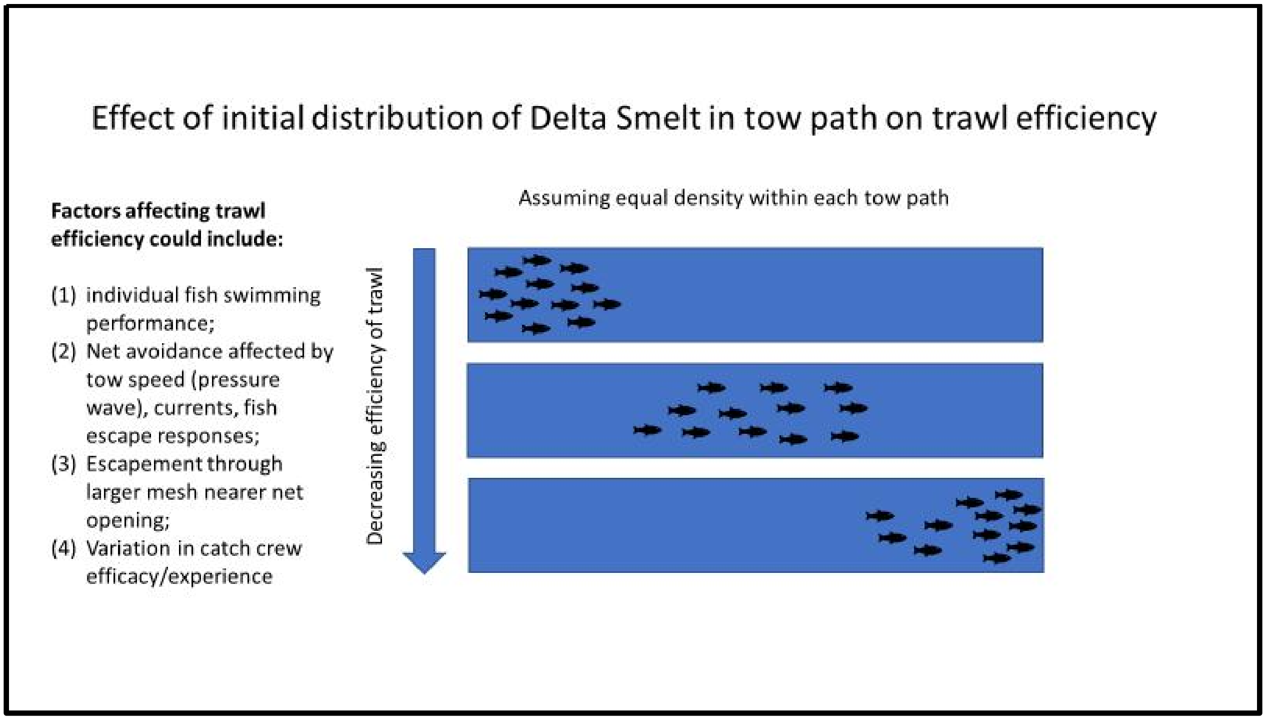
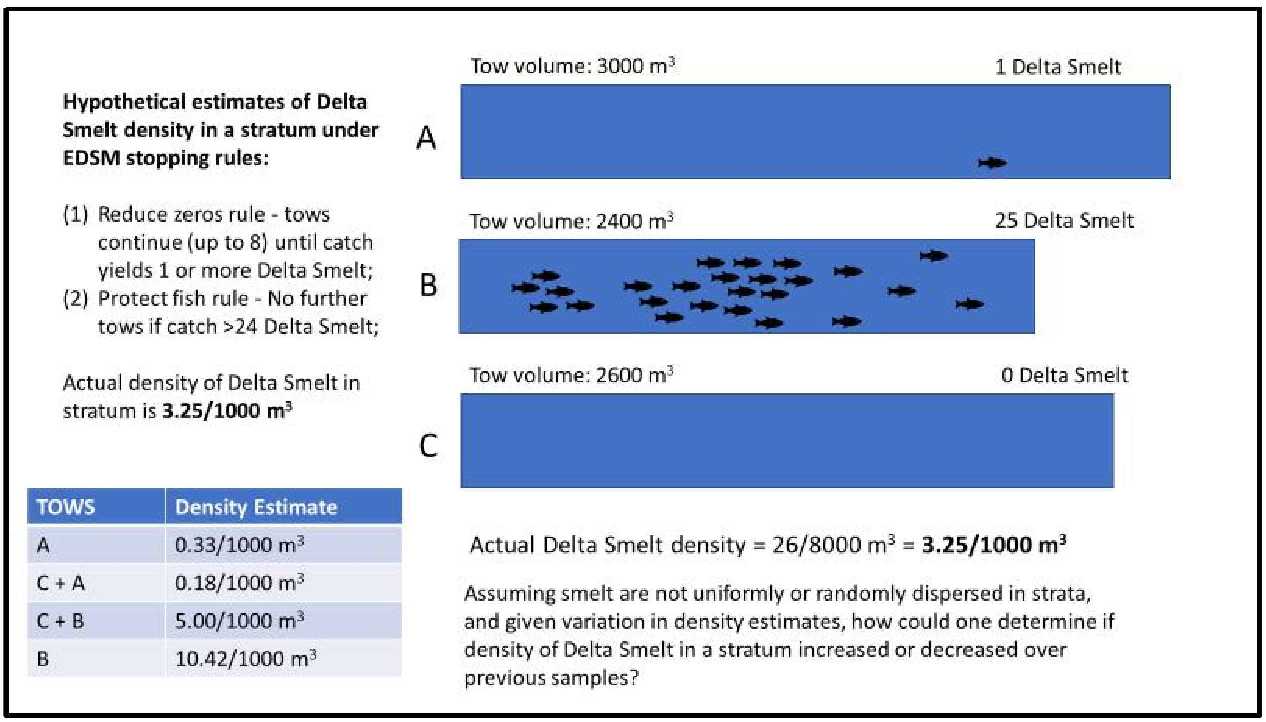


Figure 1. Illustration of the potential effect of initial fish location along hypothetical tow paths on likelihood of capture.

The stopping rules associated with the collection of adult and juvenile Delta Smelt, together with the patchy distribution and movement of fish within a sampling station, may also contribute significantly to the error and uncertainty associated with abundance estimates. It would be unusual to know the actual density of Delta Smelt at any given station and time, but Figure 2 illustrates a hypothetical situation in which a wide range of density estimates could be achieved at a station where actual densities are 3.5 fish/1000 m^3. For simplicity, this example uses a maximum of only three possible tow paths. This example uses the reasonable assumption that fish are not distributed evenly within the station. If the first tow happens to be Tow A, sampling ceases under the stopping rule and an abundance estimate that is only one tenth of the actual density is obtained. If the first tow is Tow C, the “reduce zero rule” requires sampling to continue and the abundance estimate is dependent on whether Tow A or Tow B is next. If Tow B, then an overestimate of actual density results, but if Tow A follows, density is underestimated by more than an order of magnitude. If Tow B happened to be the first at this station, the “protect fish rule” would be in effect, sampling would cease, and density would be overestimated by 300%. If the maximum of eight tows were allowed under the “reduce zeros rule”, and six of them contained no Delta Smelt, a considerably wider range of abundance estimates would be possible (on the order of 0.05 to 10.42 fish/1000 m^3). Given that the stopping rules could result in abundance estimates covering a range of over two orders of magnitude in this hypothetical exercise, it is difficult to see how the EDSM currently can be used to inform water operations in near real time, especially given the relatively few stations sampled each week. Perhaps the EDSM team should consider performing some synthetic sampling experiments, based on current abundance estimates, to study whether the increased take from relaxed stopping rules might significantly reduce estimation uncertainty and yet have negligible impact on total abundance.



TOWS	Density Estimate
A	0.33/1000 m^3
C + A	0.18/1000 m^3
C + B	5.00/1000 m^3
B	10.42/1000 m^3

Figure 2. An example of the range of abundance estimates from a station at which only 1-3 trawl tows are conducted under the “stopping rules” used to reduce zeros and protect fish when actual Delta Smelt density is 3.25 fish/1000 m^3.

The stopping rules cannot be applied to the larval Delta Smelt samples, and given that completely separate gear types with different unknown efficiencies are required to collect different life stages, it is currently unclear how the abundance estimates of larvae can be related to those of juveniles and adults.

There are potentially important sites within the designated sampling strata that are not being sampled for Delta Smelt because they are too shallow or contain too much organic material to be effectively sampled with trawls. For example, there are large shallow areas including Frank's Tract (2 m deep), Dutch Slough (2 m deep), Mildred Island (2.5 m deep) and the interior of Liberty Island (variable depth under 3 m). These areas account for large volumes and surface areas in the Delta. There are also important hydrodynamics that may influence the actual source of fish collected in different sampling strata. Table 2 includes some additional comments about each of the regions (strata).

Table 2. Notes on selection of Delta Smelt sampling strata in the Delta.

Region (Stratum)	Sampling considerations	References
Lower Sacramento[a]	Threemile Slough is a complex and critical junction with difficult hydrodynamics. This makes interpreting the significance of field sampling in this region difficult.	Monsen (2001)
Liberty Island/Cache Slough	Both exterior channels and interior of Liberty Island should be sampled. The BREACH I/II/III studies provide extensive scientific background for this region that could guide sampling.	BREACH III Studies. Lead Administrative PI: Charles Simenstad (University of Washington)
Sacramento Deep Water Ship Channel	The channel is approximately 10 m deep with no shallow water habitat. The channel supports ship traffic to the Port of Sacramento.	
Upper Sacramento	Always note whether the Delta Cross Channel is open or closed during sampling.	Gleichauf et al. 2015; Monsen et al. 2007

Region (Stratum)	Sampling considerations	References
Lower San Joaquin[b]	Sampling Dutch Slough is an important open water region in this stratum. Threemile Slough water exchange could also influence source water in this region.	
South Delta	Mildred Island is deep enough to sample. Frank's Tract is shallower. These two flooded islands have very different biological characteristics even though they are both South Delta open water habitats.	Lucas et al. 2002; Lopez et al. 2006
Suisun Bay	Always separate out Montezuma Slough/Suisun Marsh regions from the main Suisun Bay. These two regions have very different hydrodynamic characteristics.	
Suisun Bay: Montezuma Slough and Suisun Marsh	Observe what phase of tide at the time of the tow. Also note anything like observations of draining of the marsh areas into the channel.	
Suisun Bay: Main Suisun Bay	Sample in the open water habitats of both Grizzly Bay and Honker Bay. Tides are very energetic in this region.	
Mokelumne River/East Side Streams (North of the San Joaquin River;t East of the Delta Cross Channel)	This region should not be included with South Delta statistics. Always note in fieldwork whether Delta Cross Channel is open or closed as the connection to the Sacramento river drives water transport in Georgiana Slough and the Mokelumne channels.	Gleichauf et al. 2015

[a.]Lower Sacramento: Recognize that sampling in Threemile Slough represents San Joaquin sourced water half the time. Flow direction (Sacramento → San Joaquin or San Joaquin → Sacramento) is especially important to document in this channel. In Threemile Slough, the direction of the current is often different than the “flood” or “ebb” tide in the adjoining Sacramento and San Joaquin channels due to bathymetry and the connection between the two major rivers (Monsen 2001).

[b.] Lower San Joaquin: This is reasonable sampling. Sampling in the open water region called Dutch Slough is important. Also note that anything near Threemile Slough could be a Sacramento source water. It is very important to note times, direction of flow in Three Mile Slough (Sac → SJR or SJR→ SAC) and if it is ebb or flood on the San Joaquin river.

Results from the Old and Middle Rivers (OMR) Index Demonstration Project

Pumping at the State and Federal Projects (Projects) are generally constrained by the D-1641 Bay-Delta Standards by Delta outflow and export-to-inflow ratio (E:I) (DWR 2017), both of which are calculated by the DWR DAYFLOW program (http://www.water.ca.gov/dayflow/output/). However, neither Delta outflow nor the E:I ratio communicate flow conditions in the South Delta region. Therefore, during fish protection season, export facility operations must also incorporate flow restrictions based on a measured Old and Middle River (OMR) flow as outlined in the BiOps (DWR 2017).

While the charge questions asked the IRP to comment on the differences between the two methods to calculate “actual” OMR, a more relevant question is “How accurately does the Index Method predict the UGSG Gage estimates when OMR regulations are controlling South Delta pumping rates?”

Furthermore, the questions posed to the IRP on this topic seem to imply that the OMR Index and the USGS Gage Data Method were separate methodologies being equally evaluated for their success in representing “actual” OMR flow. In reality, the OMR Index values are being evaluated for their ability to accurately predict the UGGS Gage Data Method.

Use of the term “actual” OMR flow is misleading in the context of this demonstration project. “OMR flow” is a measure of direction and intensity of the flow at a tidally-averaged timescale incorporating both flows on Old River and Middle River north of the State and Federal export facilities. Although the California Data Exchange Center (cdec.water.ca.gov) has a station labeled OMR, the timeseries reported is the sum of the tidally-filtered flow measurements at the USGS stations on Old River (station OBI) and Middle River (station MDM). The “OMR “flow” was developed to communicate how strongly pumping at the State and Federal export facilities are influencing hydrodynamics in the South Delta region.

The hydrodynamics in the South Delta are tidal, with two floods and two ebbs daily. The magnitude of the tidally-averaged flow is approximately 10% of the maximum flow in either direction. It is not correct to interpret this region as a river system that is traveling northward towards the mainstem San Joaquin. Nor is it correct to assume that when the tidally-averaged flow is negative, that Old and Middle rivers are flowing “upstream” to the south towards the pumps. On a tidally averaged timescale, the residual flow is to the south, but in a Lagrangian framework, any “particle,” whether it is a drifting log, a plastic ball, or a larval/juvenile fish, may travel with the current many river miles up and downstream over a tidal cycle. To add to the complexity, the flows in the South Delta are also influenced by: 1) the filling and draining of the Delta, depending on the phase within the spring-neap tidal cycle, 2) atmospheric conditions, 3) the configuration of various temporary barriers in the channels during portions of the year, 4) the operation of the radial gates at the entrance of Clifton Court Forebay, and 5) the export rate at both the State and Federal pumping facilities.

According to the RPA action, the USGS Gage Data Method should be considered the “gold standard” definition of OMR. The critical time period of interest is during negative OMR flows, especially when the calculated OMR tidally averaged flow is between -2500 and -5000 cfs.

The OMR Index approach is being proposed as an alternative calculation method because the water operators have found that the USGS Gage Data Method specified in the RPA is difficult to use in actual operations. Table 3 compares the advantages and disadvantages of both approaches.

The correlation between the OMR Index and USGS Gage Data Method is an inadequate and inaccurate measure of how well the OMR Index predicts the values produced by the Gage Data Method. This is because any two variables with the same units can be perfectly correlated (r^2=1) and yet have major numeric differences. This will occur any time the two variables have an exact linear relationship with an intercept different from 0 and/or a slope different from 1.

The most accurate way to judge agreement between the two indices is to calculate the daily (or time-averaged) differences between them, (USGS Gage Data Method – OMR Index), and then analyze the patterns and statistics of those differences. Under which conditions do the two measures have the greatest residuals? What are the implications of those differences for fish and Project operations? These differences are the errors in

Table 3. Comparison of OMR Index Method (preferred by Reclamation) and USGS Gage Data Method (specified in RPA action).

ADVANTAGES	DISADVANTAGES
OMR Index Method	
● Vernalis flow at San Joaquin mainstem station is always unidirectional; tidal filtering unnecessary. ● Easy to extrapolate flow data when data gaps occur. ● No operational lag time for averaging required.	● Empirical relationship that includes, among other things, a very inaccurate estimate of in-Delta diversions and returns (DICU). ● Lack of analysis and understanding on the uncertainties and errors in the Index and their impacts on fish or operations.
USGS Gage Data Method	
● Data based on real-time monitoring in the tidal portion of the system. ● Incorporates effects of tides, filling and draining of the Delta, and storm surges.	● Less practical to use because of a 3-day delay to calculate the value due to tidal filtering requirements. ● Missing data because of issues with reliability of gages.[1] ● Method is still an index that involves combining information from stations in two different South Delta channels.

[1]Statements by Contra Costa Water District during the public comment period indicated that USGS has made these high priority stations since index started being used in management; data gaps are less frequent than in the past.

predictions of USGS Gage Data Method values by the OMR Index. Statistical descriptions of these differences, such as those in Table 6 of the OMR Demonstration Project report, should be the primary basis for assessing the accuracy of the OMR Index predictions.

Prediction accuracy is also difficult to assess graphically from overlaid time series data from the two methods, such as in Charts 1-21 of the OMR report. Numerically significant errors will tend to appear small in these overlaid time series, relative to the full range of variation in each variable. A clearer picture of the changing error magnitudes over time would emerge from a plot of the errors (USGS Gage Data Method - OMR Index) as a time series. The errors can also be plotted on the y-axis against other variables (e.g., USGS Gage Data Method values) on the x-axis, to better understand the sources of error.

Table 6 and scatterplots such as Chart 27 in the OMR report clearly show that there are systematic shifts in the prediction errors over time. In addition, Table 6 (OMR report) shows mean errors greater than 100 cfs, and up to nearly 500 cfs in some years, with maximum and minimum errors often exceeding 1000 cfs. Apparently, DWR has concluded that differences of these magnitudes are acceptable for the purpose of satisfying RPA action criteria for OMR flows, though no justification for this conclusion was provided. However, the IRP considers that differences of this size are of concern, especially in the context of the critical OMR range from -2500 to -5000 cfs, and encourages a more comprehensive and accurate assessment of the prediction errors and their possible sources. This assessment should include an ecologically and/or operational based threshold for error to establish that the OMR Index is adequately reproducing the USGS Gage Data.

Once differences between the two OMR methods are identified for critical periods, it would be beneficial to identify the sources of these differences. For instance, the OMR Index Method is a series of regression equations that are based on the configuration of the temporary barriers. Some of the differences could be related to inaccuracy of representing the temporary barrier configuration. There are periods of construction/deconstruction of these barriers when they are not fully operational. Alternatively, there may be regressions for certain barrier configurations that need to be improved.

It may also be possible to improve the OMR Index and reduce its prediction errors using other mathematical approaches. The IRP understands the DWR's reluctance to add complexity to the current OMR Index in the form of additional covariates. Such additions would require DWR to maintain a more complex model and to acquire and manage additional sources of model input data in real time. As an alternative, the IRP suggests an improved prediction algorithm that employs only the current value of the OMR Index and values from the USGS gages from prior days. Appendix 2 (current LOBO report) gives a full description of the proposed improved prediction model.

Draft Proposed Shasta RPA Amendment

Progress continues toward a better understanding of temperature variability in the Sacramento River and its potential consequences for juvenile Winter-run Chinook salmon. Model development based upon Martin et al. (2017), and research results thus far presented, hold considerable potential for resolving important links between the physico-chemical environment (e.g., temperature and oxygen levels) experienced by the earliest life stages of salmonids and their survival in the Sacramento River. The analysis and the integration of biophysical factors considered in the model, as well as its broad application under variable annual and interannual thermal conditions, should provide valuable guidance for temperature management targeting spawning habitat in the Sacramento River.

Application of the temperature-dependent egg mortality model and critical temperature threshold.

The Martin et al. (2017) model predicting temperature-dependent egg mortality is a parsimonious and realistic representation of temperature effects on eggs. Martin et al. (2017) found that the model gave poor predictions of observed field mortality, when its parameters were estimated from a relevant but restricted set laboratory data for egg thermal tolerances. Thus, they instead estimated model parameters from 18 years of observed field survival rates, in order to project the survival rates expected under future flow scenarios.

Despite its strengths, the IRP expects that model predictions of survival will have sizable uncertainty, resulting in wide confidence intervals (CIs) that may complicate managing for temperature-based mortality. The quality of fit of the field-parameterized model is illustrated by Figure 3 of Enclosure 3 of the draft Shasta RPA amendment. There is noticeable scatter in the relationship between observed and model-predicted survival. In addition, the sample size for fitting the model (n=18) is only marginally sufficient to estimate the two model parameters, T_{crit} and b. Such CIs were reportedly communicated to the flow management planners, along with the model's point predictions of egg survival. However, it was unclear if, and how, the uncertainty (CIs) for model-predicted survival - in addition to its point estimate - were being considered when planning for future flows.

Martin et al. (2017) propose a model of within-redd oxygen deprivation with the objective of explaining the marked discrepancy between field-measured egg survival rates and the rates predicted by the lab-parameterized mortality model. Although the oxygen deprivation is a likely hypothesis to explain this discrepancy between field observations and the model, further research is needed to eliminate other possible explanations. For example, it is possible that the assumption-laden and indirect estimates of field survival are biased. Another possibility is that temperatures within the redds are higher than the water-column temperatures predicted by the RAFT model.

This might occur due to direct radiant heating of redd gravel, coupled with greatly-reduced water movement within the gravel. Concern was expressed during public comment (J. Anderson) that the simplifying assumptions of the egg mortality model may overpredict mortality for a given seasonal temperature pattern, a point that is considered in the specific responses to charge questions below.

7-day average of the daily maximum temperatures (7DADM) vs. daily average temperature (DAT).

Managers expect that water operations could be more difficult under 7DADM due to the time lag inherent in its 7-day average. For example, daily water temperatures could have already turned to an upward trend, even as 7DADM is still falling, due to lag effects. Thus, water operations based on 7DADM compliance would require some forecasting. On the other hand, real-time operations based on DAT already have built-in lags of a few days because dam releases and downstream temperatures do not respond immediately to flow-change decisions. As a result, operations that try to maintain DAT compliance also require some forecasting.

Unless a workable management strategy under 7DADM can be achieved, the presumed biological advantages of 7DADM may not be realized. If DAT is retained as the compliance and operating metric, it is still useful to track 7DADM as an indicator of possible chronic effects on salmonid populations. As discussed more extensively below, a linearly weighted 7DADM would better represent the temperature variations in the channel on a weekly timescale for this system.

The IRP was unable to fully evaluate the tradeoffs of the two temperature criteria without further analysis. One primary concern was the need to demonstrate the ecological relevance and presumed biological benefits of considering the 7DADM over the DAT. Both the physiological costs to fish of having incomplete (lagged) temperature information as peaks or troughs occur in the system, as well as the actual likelihood of exposure to those temperatures that exceed a critical limit over a shorter period of time under the 7DADM, should be investigated. Such an analysis should overlap discussion on moving the temperature compliance point. Furthermore, questions remain about the impacts of the two temperature criteria on water supply, which directly affect the operational feasibility of the criteria. Would transitioning to 7DADM require more cold water, and at what times of year? Would that water be available during an average or dry year?

IRP RESPONSES TO SPECIFIC QUESTIONS FOR THE 2017 LOBO ANNUAL REVIEW

Responses of 2017 IRP to questions regarding Stanislaus River Watershed/Eastside Division RPA actions in WYs 2011-2017

1) **How well did the Stanislaus Operations Group (SOG) incorporate various considerations in its advice on the timing and shaping of the minimum required flow volumes for the fall pulse flow, the winter instability flows, and the spring pulse flow? What other factors (if any) should SOG include in its advice?**

Extensive monitoring is being conducted in the Stanislaus, and the integration of the field data (e.g., weir counts, water and air temperatures) with weather forecasts works to focus SOG operational recommendations on the timing and duration of pulse flows. The shaping of pulse flows seems to be based on logical current assumptions of salmonid responses to natural flow variability. For example, the transition to a three-peak pulse in October (prior to peak spawning) is intended to attract spawners while discouraging the construction of redds at streambed levels that will be dewatered later in the year. However, this qualitative logic model linking salmonid behavior to the design of flow pulses suffers from a lack of quantitative analysis to test the assumption that salmonids are responding to the pulse flows.

The primary objective of the fall pulse flows is to trigger migration up into the Stanislaus when operators are drawing down the reservoir. However, the data presented leads to some speculation about whether fall attraction flows, as currently practiced, stimulate or delay upstream migration or have an influence on encouraging straying of individuals into the basin. While it is clear that the SOG's intention in designing the fall pulses is to avoid dewatering of redds during winter base flows, it is unclear that one shape of pulse flow is better than another in actually triggering migration in the river. There is an understandable reluctance to experiment with the shape of pulses for fear of having unintended negative impacts on salmonids, but lacking objective data there is no way of knowing that current choices are any better or worse than alternatives.

A more refined justification for effects of the timing and shape of flows on salmonid migration (e.g., Sykes et al. 2009) could be developed and/or tested experimentally. Similarly, field observations could be examined to investigate the effectiveness of spring pulses in achieving their intended outcomes of reducing temperatures, inundating shallow habitats, and flushing smolts out of river and through the Delta. Such an analysis would be useful both for objectively shaping pulses and for prioritizing future locations of floodplain rehabilitation projects.

2) **Given the constraints in managing water temperatures in the Stanislaus River basin, how well did the SOG structure its approach and incorporate current science into temperature management?**

There are a number of important constraints on the effective management of temperature in the Stanislaus. One key constraint is the existing infrastructure. For example, New Melones Dam has limited temperature control capability due to low-level outlets that can only be used when reservoir depths are below 808 feet. The existence of the submerged relict structure of the original Melones Dam impedes the flow of cold water to the low-level outlets in New Melones. Furthermore, releases from New Melones flow into private reservoirs that limit the effectiveness of upstream temperature operations. Downriver, outlets at Tulloch Dam draw the coolest water available, but conflicting contract obligations constrain the availability of water from Goodwin Dam. As a result, temperature criteria for steelhead are often exceeded, in some cases for extended periods. If meeting temperature criteria is indeed a requirement of the RPA action, the SOG should consider how current infrastructure may be modified to provide the cold water needed to meet the criteria. The addition of a more flexible temperature control system at New Melones would likely be beneficial for meeting temperature targets with the least amount of water, though the capital costs are high and the infrastructure may be impractical to install or operate. If climate predictions over the next century are correct and snowmelt accounts for proportionally less of the water inflow to the Stanislaus, infrastructure improvements that allow better temperature control may become increasingly important to maintain temperature conditions to support viable populations of salmonids.

From a biological perspective, temperature criteria were established in the Stanislaus to benefit steelhead. However, the steelhead population in the Stanislaus may be too small to effectively evaluate the consequences of exceeding current temperature criteria. Furthermore, it is not clear if the current temperature criteria are protective of the remaining fall-run Chinook in the system. Although fall-run Chinook are not targeted by RPA actions, a summary of available data indicates that temperature requirements for Chinook and steelhead differ (Carter 2005) and managing temperature for one may have unintended negative consequences for the other.

3) **During extremely wet (i.e., water years 2011 and 2017) and extremely dry (i.e., water years 2012-2016) hydrologic conditions, SOG faced challenges in managing flows, reservoir storage, and temperature. How well was scientific information considered in the SOG's decision-making process under these extreme hydrologic conditions?**

The wide range of hydrologic conditions over the past seven years demonstrate the complex challenges of meeting targets for discharge and temperature under both wet and dry water years. The result is desired outcomes for people and ecosystems are decreasingly met in these extreme conditions. In WY2017, fall pulse flows were managed to prioritize flood risk reduction, with fish managers relying on luck to prevent redds established during elevated flows to remain inundated through fry emergence.

Unfortunately, forecasts of climate change indicate that normal water years are decreasingly likely to occur and that extreme conditions are likely to become more common. While California has been subject to extreme hydrologic conditions through its geologic history, the current hydrosystem, including the reservoirs, diversions, and relocated spawning habitats, are novel. Thus, while historical data can provide a guide for how fish might respond to extreme events in a natural river system, it provides limited insight into how best to protect fish in a highly manipulated system. Furthermore, the IRP was unable to find a focused attempt to summarize how experiences of the operators or the fish were synthesized in any way. Thus, it was unclear how well scientific information was integrated into the decision-making process during these past extreme events. These questions and issues are not unique to the Stanislaus River, and there are not currently any universally acceptable approaches to recovering species subject to hydrologic extremes. However, some key concepts can be applied in the future as a general framework for thinking about managing extreme events in modified systems. These concepts emphasize robustness (Herman et al. 2015), a condition where the species' sensitivity to the disturbances is reduced, and resiliency, where species are able to recover following a disturbance due to diversity of habitats and population life strategies (Walker et al. 2004).

First, a synthesis of lessons learned would be an important exercise for operators and managers. Such a synthesis should summarize the hydrologic year at a weekly to monthly time scale, outline operational decisions in response to data and associated releases, examine measures of how operations and extreme conditions impacted fish, and identify what additional information and/or alternative actions would be needed in a future event.

Second, there is a growing body of literature on managing novel ecosystems that is relevant to the Stanislaus. Even absent the effects of extreme conditions, the relocation of habitats downstream of the dams to formerly unsuitable locations, the loss of key species (e.g., Spring-run Chinook), and the introduction of invasive predators has already sufficiently modified the ecosystem structure and function to classify the Stanislaus as a novel ecosystem. The literature (Hobbs et al. 2006) argues to examining the persistence and value of the new ecosystem, and for managing these types of system in a way that is fundamentally different from current approaches. For example, Seastedt et al. (2008) suggest that, rather than attempting to recreate historical conditions and species compositions, management should emphasize maintaining genetic and species diversity, as well at the biogeochemistry that favor the desirable species. If the system is stable, what are the costs and risks of attempting to guide the system to a more desirable state? If the system is changing, what are the costs and benefits of maintaining current species compositions? Identifying where sites fall along a range of wild to intensively-modified will help managers prioritize conservation of the least impacted systems and identify for which systems maintaining or restoring historical species levels and compositions is not feasible.

The increasing frequency of extreme events raises the need to revisit the guiding vision for the basin. In more natural river systems subject to extreme events, critical refuges from droughts and floods are provided by side channels, floodplain habitats, deep pools, and other complex habitats that are largely missing from the current Stanislaus River. Thus, managers may need to reconsider what outcomes for the basin are feasible. Gravel augmentation and other habitat enhancements are likely to become increasingly important for species to survive extreme events, but there is currently no evidence that the agencies will be able to create enough of these habitats to produce a measurable benefit. The RPA action requiring the addition of 50,000 yd^3 of gravel by 2014, and 8,000 yd^3 annually thereafter, has been stalled by lack of funding, land access, and other issues. Other RPA actions lack measurable, time bound objectives (e.g., floodplain restoration and predation management) that link to biologically-relevant outcomes. While effects of RPA actions can be measured in a variety of ways (e.g., primary and secondary production, diversity, etc.), some evidence that the habitat projects meaningfully lead to recovery is needed. What area of new habitat is expected from 50,000 yd^3 of gravel, if that could be achieved? How many fish can that area support? Similarly, are small side channel projects at an adequate scale to produce a biological response, or are landscape-scale projects needed to provide adequate habitat for protecting fish during dry and/or wet years? Will predation mitigate much of the increased production in salmonids from the floodplain and gravel projects?

<u>Responses of 2017 IRP to questions regarding the Enhanced Delta Smelt Monitoring (EDSM) Program</u>

4) **How well is the EDSM program designed and carried out to provide usable results to inform the implementation of the 2008 Fish and Wildlife Service Biological Opinion?**

USFWS should proceed as rapidly as possible to evaluate whether EDSM can accomplish key objectives, especially the entrainment estimation goal. Because the current EDSM field program requires substantial money, material and person-hours, its continuance at the current level of effort may depend upon demonstrating promise for meeting as many of these objectives as possible. This was the first year of field sampling for the EDSM and the program continues to change (e.g., beginning with one sampling crew and expanding to three). This question might best be addressed after the sampling and abundance estimation approaches stabilize.

5) **How complete are the EDSM methods for providing improved understanding of abundance and/or distribution of Delta Smelt? What modifications could be considered to further improve understanding?**

Several issues that present challenges to improving the understanding of both abundance and distribution of Delta Smelt in the Delta are discussed in the IRP's general consideration of the EDSM program in this report – see "Enhanced Delta Smelt Monitoring (EDSM) Program." The section includes a discussion on limitations of currently used sampling gear, sampling volume calculations, inability to adequately sample shallow waters, potential "hot spots" of abundance, effects of tidal stage on the distribution of Delta Smelt within and between sampling strata, and the potential effect of stopping rules on estimates of abundance.

As is acknowledged in the EDSM report (Section 5.1), it remains unclear if EDSM can fulfill its seven stated objectives (Section 2 of the EDSM report) to inform the BiOps and water operations, particularly with regard to entrainment losses at the pumping facilities. In 2017, emphasis was on progress towards Objectives one (abundance estimation on a weekly time scale), two (spatial distribution), and five (comparison to existing surveys). The IRP encourages USFWS to push forward as rapidly as possible to evaluate if EDSM can also fulfil the other four objectives, especially the entrainment estimation goal. A reliable estimate of population size is also essential for determining a jeopardy level and allowable take.

Responses of 2017 IRP to questions regarding results of the Old and Middle Rivers (OMR) Index demonstration project

6) **How adequate are the two different OMR estimation methods (i.e., the OMR index equation method vs. the gage data method) for estimating actual OMR flows?**

The evaluation of the two method is inadequate. According to the 2009 RPA with 2011 requirements, and to most reliably represent hydrodynamics of the Delta, the USGS Gage Data Method should be considered the “gold standard” definition of OMR. The critical conditions of interest are negative OMR flows, especially when the calculated OMR tidally averaged flow is in the range -2500 to -5000 cfs. The OMR Index Method should be compared to the USGS Gage Data Method under these critical operating conditions to evaluate how well it represents the USGS Gage Data Method calculation.

7) **How complete is the evaluation of the two methods, including their effects on Central Valley Project/State Water Project operations in the Delta?**

The report provided a general evaluation, but did not focus on specific critical flow conditions such as those specified in the RPA action (-2500 cfs to -5000 cfs). The analysis should not only examine when the two indexes do not match, but also the underlying mechanisms causing the discrepancy. The OMR Index is a series of regression equations based on the configuration of the temporary barriers. It is important to evaluate the agreement of the model for all possible barrier configurations. Detailed critique of the evaluation, and recommendations for improvement, are provided in the general text on OMR flows above and in Appendix 2.

8) **When OMR index values differ from the gage-based OMR measurement, how well are these differences evaluated and understood?**

There was no apparent attempt to understand the reasons for the differences between the gage-based measurement and the index values. The OMR report briefly mentions that wetter years have smaller differences, but the explanation was unclear. The USGS Gage Data Method should be viewed as the “gold standard” to be attained, even though errors and omissions can occur with field sampling at the Old and Middle River flow stations due to equipment malfunctions. The OMR Index equations should be frequently reviewed and improved over time. It is very likely that there are certain barrier configurations where the OMR Index equation could be improved. There are statistical approaches that would improve Index-based predictions of the USGS Gage Data values (see Appendix 2 of this report). Greater effort directed toward understanding the prediction errors of the OMR Index, so that it can be applied more knowledgably, is strongly encouraged. If such understanding is not obtained and applied, then an algorithm such as Appendix 2 would be the only pathway to improved predictions of gage-based measurements.

<u>Responses of 2017 IRP to questions regarding the draft proposed Shasta RPA amendment</u>

<u>Temperature-dependent egg mortality model and critical temperature threshold</u>.

9) **How appropriate is the application of the temperature-dependent egg mortality model (including temporal and spatial scales; Martin et al. 2017) to understand early life history temperature-dependent mortality of Winter-run Chinook salmon and temperature management planning?**

In general, the application of the temperature-dependent egg mortality is a productive step forward in understanding one critical aspect of potential temperature-related mortality. In combining CE-QUAL-W2 modeling in the reservoir, the RAFT modeling in the river, and an understanding of conditions in the redds that may create survival problems for embryos and alevins, this approach represents a powerful predictive model for salmon vulnerability to temperature exposure. Two major concerns in its application are:

- The predictions of the oxygen diffusion model should be tested under field conditions because of the model's apparent sensitivity to extremely small changes in flow velocity. Eggs within a redd likely experience flow conditions that are spatially variable and temporally dynamic. The idea that a difference of <0.1 cm/sec (Martin et al. 2017) drives a 3-degree shift in critical temperatures would be strengthened by some empirical support for *in situ* redd conditions as well as oxygen depletion envelopes.

- The model depends on estimating background mortality that is additive with the temperature model. However, it may be problematic to apply such a density dependent model that lacks any mechanistic basis or site-specific information (see Assumptions, Q. 10).

Although a relatively minor concern, the comparison of temperature-dependent conditions in artificial and natural redds, using velocity as the hydraulic component of concern could have been enhanced by comparisons with other complex hydraulics which are known to create conditions that aerate redds. Gore et al. (2008) offer a compendium of hydraulic conditions that affect the physiology of aquatic organisms.

10) Is additional information needed to support the temperature-dependent egg mortality model assumptions, parameter estimates, and conclusions for Winter-run Chinook salmon?

The integration of biophysical factors that have been considered in the model, as well as its broad application under variable annual and interannual thermal conditions, is encouraging and represents a positive step toward quantitatively linking annual precipitation conditions, reservoir volume, temperature management, and biological outcomes in the river. However, more emphasis could be placed on parameter estimation, as well as sensitivity analyses, to further develop confidence intervals around the embryonic mortality that can be attributed to temperature. Survival, and hence mortality, are complicated metrics to estimate for juvenile salmonids, even with the benefit of many marked and recaptured individuals in a system (e.g., Hartson and Kennedy 2015, Myrvold and Kennedy 2016). Admittedly, a mark-recapture approach to estimating survival estimates cannot be easily applied to the early life stages and large spatial scales represented in this current work, but more research could be proposed to more confidently separate the temperature-dependent mortality that is of fundamental interest from other mechanisms of mortality in the system.

The components of survival across two very different life stages, egg and fry, would appear to be confounded. The research upon which the model is based (Martin et al. 2017), correctly assumes that thermal tolerance of fry is higher for fry than for embryos. However, the assumption that fry mortality is temperature-independent, at the experienced temperatures, is also simplistic and ignores bioenergetics constraints on juveniles near temperature thresholds (Myrvold and Kennedy 2015), as well as carryover effects of high temperature experienced throughout alevin and fry stages. If the Martin model is to be widely applied in the interpretation of egg survival, additional consideration of how uncertainties about fry and juvenile mortality affect the conclusions for thermal jeopardy of eggs may be required.

Based on the IRP's interpretation of the model, mortality is separated into two component models – an egg mortality model and a fry mortality equation. The egg mortality is based upon a probabilistic temperature-dependent egg survival model. This survival model is based upon laboratory data reported from two studies (Jensen and Groot 1991, USFWS 1999) that converge on 15.2 °C or 15.4 °C as a critical temperature for embryos. When applied to field temperature data from 1996 to 2015, these temperatures "failed to explain significant variation in the percent of embryos surviving through to the fry stage" (Martin et al. 2017). The workshop presentation (Danner, NMFS) and Martin et al. (2017) posed two hypotheses for why the lab model's predicted survival, estimated at ~20%, had so little predictive capability for observed survival, estimated at 5-50%. The alternate hypotheses are:

- Other factors may drive variation in annual survival; or
- Temperature tolerance under field conditions may be lower than under idealized laboratory conditions.

Little research has been conducted on the first alternative hypothesis. However, there is some modeling evidence supporting the second hypothesis based upon a theoretical, flow-based oxygen depletion of the *in situ* thermal tolerance. By dropping the critical temperature more than 3 °C to 12.0 °C, the model captured more of the observed variation in survival, including the outlier years of unusually high temperatures (and low observed survival at 5%) in 2014 and 2015.

The model also incorporates a temperature-independent background survival probability from egg to fry stage that essentially uses a density dependent fry survival model based upon the Beverton-Holt relationship. This attempts to estimate background mortality based largely upon the number of breeders (redds and carcasses) that are estimated annually and is imposed *after* the temperature dependent mortality at the egg stage.

The IRP discussed some concerns regarding the general approach to quantifying *observed mortality* in the system over the period of survival modeling (Martin et al. 2017). Generally, the temperature-dependence in the Martin model assumes that other sources of mortality in the system are understood. This is a difficult assumption to accept given all of the other attributes of this novel ecosystem. In short, the survival estimate is based on the comparison of final survivors, or the count of out-migrating fry, derived from rotary screw trap sub-sampling compared to counts of redds and other evidence of spawning in the system more than 15 miles upstream. For example, if in a model system that begins with 10,000 eggs, equally distributed within four separate redds (2,500 eggs per redd), and if the temperature dependent model in year X predicts that *half* of those survived, 5,000 alevins should result. Next, based upon the Beverton-Holt relationship, there is an expected 40% additional mortality due to fry survival. This results in an expected catch (from an efficiency-corrected rotary screw trap estimate) of 2000 juvenile fish in the sampled population at Balls Ferry. Each of these steps includes parameter estimation with confidence intervals that are not fully considered or presented in the survival numbers. These parameter estimations are also not treated in a sensitivity analysis in order to quantify how uncertainty in true spawner density, egg abundance per redd or capture efficiency of fry affects final survival estimates.

Data are not presented to support density dependent effects at the current population levels through either of the hypothesized mechanisms of superimposition or competition. If superimposition is the primary mechanism of density dependent mortality, the spatially explicit redd information should be able to inform the extent of that relationship. In addition, the background survival parameter should be considered prior to, at least in part, the application of the temperature-dependent survival. This is not how it is currently applied, and the background mortality is considered independently, and in addition to (*i.e.,* after), the thermally-induced embryo mortality. Alternatively, if fry mortality is the major mechanism, then it would make sense to apply it to post emergence abundance estimates (after temperature-dependent embryo mortality). In such a scenario, the predicted relationship that uses the Beverton-Holt relationship should not be based on the number of redds, or spawners, but on the surviving densities of the egg mortality model. As it has been presented, densities for the carrying capacity model are based upon densities of females or redds, but to the extent that the Beverton-Holt relationship represents survival for alevins and fry, it should consider the mortality imposed at the egg stage and its relevance for competition at later life stages. However, in this case, as Martin et al. (2017) states, "observation suggests that most fry only spend a few days between emergence and Red Bluff Diversion Dam passage." It is unlikely that the density dependence captured by the Beverton-Holt relationship makes sense mechanistically over such short time periods, broad spatial areas, and overall low densities. The model could be improved with an increased understanding of the extant mechanisms and timing of density-dependent survival in the system. In addition, more research designed to test the density dependence relationship used in the background mortality estimates of the Martin model would be useful.

Some assumptions in the model that should be more carefully considered, as well as some opportunities for further research include:

- Separating pre-hatch and post hatch survival. Future research could address this, but it could be a significant undertaking that introduces more uncertainty into the model. However, as it stands, the results and their application hinge upon the confidence with which embryonic mortality is assigned, particularly when it is recognized that survival from redd locations (embryos) to Balls Ferry (fry) is likely complicated.

- Much confidence is placed on the starting size of the embryo population. This is apparently based on both redd surveys and carcass surveys, but clearly not every redd is necessarily viable. It is widely recognized that females sometimes dig false redds and, conceivable, this occurs more under some environmental conditions than others. The Beverton-Holt term seeks to deal with this in a density dependent matter, but there may be other density independent concerns about turning redd data into embryo abundance.

- A conceptual diagram presented during the review, (E. Danner, NMFS, Slide 23,) nicely summarized how survival is estimated as a daily time step through the residence of the redd. However, it is unclear how, or if, variable temperatures experienced by fry are incorporated into a single survival number for those days that exceed the critical temperature (T_{crit}). A public opinion presentation at the review questioned whether egg mortality risk remains constant across the incubation period or should be treated as a dynamic property based upon the duration of egg incubation. Questions were raised with the egg mortality model about whether metabolic demands just after fertilization were as high as those 20-40 days post fertilization. This is not the general case wherein, during development, the embryo requires increasing O_2 as egg mass is converted into tissue. Given the high degree of precision that the model seeks to capture in terms of the temperature-flow-mortality relationships, it would seem appropriate to consider that temperature thresholds (and the associated critical O_2 levels) immediately following fertilization are lower than that required several weeks later. As is presently considered, the model might overestimate the duration or extent over which eggs are subject to thermal stress. Incorporating this level of specificity into an already complicated and assumption-laden model would be nice, but the practicality of temporally modifying system-level temperature targets over a potentially extended period of fertilization dates would be too cumbersome. Managing for the peak of spawning under a fixed temperature criterion (as appears to be the case) might be the most logical choice, but considering variation in spawn timing and stage dependent embryo requirements would also be an important refinement to the model.

In proposed future applications of this model (e.g., draft proposed 2017 Shasta RPA amendment document, p. 21), meeting temperature-dependent mortality targets defines the amount of reservoir storage required. However, this mortality is not solely dependent on the temperature of reservoir releases into the Sacramento River. It is also affected by the spatial distribution (and abundance) of redds (temperature-independent mortality). So, it is not entirely clear how future population dynamics or dispersion would or could affect operations. At the very least, these spatial considerations are critical for establishing temperature compliance locations. For example, there was a broader than expected dispersion of spawning fish approaching 10 miles downstream in the warmer years of 2014 and 2105 (E. Danner, NMFS, Slide 40) thus exposing these particular redds, a subset of the entire modeled population, to higher mortality probability. Within any given year, this means that a truly responsive water management plan should consider where and how many redds there are in the system in addition to the projected temperature dependent mortality. While this may or may not represent a meaningful or feasible consideration within any given year, it is increasingly appreciated that salmonid populations can exhibit rapid evolutionary responses to environmental change (Hegg et al. 2013, Waples et al. 2017). Persistently managing temperature in the system in a way that minimizes the opportunity for the selection of traits that may confer resilience in a population, may itself be counterproductive. Thus, the consideration of a water release schedule that can respond to changes in distribution and abundance may be desirable in years where population levels and cold-water storage can accommodate it.

11) How appropriate is the application of the Australian model (Mount et al. 2016) in defining objectives for Winter-run Chinook salmon and temperature management planning (e.g., water year type, Shasta Reservoir storages, and temperature-dependent mortality of Winter-run Chinook salmon)? What additional insight does this model provide for defining objectives?

Mount et al. (2016) provide a conceptual model for management of drought-prone regions like some of those in California and the Murray-Darling system in Victoria, Australia. The Murray-Darling is the largest river system in Australia and is greatly influenced by the ENSO (El Nino Southern Oscillation) (Allan et al. 1996). Rainfall variation in the Murray-Darling is estimated by the Southern Oscillation Index (SOI), a measure of air pressure difference between Tahiti and Darwin, Australia. These variations are phase-locked into the annual cycle (Nicholls 1991). Thus, management of the Murray-Darling can be anticipated on a more predictable cycle than for management of California rivers. If such predictions were available for the Sacramento-San Joaquin system, it might be possible to refine predictions of water availability into the future.

Thus, the scientific appropriateness of the model is on somewhat shaky footing. Mount and colleagues warn that there are several considerations to their proposed protocol:

- Duration of Drought - Currently, drought in California is relatively short-term. However, climate change over the coming decades may make California systems much more like Australia.

- Differences in Snowpack Storage - little or no snowpack storage drives Victorian systems while California systems are largely dependent upon snowpack for delivery of the cold water that is used for temperature management. If, as mentioned earlier, snowpack coverage is declining, the ability to regulate cold-water flows is also in jeopardy.

- Role of groundwater - Australian systems depend upon a well-driven system and groundwater recharge.

- Targets of conservation - While California tends to focus upon recovery or maintenance of salmonid populations, Australian conservation efforts are aimed at restoration of ecosystem integrity.

- Different species of concern – Australian species of concern focus primarily upon endemic and endangered species, most of which are considered to be forage fish rather than top carnivores. Indeed, much of the management of the Murray-Darling is focused upon the reduction or elimination of introduced exotics: Rainbow Trout (*O. mykiss*), Brown Trout (*S. trutta*), Atlantic Salmon (*S. salar*), and Brook Trout (*S. fontinalis*). Thus, some management targets in California might be substantially different from those in Victoria.

It would be useful to consult Young (2001) to understand the differences between the two ecosystems.

One thing that Mount and colleagues do suggest that is important to both agencies is the essential nature of **planning for drought rather than reacting to drought**.

The Mount et al. (2016) system may be useful, but the appropriateness of the model can and should be tested by establishing an operations-decision-tree using Mount's guidance. It would then be relatively easy to go back to each year over the past decade and evaluate at what junctures in operations Mount and colleagues' concepts would have led to a different operational decision. If those alternate (Mount) decisions would have led to a better outcome (e.g., increased reliability of meeting temperature targets), then it could be assumed that the Mount et al. system is preferable to current decision/options.

12) How appropriate are the interim temperature-dependent mortality objectives for informing the development of life-stage specific survival objectives? What, if any, guidance should be incorporated to refine the objectives?

The temperature dependent mortality limitations are an adequate first step in establishing interim criteria for reservoir management. The application of exceedance levels based upon water conditions is a logical first step and linking it with the conceptual model of Mount et al. (2016) is a useful extension of this framework. However, it was unclear what data, models, or logic were used to establish the exceedance levels (30% down to 3%). From the historical temperature profiles (1990 and through to 2015), it appears that in six years early life stages would face zero mortality while in at least seven years they would suffer mortality up to 70%. However, total survival, the “product of temperature dependent survival and background survival,” would appear to average across the survival landscape that is consistently between 20% and 30% (i.e., mortality > 70%) across years 1996-99 and 2002-2014 (NMFS, Appendix 4, 2017). Much of this mortality is for a population that, historically, was unlikely to have spawned in this section of the Sacramento River, or at these temperatures. Before establishing guidelines for a model that is focused on temperature-dependent embryo mortality, there is a need to better understand how other sources of mortality (NMFS, Appendix 4, 2017) affect the number of individuals expected to make it to the Delta.

Thus, the approach is limited by not separating temperature dependent mortality from other sources of mortality, as well as by other sources of error in applying Martin et al.’s (2017) temperature mortality model. Furthermore, the identified thresholds of temperature-dependent mortality need to be biologically justified, and managers should confirm that other important life-cycle stages that require temperature consideration are not neglected. Finally, it was not clear how compliance (based upon survival) can be accurately measured, how uncertainty would be incorporated or what the consequences of exceeding mortality thresholds would be.

7-Day Average of the Daily Maximum temperatures (7DADM) vs. Daily Average Temperature (DAT).

13) How appropriate is the use of 7DADM for the temporal and spatial scale involved in temperature management implementation?

The purpose of averaging any timeseries dataset is to smooth out short term fluctuations while maintaining the general trend in the data for a specified period of interest. For this proposed application of analysis of temperature trends downstream of Shasta, the IRP identified:

a) limitations to the 7DADM approach,
b) additional averaging approaches that should be considered,
c) the need for evaluating the operational feasibility of the criteria, and

d) the need for identify biologically-relevant criteria for the types of temperature fluctuations that typically occur in the system.

The 7DADM proposed by NOAA Fisheries is an arithmetic moving average, which assigns equal weight to all values in the timeseries. This approach is often used to filter out fluctuations in data centered around the mean. However, for datasets that are not centered on the mean, the moving average will create a lag in the data that can bias the average by the previous data point. This type of averaging lag is evident in Slide 7 of the workshop presentation (Sawyer, NOAA, Temperature Metrics: DAT and 7DADM).

Alternate averaging approaches should be considered in addition to 7DADM. For example, the weighted moving average is a simple statistical approach in which the most recent data are given the most weight while each previous point has a linearly decreased weight.

7DADM Linear Weight=
(T(7)*7+T(6)*6+T(5)*5+T(4)*4+T(3)*3+T(2)*2+T(1)*1)/(7+6+5+4+3+2+1)

This approach represents the fluctuating temperatures downstream of Shasta Dam better than an arithmetic mean average (Figure 3). This approach is consistent with the length of averaging specified in the U.S. EPA (2003) report, but is a better averaging approach in a river system where temperatures fluctuate at a timescale of less than a week. In addition, managers may consider if averaging over a shorter period (e.g., 3DADM, 4DADM) would be more protective of fisheries, as well as the impacts on cold water resources over time.

Multiple averaging schemes that could be used to specify an average daily maximum at Jelly's Ferry are illustrated in Figure 3. The proposed 7DADM significantly lags the observed data. However, both a 3-day and 4-day average daily maximum both follow the sharp rise in observed temperature with less lag at the temperature peak. The simple linear weighting of the 7DADM improved the lag in temperature, while still keeping the peak maximum temperature computed with the 7DADM.

Furthermore, as previously noted, the operational feasibility of the two criteria need further evaluation. For example, it may not be appropriate to adopt 7DADM for temperature compliance until water operators can devise a feasible strategy for real-time system operations under 7DADM that would maintain temperature compliance at least as well as the current DAT-based strategy. A starting point for this task might be to clearly articulate the current DAT-based strategy, including its forecasting and lagged system-response features. Such an analysis would likely require use of hydrodynamic and temperature models, which already exist for the basin. Analysis should then be conducted to modify the current strategy to represent operations with 7DADM. The proposed 7DADM strategy could then be applied to a time series of historical flows to evaluate its success relative to what was actually achieved under the DAT strategy.

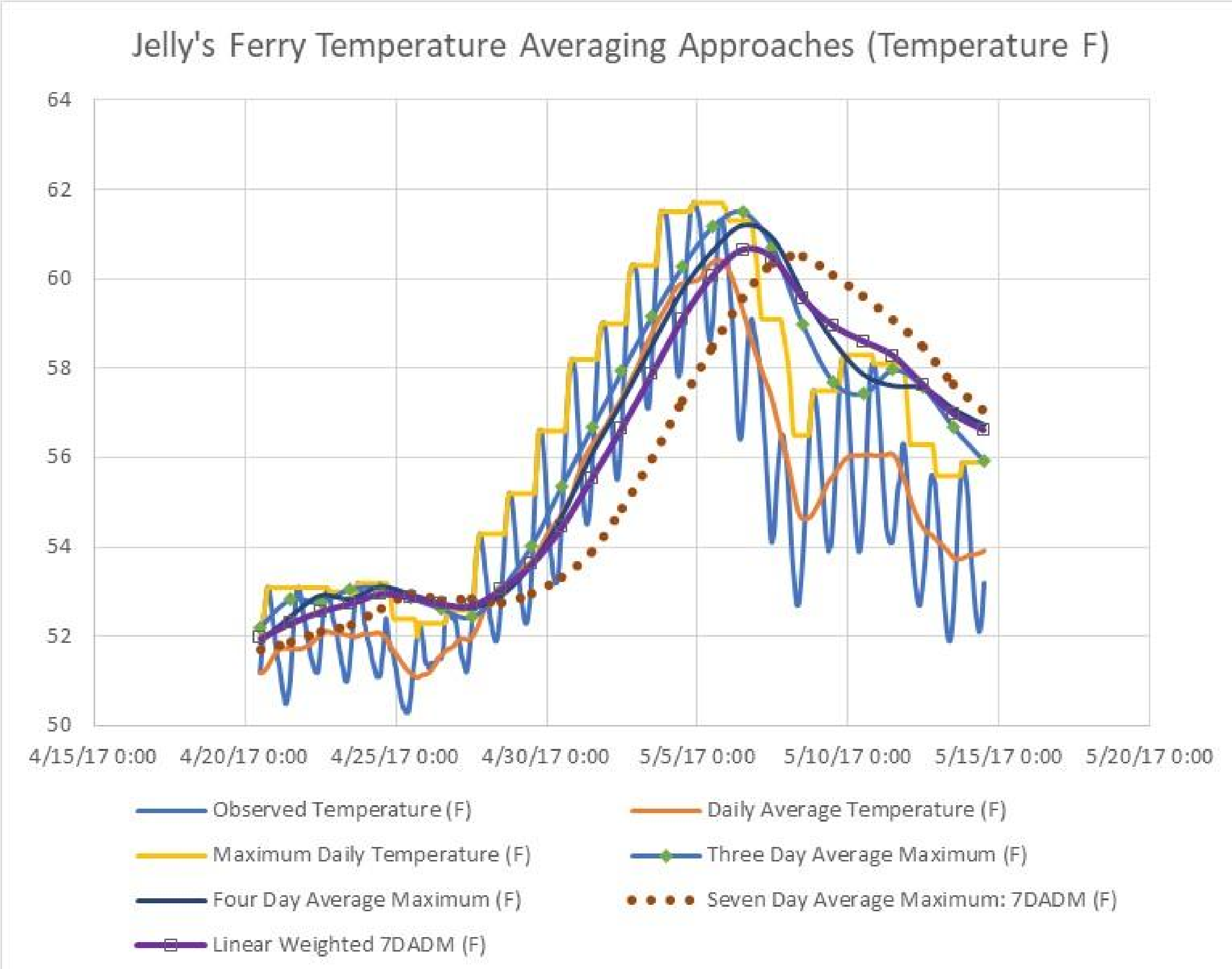


Figure 3. Comparison of averaging approaches that could be used to specify daily average temperature (DAT) at Jelly's Ferry.

This comparison should be conducted during wet, average, and dry years to examine the timing and volumes of water demands for meeting the two criteria.

Finally, evaluation of the two temperature targets requires some demonstration of the ecological relevance and presumed biological benefits of 7DADM and DAT. The justification for applying 7DADM was based on an U.S. EPA (2003) report that was not provided to the IRP, the results of which are derived from data in another physiographic region. A demonstrated ecological relevance is needed for any temperature target, where relevance should be defined by the physiological costs to fish of the combined magnitude and duration of peak temperatures.

14) How well is best available science incorporated into the draft proposed amendment's rationale for the change in temperature metric (from DAT to 7DADM) and location (farthest downstream, to the California Data Exchange Center CCR gage location [as a surrogate for the most downstream redd location])?

The rationale in the draft proposed amendment to change the temperature metric and move the temperature compliance point upriver to the CCR gage location appears to rely on a combination of sources including:

(a) the U.S. EPA (2003) report recommending temperatures that should support survival of Chinook Salmon eggs and alevins;

(b) the model described in Martin et al. (2017) which predicted that slower flowing waters at elevated temperatures would not provide a sufficient oxygen supply for embryo survival in redds; and

(c) recommendations of the 2014 and 2015 LOBO panels, which offered an expert opinion and recommendation for conserving cold-water resources by moving the temperature compliance point upstream so as to maintain cold water flows only to the downstream river reaches actually used by Chinook Salmon for spawning.

It is the 2017 IRP's understanding that assessing the quality of data in the U.S. EPA (2003) report was not within the scope of our charge. Since this literature source appears to be the sole basis for the agencies' consideration of 7DADM, the IRP has little ability to determine if choices made by the agencies are using the best available science.

The recommendations of the previous LOBO panels were a matter of logic and expert opinion. The reasoning was that temperature and flow volumes, which cleaned fine sediments from redds, interacted with dissolved oxygen to create microhabitat conditions that would support survival of eggs and embryos. It was also clear from experience in previously dry years that limited cold-water resources were insufficient to meet temperature criteria set for river reaches that were downstream of the areas used by salmon for spawning. Rather than raising the temperature compliance target (an option considered by the agencies), the LOBO IRP recommended moving the compliance point upriver where sufficiently cold temperatures could be maintained over all, or most of, the redds yet still conserve some cold-water storage for use later in the year to improve conditions for other life stages. The Martin et al. (2017) paper provides some additional support for the reasoning and expert opinions of the previous LOBO panels. However, even taken collectively, the information relied upon for moving the temperature compliance point upriver falls short of being conclusive evidence supporting the use of best available science as a foundation for the decision.

REFERENCES

Allan, R. J., G. S. Beard, A. Close, A. L. Herczeg, P. D. Jones, and H. J. Simpson. 1996. Mean sea level pressure indices of the El Niño-Southern Oscillation: relevance to stream discharge in southeastern Australia. *CSIRO Division of Water Resources, Divisional Report* 96/1.

Bennett, W. A. and J. R. Burau. 2014. Riders on the storm: Selective tidal movements facilitate the spawning migration of threatened Delta Smelt in the San Francisco Estuary. *Estuaries and Coasts*: 1-10.

Bozic, S. M. 1979. Digital and Kalman filtering. Edward Arnold, London, U.K.

Brown, C., Y. Ghile, M. Laverty and K. Li. 2012. Decision scaling: Linking bottom-up vulnerability analysis with climate projections in the water sector, *Water Resources Research* 48, W09537, doi:10.1029/2011WR011212.

Calow, P. (Editor). 1998. The encyclopedia of ecology and environmental management. Blackwell Science Ltd., Oxford (see pp. 329-330).

Cayan, D. R., E. P. Maurer, M. D. Dettinger, M. Tyree and K. Hayhoe. 2008. Climate change scenarios for the California region. *Climatic Change* 87 (Supplement 1): S21-S42.

Cloern, J. E., A. Robinson, A. Richey, L. Grenier, R. Grossinger, K. E. Boyer, J. Burau, E. A. Canuel, J. F. DeGeorge, J. Z. Drexler, C. Enright, E.R. Howe, R. Kneib, A. Mueller-Solger, R. J. Naiman, J. L. Pinckney, S. M. Safran, D. Schoellhamer, C. Simenstad. 2016. Primary production in the Delta: then and now. *San Francisco Estuary & Watershed Science* 14(3) Article 1:1-7. http://dx.doi.org/10.15447/sfew.2016v14iss3art1

Cunningham, R. B. and D. B. Lindenmayer. 2005. Modeling count data of rare species: some statistical issues. *Ecology* 86: 1135-1142.

Easterling, D. R., K. E. Kunkel, J. R. Arnold, T. Knutson, A. N. LeGrande, L. R. Leung, R. S. Vose, D. E. Waliser and M. F. Wehner. 2017. Precipitation change in the United States, pp. 207-230. In: Wuebbles, D. J., D. W. Fahey, K. A. Hibbard, D. J. Dokken, B. C. Stewart and T. K. Maycock (eds.). Climate Science Special Report: Fourth National Climate Assessment, Volume 1, U.S. Global Change Research Program, Washington, D.C., USA. doi: 10.7930/J0H993CC.

Franks, S. E. and R. T. Lackey. 2015. Forecasting the most likely status of wild salmon in the California Central Valley in 2100. *San Francisco Estuary & Watershed Science* 13(1) Article 1:1-11. http://dx.doi.org/10.15447/sfew.2015v13iss1art1.

Gleichauf, K., P. J. Wolfram, N. E. Monsen, O. B. Fringer, S. G. Monismith. 2015. Dispersion Mechanisms of a Tidal River Junction in the Sacramento-San Joaquin, Delta California. *San Francisco Estuary & Watershed Science* 12(4) Article 1:1-23. http://escholarship.org/uc/item/6js9z7bc

Gore, J. A 1982. Benthic invertebrate colonization: source distance effects on community composition. *Hydrobiologia* 94: 183-194.

Gore, J. A., J. Mead, T. Penczak, L. Higler and J. Kemp. 2008. Processes Influencing Aquatic Fauna. Pp. 62- 87. *In:* D. Harper, M. Zalewski, and N. Pacini (eds.) *Ecohydrology: Processes, Models and Case Studies. An approach to the sustainable management of water resources.* CABI Publ., London.

Gore, J. A., and A. M. Milner. 1990. Island biogeographic theory: can it be used to predict lotic recovery rates? *Environmental Management* 14: 737-753.

Hartson, R. B. and B. P. Kennedy. 2015. Competitive release modifies the impacts of hydrologic alteration for a partially migratory stream predator. *Ecology of Freshwater Fish* 24(2): 276-292.

Hegg, J., B. P. Kennedy, P. Chittaro and R. Zabel. 2013. Spatial structuring of an evolving life-history strategy under altered environmental conditions. *Oecologia* 172:1017-1029.

Herman, J. D. Reed, P. M., Zeff, H. B. and G. Characklis. 2015. How should robustness be defined for water systems planning under change? *Journal of Water Resources Planning and Management* 141(10): 1-14.

Hobbs, R. J., S. Arico, J. Aronson, J. S. Baron, P. Bridgewater, V. A. Cramer, P. R. Epstein, J. J. Ewel, C. A. Klink, A. E. Lugo, D. Norton, D. Ojima, D. M. Richardson, E. W. Sanderson, F. Valladares, M. Vilà, R. Zamora and M. Zobel. 2006. Novel ecosystems: theoretical and management aspects of the new ecological world order. *Global Ecology and Biogeography* 15(1):1-7.

Hughes, D. L., J. R. Olson, M. P. Brossett and J. A. Gore. 2010. Development of ecoregional and subecoregional reference conditions. Pp. 65-8n *in:* D.L. Hughes, M.P. Brossett, J.A. Gore, and J.R. Olson (eds.) *Rapid Bioassessment of Stream Health* CRC Press, Boca Raton.

Hylen, A., K. Korsbrekke, O. Nakken and E. Ona. 1995. Comparison of the capture efficiency of 0-group fish in pelagic trawls. pp. 145-156, In: Hylen, A. (ed.). *Proceedings of the Sixth IMR-PINRO Symposium*, Institute of Marine Research, Bergen, Norway.

Jensen , J. O. T. and E. P. Groot. 1991. The effect of moist air incubation conditions and temperature on Chinook salmon egg survival. *Fisheries Bioengineering Symposium Proceedings – American Fisheries Society* 10:529-538.

Kelly, M. H. and J. A. Gore. 2008. Florida river flow patterns and the Atlantic Multidecadal Oscillation. *River Research and Applications* 24: 598-616.

Lopez, C. B., J. E. Cloern, T. S. Schraga, A. J. Little, L. V. Lucas, J. K. Thompson and J. R. Burau. 2006. Ecological Values of Shallow-Water Habitats: Implications for the Restoration of Disturbed Ecosystems. *Ecosystems* 9(3): 422-440.

Lucas, L. V., J. E. Cloern, J. K Thompson and N. E. Monsen. 2002. Function Variability of Habitats Within the Sacramento-San Joaquin Delta: Restoration Implications. *Ecological Applications 12(5)*: 1528-1547.

Mantua, N. J., S. R. Hare, Y. Zhang, J. M. Wallace, and R. C. Francis. 1997. A Pacific interdecadal climate oscillation with impacts on salmon production. *Bulletin of the American Meteorological Society* 78: 1069-1079.

Martin, B. T., A. Pike, S. N. John, N. Hamda, J. Roberts, S. T. Lindley, and E. M. Danner. 2017. Phenomenological vs. biophysical models of thermal stress in aquatic eggs. *Ecology Letters* 20(1): 50-59.

McKenzie, D. I., J. D. Nichols, J. A. Royle, K. Pollack, L. Bailey and J. Hines. 2005. Occupancy estimation and modeling. Academic Press, Elsevier Inc., Oxford, U.K.

Monsen, N. E. 2001. A Study of Sub-Tidal Transport in Suisun Bay and the Savramento0San Joaquin Delta, California. PhD Thesis. Stanford University.

Monsen, N. E., J. E. Cloern, and J. R. Burau. 2007. Effects of Flow Diversions on Water and Habitat Quality: Examples from California's Highly Manipulated Sacramento-San Joaquin Delta. *San Francisco Estuary & Watershed Science* 5(3) Article 2. http://repositories.cdlib.org/jmie/sfews/vol5/iss3/art2

Mount, J., B. Gray, C. Chappelle, J. Doolan, T. Grantham, and N. Seavy. 2016. Managing water for the environment during drought: Lessons from Victoria, Australia. Public Policy Institute of California, San Francisco, CA. June 2016.

Munson, A. B. and J. J. Delfino. 2007. Minimum wet-season flows and levels in southwest Florida rivers. *J. American. Water Resources Association* 43: 522-532.

Munson, A. B., J. J. Delfino and D. A. Leeper. 2005. Determining minimum flows and levels: the Florida experience. *J. Amer. Water Resources Assoc.* 41: 1-10.

Myrvold, K. M. and B. P. Kennedy. 2015. Variation in juvenile steelhead densities in relation to instream habitat and water characteristics. Transactions of the American Fisheries Society 144: 577-590.

Myrvold, K. M. and B. P. Kennedy. 2016. Juvenile steelhead movements in relation to stream habitat, population density and body size: consequences on individual growth rates. *Canadian Journal of Fisheries and Aquatic Sciences 73(10): 1520-1529.*

Nicholls, N. 1991. The El Niño Southern Oscillation and Australian vegetation, pp. 23-36, *In*: Henderson-Sellers, A and A. J. Pittman (eds.), Vegetation and climate interactions in semi-arid regions. *Advances in Vegetation Science* 12. Springer, Netherlands.

Seastedt, T. R., R. J. Hobbs and K. N. Suding. 2008. Management of novel ecosystems: are novel approaches required? *Frontiers in Ecology and the Environment* 6(10):547-553.

State Water Resources Control Board. 2016. Substitute Environmental Document in Support of Potential Changes to the Water Quality Control Plan for the San Francisco Bay-Sacramento/San Joaquin Delta Estuary: San Joaquin River Flows and Southern Delta Water Quality. Recirculated Draft (SCH#2012122071). September. (ICF 00427.11.) Sacramento, CA.

Stenseth, N. C., G. Ottersen, J. W. Hurrell, A. Mysterud, M. Lima, K.-S. Chan, N. G. Yoccoz, and B. Adlandsvik. 2003. Studying climate effects on ecology through the use of climate indices: the North Atlantic Oscillation, El Niño Southern Oscillation and beyond. *Proceedings of the Royal Society of London.,* B 270: 2087-2096.

Swanson, C., P. S. Young and J. J. Cech Jr. 1998. Swimming performance of delta smelt: maximum performance, and behavioral and kinematic limitations of swimming at submaximal velocities. *Journal of Experimental Biology* 201: 333-345.

Sykes, G. E., C. J. Johnson and J. M. Shrimpton. 2009. Temperature and flow effects on migration timing of Chinook Salmon smolts. *Transactions of the American Fisheries Society* 138:1252-1265.

Walker, B., C. S. Holling, S. R. Carpenter and A. Kinzig. 2004. Resilience, adaptability and transformability in social–ecological systems. *Ecology and Society* 9(2): 5.

Wilby, R. L. 2016. Managing rivers in a changing climate. Pp. 353-371 *in:* D. J. Gilvear, M. T. Greenwood, M. C. Thoms, and P. J. Wood (eds.) *River Science. Research and Management for the 21st Century.* Wiley Blackwell, West Sussex.

U.S. Environmental Protection Agency. 2003. EPA Region 10 Guidance for Pacific Northwest State and Tribal Temperature Water Quality Standards. EPA 910-B-03-002. Region 10 Office of Water, Seattle, WA. USFWS. 1999. Effect of temperature on early-life survival of Sacramento River Fall- and Winter-run Chinook salmon. Red Bluff, CA: Northern Central Valley Fish and Wildlife Office.

Young, W. D. (ed). 2001. *Rivers as Ecological Systems. The Murray-Darling Basin.* Murray Darling Basin Commission / CSIRO Land and Water, Canberra.

Waples, R. S., A. Elz, B. D. Arnsberg, J. R. Faulkner, J. J. Hard, E. Timmins-Schiffman and L. K. Park. 2017. Human-mediated evolution in a threatened species? Juvenile life-history changes in Snake River salmon. *Evolutionary Applications* 10(7):667-681.

APPENDIX 1 – Review Materials for 2017 IRP Review

Review Materials Available to the 2017 LOBO Independent Review Panel

I. The following documents were provided in electronic format as required reading by the IRP prior to the 4-day workshop in Sacramento, CA on 4-7 December 2017:

Stanislaus River Watershed/Eastside Division RPA Actions

1. Stanislaus Operations Group (SOG) Annual Report of Activities (water year 2017)
2. Stanislaus Operations Group (SOG) Additional Materials

Enhanced Delta Smelt Monitoring (EDSM) Program

3. Summary of Enhanced Delta Smelt Monitoring (EDSM) Monitoring for LOBO Independent Review Panel

Results from the Old and Middle Rivers (OMR) index demonstration project

4. *Draft* Old and Middle River Flow Review: USGS Gauge Readings versus Index Calculations for 2011 through 2017
5. Extension of the Old and Middle River index demonstration project letter (March 13, 2017)

Draft proposed Shasta RPA amendment

6. NMFS' draft proposed Shasta RPA amendment: http://www.westcoast.fisheries.noaa.gov/publications/Central_Valley/Water%20Operations/nmfs_s_draft_proposed_2017_rpa_amendment_-_january_19__2017.pdf
7. Sacramento River Temperature Task Group (SRTTG) Meeting Materials (water year 2017).

II. The following additional reports were made available in electronic format for supplemental use in providing historical context for the IRP:

Enhanced Delta Smelt Monitoring (EDSM) Program

1. Smelt Working Group (SWG) Annual Report of Activities
2. 2015 DSEM/DSEE Proposal Report
3. The Enhanced Delta Smelt Monitoring Program Standard Operating Procedures
4. Enhanced Delta Smelt Monitoring Preliminary Abundance Analysis

5. Enhanced Delta Smelt Monitoring Preliminary Abundance Analysis Larval/Juvenile Life stages
6. Enhanced Delta Smelt Monitoring Preliminary Abundance Analysis Phase 3 Sampling
7. DSM TN 23. Fish density estimation in a zero inflated field with doubly truncated geometric sampling

Results from the Old and Middle Rivers (OMR) index demonstration project

8. NMFS Response Re: Extension of the OMR Index Demonstration Project (June 27, 2017)

Draft proposed Shasta RPA amendment

9. Martin, B. T., A. Pike, S. N. John, N. Hamda, J. Roberts, S. T. Lindley, and E. M. Danner. 2017. Ecology Letters 20(1):50-59 Phenomenological vs. biophysical models of thermal stress in aquatic eggs.
10. Mount, J., B. Gray, C. Chappelle, J. Doolan, T. Grantham, N. Seavy. 2016. Managing Water for the Environment During Drought: Lessons from Victoria, Australia. Public Policy Institute of California, San Francisco, CA. June 2016.
11. The Central Valley Temperature Mapping and Prediction (CVTEMP): http://oceanview.pfeg.noaa.gov/CVTEMP/
12. Reclamation's request for Shasta RPA adjustment: http://www.westcoast.fisheries.noaa.gov/publications/Central_Valley/Water%20Operations/bureau_of_reclamation_s_request_for_shasta_rpa_adjustments_-_august_2__2016.pdf
13. NMFS and Reclamation letter exchange regarding the Shasta RPA adjustment: http://www.westcoast.fisheries.noaa.gov/publications/Central_Valley/Water%20Operations/nmfs_response_to_reclamation_s_request_for_shasta_rpa_adjustments_-_august_17__2016.pdf
14. Reclamation's initial response to NMFS' draft proposed Shasta RPA amendment: http://www.westcoast.fisheries.noaa.gov/publications/Central_Valley/Water%20Operations/reclamation_s_response_to_nmfs_s_draft_proposed_2017_rpa_amendment_-_january_25__2017.pdf
15. Reclamation's detailed response to NMFS' draft proposed Shasta RPA amendment: http://www.westcoast.fisheries.noaa.gov/publications/Central_Valley/Water%20Operations/reclamation_s_detailed_comments_on_nmfs_s_draft_proposed_2017_rpa_amendment_-_march_22__2017.pdf

Technical Team Reports and Other Materials

16. Delta Operations for Salmonids and Sturgeon Group (DOSS) Annual Report of Activities
17. Sacramento River Temperature Task Group (SRTTG) Annual Report of Activities
18. Clear Creek Technical Team (CCTT) Annual Report of Activities
19. American River Group (ARG) Annual Report of Activities
20. Interagency Fish Passage Steering Committee (IFPSC) Annual Report of Activities
21. Summary Matrix of the NMFS and Service Coordinated Long-term Operation BiOps RPA actions

III. *The following background materials also were available to the IRP:*

- 2016 Long-term Operations Biological Opinions (LOBO) Informational and Update Meeting, December 6, 2016 http://deltacouncil.ca.gov/events/science-program-review/2016-long-term-operations-biological-opinions-lobo-annual-science
- Letter from National Marine Fisheries Service: Re: Proposed Modification to the Annual Review Schedule Required as Part of the Reasonable and Prudent Alternative in the National Marine Fisheries Service's 2009 Biological Opinion on the Coordinated Long-term Operation of the Central Valley Project and State Water Project, June 30, 2016 http://deltacouncil.ca.gov/docs/letter-national-marine-fisheries-service-re-proposed-modification-annual-review-schedule
- Letter from U.S. Bureau of Reclamation: Proposed Modification to the Annual Review Schedule Required as Part of the Reasonable and Prudent Alternative in the National Marine Fisheries Service's 2009 Biological Opinion on the Coordinated Long-term Operation of the Central Valley Project and State Water Project, April 25, 2016 http://deltacouncil.ca.gov/docs/letter-us-bureau-reclamation-proposed-modification-annual-review-schedule-required-part
- 2015 Annual Science Review:
    - Report of the 2015 Independent Review Panel (IRP) on the Long-term Operations Biological Opinions (LOBO) Annual Review (December 6, 2015). http://deltacouncil.ca.gov/docs/delta-isb-delta-science-program-isb-products-lobo/report-2015-independent-review-panel-irp-long
    - Federal Agencies' Response to the 2015 Independent Review Panel's Report
- 2014 Annual Science Review:
    - Report of the 2014 Independent Review Panel (IRP) on the Implementation of the Long-term Operations Opinions Reasonable and Prudent Alternative (RPA) Actions (December 11, 2014)
        - http://deltacouncil.ca.gov/sites/default/files/2014/12/2014-12-11-LOBO-2014-Report-Panel-Final.pdf

    - Federal Agencies’ Response to the 2014 Independent Review Panel’s Report http://deltacouncil.ca.gov/docs/delta-science-program-independent-review-lobo-science-program/response-delta-science-program
- 2013 Annual Science Review:
    - Report of the 2013 Independent Review Panel (IRP) on the Implementation of the Long-term Operations Opinions Reasonable and Prudent Alternative (RPA) Actions (December 7, 2013). http://deltacouncil.ca.gov/sites/default/files/documents/files/LOBO_2013_Report_Final_120613_FINAL.pdf
    - Federal Agencies’ Response to the 2013 Independent Review Panel’s Report http://deltacouncil.ca.gov/sites/default/files/documents/files/Joint_Federal_Response_to_DSP_Final_2_3_14.pdf
- 2012 Annual Science Review:
    - Report of the 2012 Independent Review Panel (IRP) on the Implementation of the Long-term Operations Opinions Reasonable and Prudent Alternative (RPA) Actions (December 1, 2012). http://deltacouncil.ca.gov/sites/default/files/documents/files/Report_2012_DSPIRP_LOOAR_120112_final.pdf
    - Review Materials, Background Information and Presentations http://deltacouncil.ca.gov/2012-long-term-operations-opinions-annual-review-%E2%80%93-review-materials-background-information-and-prese
    - Federal Agencies’ Response to the 2012 Independent Review Panel’s Report (July 19, 2013). http://deltacouncil.ca.gov/sites/default/files/documents/files/Federal_Agencies_Response_to_the_Panels_Report_July19_2013.pdf)
- 2011 Annual Science Review:
    - Report of the 2011 Independent Review Panel (IRP) on the Implementation of Reasonable and Prudent Alternative (RPA) Action Affecting the Operations Criteria And Plan (OCAP) for State/Federal Water Operations (December 9, 2011). http://deltacouncil.ca.gov/sites/default/files/documents/files/IRP_OCAP_RPA_2011_Final_Report_v2.pdf
    - Review Materials, Background Information and Presentations http://deltacouncil.ca.gov/science-program/2011-ocap-review-materials-background-information-and-presentations
    - Federal Agencies’ Detailed Response to the 2011 Independent Review Panel’s Report (June 20, 2012) http://deltacouncil.ca.gov/sites/default/files/documents/files/2012-06-20%20Joint%20Fed%20Resp%20to%20DSP%20for%20IRP.pdf
- 2010 Annual Science Review:
    - Report of the 2010 Independent Review Panel (IRP) on the Reasonable and Prudent Alternative (RPA) Actions Affecting the Operations Criteria and Plan (OCAP) for the State/Federal Water Operations (December 9, 2010).

http://deltacouncil.ca.gov/sites/default/files/documents/files/workshop_OCAP_2010_IRP_RPA_Final_Report_121310_0.pdf

  - Review Materials and Presentations http://deltacouncil.ca.gov/events/science-program-workshop/workshop-ocap-integrated-annual-review
  - Joint Department of Commerce and Department of the Interior Response to the Independent Review Panel's (IRP) 2010 Report of the Reasonable and Prudent Alternative (RPA) Actions Affecting the Operations Criteria and Plan (OCAP) for the State/Federal Water Operations (March 9, 2011) http://deltacouncil.ca.gov/sites/default/files/OCAP_2010/workshop_OCAP_2010_review_detailed_response_letter_032111.pdf
- NMFS 2009 RPA with 2011 amendments http://www.westcoast.fisheries.noaa.gov/publications/Central_Valley/Water%20Operations/Operations,%20Criteria%20and%20Plan/040711_ocap_opinion_2011_amendments.pdf
- USFWS BiOp on the Long-Term Operational Criteria and Plan (OCAP) for coordination of the Central Valley Project and State Water Project (pages 279-282 and 329-356) https://www.fws.gov/sfbaydelta/Documents/SWP-CVP_OPs_BO_12-15_final_OCR.pdf
- National Academy of Sciences March 19, 2010, report https://www.nap.edu/catalog/12881/a-scientific-assessment-of-alternatives-for-reducing-water-management-effects-on-threatened-and-endangered-fishes-in-californias-bay-delta
- VAMP peer review report 2010: http://www.sjrg.org/technicalreport/2009/2010-VAMP-Peer-Review-Panel-Report.pdf
- State Water Board's 2010 Delta Flows Criteria Report: http://www.waterboards.ca.gov/waterrights/water_issues/programs/bay_delta/deltaflow/docs/final_rpt080310.pdf

IV. ***The following additional materials were made available at the Workshop in Sacramento for supplemental use of the IRP*:**

- **Public Comments**, including the following document:

  Anderson, J. (December 6, 2017) Comments on the egg mortality model used to develop the Shasta RPA – an analysis funded by San Luis & Delta-Mendota Water Authority

APPENDIX 2 – Improved predictions of the USGS tidally-filtered, 3-day-average, OMR flow estimate (GS)

The IRP suggests that the following approach can predict the USGS tidally-filtered estimates of OMR flow with greater accuracy than that of the Index method alone.

According to the OMR report (p. 2), the IND value for yesterday is used for today's operating decisions. Thus, we set day k equal to "yesterday relative to today's operations. We also assume that IND(k) is used as a prediction of GS(k), for any day k (this was not clear in the report). GS is a tidally-filtered 3-day average of raw data from 2 gages. The report states that GS is available only after a 3-day lag (p.2). Some of this delay must be due to institutional delays, because all of the raw data need to calculate GS(k-1) has been acquired by day k. That is, USGS gage data for days (k-2), (k-1) and k should be available on day k, to calculate GS(k-1), as the 3-day, tidally-filtered average centered on day (k-1). Thus, we assume that with communications improvements, the true, measured value of GS(k-1) could be available on day k.

Our proposal for improving on the IND prediction is inspired by the dynamic error-correcting properties of a Kalman filter (Bozic 1979). Our method assumes that GS changes fairly slowly on a daily time scale. In other words, it assumes that, on most days, the value of GS on day k, denoted by GS(k),is not very different from the flow on the previous day, GS(k-1). This assumption is reasonable because GS(k) and GS(k-1) share input data from 2 of the 3 days used in the calculation of each value. Hence, GS(k-1) should be a fairly good prediction of GS(k).

Thus, on any day k, the values of IND(k) and GS(k-1) should both be available as independent, fairly accurate, predictions of GS(k). We suggest using a weighted average of these two predictions, to give a new prediction of GS(k) that is closer to its true value than would be predicted by IND(k) alone.

The weighted-average predicted value, GSP(k), would be calculated as:

$$GSP(k) = w(k)*IND(k) + (1-w(k))*GS(k-1) \qquad (1)$$

where the weight, w(k), can vary from day to day, and could be positive or negative. To specify a weight value, w(k), assume that weights will also change little from day to day. Thus, as a value for w(k), we propose using the estimated weight value from day (k-1) that would have made the predicted and observed values of GS(k-1) perfectly agree. On day k, the true value of GS(k-1) is known. And the predicted value of GS(k-1) is given by the right-hand side of Equation 1, substituting (k-1) for k. So, the estimated weight, $w_E(k-1)$, that makes the predicted and observed values agree for day (k-1), will satisfy:

$$GS(k-1) = w_E(k-1)*IND(k-1) + (1-w_E(k-1)*GS(k-2) \quad (2)$$

Solving for $w_E(k-1)$ allows it to be directly calculated from quantities that are all known on day k:

$$w_E(k-1) = [GS(k-1) - GS(k-2)] / [IND(k-1)-GS(k-2)] \quad (3)$$

We propose to let the weight w(k) be equal to its most recent optimal value, $w_E(k-1)$, in order to implement Equation 1 on day k. As an option, one should get a more robust estimate for w(k) by setting it equal to the average of the w_E estimates from several (4 or 5) of the most recent previous days. Finally, to allay DWR's concern with missing data for GS(k-1) and/or GS(k-2), one could just use IND alone as the predictor, on days when one or both of these GS values are missing.

This method appears promising to the IRP, and its accuracy should be relatively easy to assess using historic data sets for GS and IND. If our method does indeed reduce prediction errors, then DWR may be motivated to obtain GS(k-1) values on any day k, and hence be able to implement the method. If this is not feasible, and the most recent available value of GS is from day (k-2), then one can instead use a weighted average of IND(k) and GS(k-2), by modifying the above approach.