Independent Science Review

Evaluation of key assumptions underlying analyses of delta smelt survey data

Reviewer #2

Charge Questions

1. Are goals, objectives, hypotheses and questions clearly articulated and internally consistent?

The proposal very clearly lays out 2 major topics (catchability and false zeros; autocorrelation) with 2 testable questions under each topic. The proposal explains what should be done if significant effects on catchability (equation 3) or autocorrelation are found (equation 4). Thus, there is a logical consistency in how the questions are formulated as testable hypotheses, and how the results COULD be incorporated into the fundamental estimation of CPUE values.

It is not clear how far the proposal will go in this next step of deriving new estimators for CPUE that include non-constant catchability and autocorrelation, or both. I list this issue here (plus other places) because it relates to the internal consistency of this question. Part of consistency is whether the analyses ultimately end up answering the stated questions. The proposal is very clear about identifying whether catchability and autocorrelation are problems that need to be dealt with. But the proposal lacks a bit in closing the circle and determining whether these issues are important in the actual estimation of CPUE in the context of how CPUE is used. I do not see any comparisons like this for question 1b. The autocorrelation analysis (question 2b) comes closer in that there is comparison of without and with CAM for reduced datasets, but the comparison is still within the new analyses to assess the effects of CAM and not a comparison to the old CPUE estimates. In terms of ensuring these analyses are of full value, new CPUE estimates should be generated, perhaps with non-constant catchability and then with autocorrelation correction and then with both. The old way of generating CPUE should be repeated on this same dataset used in the proposal, and a comparison made between the new and old estimates. This would complete the logic of the proposal. Otherwise, we will end up with two more issues on the long list of **POSSIBLE** things that **COULD** make the CPUE values biased and highly uncertain without knowing the management implications because the link back to the original estimates is missing.

2. Will the proposed work contribute to our overall understanding of Delta Smelt abundance and distribution?

As the proposal nicely states as a basic limitation of the proposal, all of the proposed analyses use the sampling design (station locations, frequency) as it exists. At some point, the many re-analyses of the same data from the monitoring is trying to get "blood from a stone." This proposal brings some new investigators to the data analyses, armed with some statistical approaches not previously used. This proposal can add to the evolution of our understanding of the monitoring data but not to a revolution. Thus, I do not see how refining the CPUE estimates will affect how the monitoring data are used for spatial distributions. Those types of analyses are usually done on a relative basis and semi-quantitative so adjusted CPUE values (even with lower variance) would not change the interpretations for the data when used for spatial distribution mapping. One of the biggest issues on the monitoring and spatial

distribution is whether there are places not sampled where delta smelt occur; no analysis of data from the existing sampling can determine the effects of smelt in unsampled places on CPUE and assessment of population status.

Perhaps if entrainment analyses, such as Kimmerer's[a], were repeated, which used information about how close delta smelt were located to the pumping facilities, revised spatial information from adjusted CPUE value could affect the estimation of entrainment. However, in my opinion, this is unlikely given the likely small to moderate changes in determination of vulnerabilities to entrainment that would result from updated CPUE values and the high uncertainties in other aspects of the entrainment analyses.

There is some potential for the proposal to influence the trends analyses done with the monitoring data. This is why I suggested that old and new estimates be generated and compared as part of my comments to question 1. It would seem, although not explicitly stated in the proposal so I may be wrong, that a major benefit of the new CPUE estimates is their lower variance. This would not greatly affect the trends themselves but would inform the science and resource management community about how much confidence to have in large percent changes but small magnitude changes in the indices. For example, I have heard people interpret a change in an index of 1 to 1.5 as a meaning the population has increased by 50%, when the index used to be at values of 10 and 20 and 1 to 1.5 is just noise.

3. Are the budget and the schedule reasonable and adequate for the work proposed?

If the proposal is expanded to include analyses comparing old and new estimates, then the budget is reasonable but not a bargain. Charging consulting rates ($1200 per day) for a relatively large effort (60 days) that is research-oriented (grant-like) is OK, as it is the total dollars (rate x days) that matter. About $89K is reasonable for a total budget, although most of the data has recently been analyzed by the PI so the time consuming part of data acquisition and QA/QC should be mostly done. I would like to see a systematic set of analyses comparing the old values and the new CPUE results. Also, I would also ask for 3-5 journals for the manuscript (which is critical to ensure the proposal results are fully used) to be named to ensure that the manuscript is submitted to a highly credible journal. This is not a reflection of the PI but the new reality of so many journals, some of which use questionable processes for peer review – i.e., now you can pay and publish.

4. Does the proposed work take a feasible approach to addressing questions such as spatial autocorrelation in the data, correction factors for covariates affecting catchability and uncertainties concerning abundance indices? Are there other approaches that would be more appropriate for this analysis?

The proposed analyses are appropriate for identifying and quantifying the effects of autocorrelation and non-constant catchability. How to then revise and improve the estimation of CPUE and the abundance indices is more vaguely described and it is not clear to me how far the proposal will go in actually deriving new CPUE values and documenting the improvement by the new methods over older estimation methods and the degree of agreement and disagreement between new and old values.

[a]Kimmerer, W.J. 2011. Modeling delta smelt losses at the south Delta export facilities. San Francisco Estuary and Watershed Science [online serial] 9(1): article 5.

Miller, W.J. 2011. Revisiting assumptions that underlie estimates of proportional entrainment of delta smelt by State and Federal water diversions from the Sacramento-San Joaquin Delta. San Francisco Estuary and Watershed Science [online serial] 9(1): article 4.

In my opinion, identification of the issues is the first step, then following through and generating better estimates is the second step, and then enabling a bridge from the old estimates to the new estimates is the final step and critical for management. The proposal is very strong in identification and gets progressively weaker (or maybe just more vague) with the next two steps. Thus, how far the proposal will go in the generation of "correction factors" (as asked in the charge question) and how these will be compared to earlier estimates should be part of the proposal and clearly agreed upon by all involved. I would suggest the development of a short plan of analysis of how new and old estimates will be compared.

The proposal deals very well with identification and quantification of catchability and autocorrelation issues and also explains that all of the proposed analyses use the sampling design as it exists, which is a limitation of the proposed analyses. However, I suggest they also look at some approaches one may consider as being in between dealing with the data as is and trying to determine information where data were not collected or were not collected frequently enough. I am thinking about approaches such as bootstrapping and power analyses to determine the ability of the monitoring data to generate credible CPUE-based indices for trend analysis and deriving spatial distributions. The final index values are often mis-interpreted by people giving them too much confidence or by people dismissing the indices completely. Using approaches to quantify the robustness of the indices to various combinations of stations and sampling frequencies would be very useful. Analysis such as subsampling stations, sampling frequencies, certain year types, etc. could provide very useful information about the influence points (e.g., stations, years) and robustness of the indices. Similarly, doing a power analysis on the derived indices under various assumptions of variability in the data would also provide information on how the indices should be properly interpreted. Note that I was asked to review this proposal and not to propose my own project. But I do see a relationship between this proposal that will refine the CPUE indices and the approaches like I suggest (bootstrapping, power analysis) that quantify how old and new CPUE estimates should be best interpreted.

5. Will the results from this proposal add value to other work that uses these survey data, and if so, how much confidence can be added from the analyses?

I think the methods are transferable but the results for delta smelt are not transferable to other species. False zeroes, non-constant catchability, and autocorrelation likely have a strong species-specific (and even life stage-specific) component. For example, one cannot infer that autocorrelation is important for striped bass or other species based on it being identified as important form the delta smelt analyses. With this proposal, the methods will now be known and available to the Delta science community and thus the delta smelt analyses would serve a template for the methods to be applied to other species.

The proposal mentions life cycle modeling as a downstream user of the results. Revised CPUE estimates could influence the life cycle modeling that relies on fitting the population dynamics model to the monitoring data indices, such as Maunder[b] and the ongoing effort of Newman. Again, the degree of change likely in the CPUE indices would probably not change the modeling results very much. Revised CPUE value would have less effect on other life cycle modeling, such as mine[c], that uses the data in a less formal way for model calibration and corroboration. There are also statistical analyses that use the

[b]Maunder, M.N., and R.B. Deriso. 2011. A state-space multi-stage lifecycle model to evaluate population impacts in the presence of density dependence: illustrated with application to delta smelt. Canadian Journal of Fisheries and Aquatic Sciences 68:1285-1306.

[c]Rose, K.A., W.J. Kimmerer, K.P. Edwards, and W.A. Bennett. 2013. Individual-based modeling of Delta Smelt population dynamics in the upper San Francisco Estuary: I. Model description and baseline results. Transactions of the American Fisheries Society 142: 1238-1259.

data related to trend analysis[d] and quantifying habitat quality and quantity[e]. The latter is highly spatial-specific so revised CPUE estimates that account for non-constant catchability and autocorrelation would have the largest effect. In general, I do not think the new estimates would change the major conclusions of these population dynamics and statistical modeling analyses, but that is purely a guess. The only way to really know is to generate the new estimates, compare to them to the old estimates, and perhaps repeat some of these downstream analyses with the new estimates.

6. Are the proposed analyses suited to constructively informing management actions such as those associated with the existing biological opinions?

The proposed analyses can inform the management actions, but unless the analyses show **GREATLY** revised estimates of CPUE, new estimates would not likely affect the decision-making. I do not think the proposed adjustments can change the CPUE values to a large enough extent to change the assessment of the state of the delta smelt population, especially since all of the analyses in the proposal are data-based and use the data as it currently exists. New estimates can be added to the older estimates, which would increase confidence, but if the new estimates are orders of magnitude different then old estimates, this would lead to further investigation before the new values are adopted and the old values completely dismissed.

I suspect the most likely implications of the new analyses could be to discourage over-interpretation of the indices, which would be very useful in the management discussions. I am not convinced that refining and improving the CPUE indices within the existing monitoring data would greatly change the ultimate decisions. The big issues that could "shake-up" management are where the monitoring is not sampling and whether the current station configuration (locations) is still effective at capturing Delta smelt in a way that the data can be interpreted as representative of the population. The proposed analyses do not directly address this. But more examinations of the data, if done so the final implications are documented, would help inform the decision-making process.

7. Does the proposal address the most important potential data limitations relevant to questions about Delta Smelt entrainment and the influence of Fall outflow?

My answer is similar to my responses to questions 2 and 6. In summary, I do not think the revised indices would **GREATLY** affect entrainment estimates because of the high uncertainty in other aspects than spatial distribution that are involved with estimating entrainment effects and because the relationship to outflow is usually done with CPUE values aggregated over time and space. If all of the CPUE values are adjusted, including somewhat unevenly, then this would likely shift the relationship but not likely greatly change the relationship (e.g., linear, monotonic increasing, dome-shaped).

If and when the absolute values of the derived indices are used for management (e.g., RPAs), then revised CPUE values (e.g., shifted down) would need to be rectified with benchmark or threshold values that were based on the old CPUE estimates. But depending on the evenness of the adjustments to CPUE over time and space, this rectification could be very simple (e.g., scale change) to very complicated (e.g., CPUE in years of high outflow adjusted very differently than in years of low outflow).

[d]Mac Nally, R., J.R. Thomson, W.J. Kimmerer, F. Feyrer, K.B. Newman, A. Sih, W.A. Bennett, L. Brown, E. Fleishman, S.D. Culberson, and G. Castillo. 2010. Analysis of pelagic species decline in the upper San Francisco Estuary using multivariate autoregressive modeling (MAR). Ecological Applications 20: 1417–1430.

[e]Nobriga, M.L., T.R. Sommer, F, Feyrer, and K. Fleming. 2008. Long-term trends in summertime habitat suitability for delta smelt (*Hypomesus transpacificus*). San Francisco Estuary and Watershed Science [online serial] 6(1): article 1.

Summary Comments

I think this proposal has high merit, given the wide use (and sometimes mis-use) of the monitoring data for delta smelt. The PI is very well qualified, and the total budget is reasonable with the addition of a systematic "old estimates versus new estimates" comparison. My concerns center on ensuring the analyses are carried to an end point that best helps how the data are used for management and to avoid having another analysis done somewhat independently that just adds to the noise and confusion rather than to clarify some important issues. The potential for this proposal to clarify some important issues is clearly there, but not the extent and completeness in the current proposal that I would like to see. However, my concerns can be addressed with some additional analyses.