Proposal 'Evaluation of Key Assumptions Underlying Analyses of Delta Smelt Survey Data'
Response to Reviewers
Robert J. Latour
June 2015

Reviewer #1

1.) Are goals, objectives, hypotheses and questions clearly articulated and internally consistent?
The Reviewer indicated that the goal of the temporal/spatial autocorrelation component of project was not clear. Below is some additional text that will hopefully provide additional clarification. Observed counts of fish relative abundance from a survey where samples are taken close in time (e.g., same day, month) or space (e.g., near each together geographically) can sometimes be related and more similar to one another than those take farther apart temporally or spatially. If either of these patterns is present in the residuals of statistical models fitted to such data, then the key assumption of independent and identically distributed residuals is violated. Such assumption violations may bias parameter estimates used to assess changes in catchability and subsequent model-based predicted annual indices of relative abundance. Autocorrelation among residuals can also increase Type I error rates (falsely rejecting the null hypothesis of no effect for covariates) which implies conclusions regarding the statistical significance of covariates may be suspect.

The second objective of the proposal is designed to evaluate the degree of temporal/spatial autocorrelation among residuals of statistical models fitted to the FMWT delta smelt survey data. If either form of autocorrelation is present in the data and not properly accounted for, then statistical modeling results and associated predictions may be incorrect. Since survey data are the backbone for understanding patterns in delta smelt relative abundance (and thus overall abundance assuming the survey data are representative), detecting and, if present, accounting for autocorrelation is important to ensuring that inferences about the significance of covariates are correct and model based predicted indices are unbiased.

2.) Will the proposed work contribute to our overall understanding of Delta Smelt abundance and distribution?
I agree with the Reviewer that the proposed work is an extension of the Latour (2015) manuscript and also that the value of the project will depend on identifying meaningful covariates for testing. Covariates measured synoptically with sampling can be evaluated in a straightforward manner, however, derivation of other covariates of interest may require analyses of auxiliary data. Communication with CAMT members and associated researchers will aid in identifying those covariates and datasets and additional analyses will be conducted as needed.

3.) Are the budget and the schedule reasonable and adequate for the work proposed?
I appreciate the Reviewer's comment that more time should be spent on assessing the importance of covariates hypothesized to affect catchability, however, if temporal/spatial autocorrelation of model residuals is notable and not accounted for, then virtually all model-based inferences of covariate significance and predications of relative abundance have the potential to be biased. Ensuring the results are sound is a key guiding principle of the project, so investing time to analytically account for correlated residuals is thus critical. I am committed to working to provide a rigorous and comprehensive treatment of both proposal objectives.

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 could be used in the analysis?

I will ensure that the recent work cited by the Reviewer is incorporated into my analyses, to the extent possible. An advantage of the zero-inflated GLMs is that the effects of covariates such as tidal stage (as mentioned by the Reviewer) on the probability of a false zero survey catch can be explicitly evaluated. The lack of a formalized list in the proposal was not meant to give the impression that only a small suite of covariates will be considered. I am open to including any and all covariates of interest, particularly those based on sound ecological or biological hypotheses (see comments above in bullet (2)).

Since the zero-inflated negative binomial GLM model structure proved useful for analyzing the FMWT delta smelt survey data (Latour 2015), it is offered as a starting point in the proposal. However, other model structures (e.g., traditional and zero-inflated GAMs) will also be considered. Identifying the proper underlying model structure is paramount to the success of the project, so this issues will receive considerable attention.

Although not explicitly stated in the proposal, I plan to further examine patterns in predicted false zero survey observations. For example, constructing spatiotemporal patterns in false zeros could prove useful in detecting systematic changes in habitat use/preferences over time. Moreover, the predicted false zero survey observations could be related to environmental covariates as a means of further identifying potential drivers of delta smelt population dynamics. These analyses will hopefully address (at least in part) the remainder of the Reviewer's comments to this question. However, it is important to recognize that the proposed work involves applying statistical models to existing survey data. By definition, these models are not mechanistic or dynamic, so modeling processes is not possible within the proposed analytical framework. The value of statistical models is that they can elucidate relationships that may be in play mechanistically and the result of underlying biological, ecological, behavioral, etc. dynamics. Once results are generated from statistical models, it is important place them in a biological, ecological context and postulate potential explanatory mechanisms. This will certainly be a point of emphasis as the project moves forward.

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

Admittedly, the proposal is vague in terms of articulating how the work will link to other ongoing/planned research. However, following submission of the proposal, I travelled to California to attend a DSST meeting and interact with other CAMT sponsored investigators. This trip was very informative and it allowed me to develop a better understanding of the Entrainment (Grimaldo et al.) and Fall outflow (Fleishmen et al.) research projects. This proposed work dovetails with those projects in numerous ways. For example, predicted patterns in the spatial distribution and habitat use of delta smelt (both counts and false zeros) will help inform models of entrainment since delta smelt interactions with intake pumps is largely a spatially driven process. For the Fall outflow study, understanding the covariates that significantly affect relative abundance and probabilities of false zero are key

to formulating the functional form of the encounter probability equation inherent to the patch-occupancy modeling framework outlined in the proposal.

6.) Are the proposed analyses suited to constructively inform management actions, such as those associated with the existing biological opinions?
The Reviewer highlighted an important intended outcome of the proposed work; namely, development of potentially improved indices of delta smelt relative abundance and statistical relationships among survey catches and external covariates. Both of these are embedded in the management process and existing biological opinions. More broadly, this project is intended to aid management efforts by evaluating potential sources of scientific uncertainty in delta smelt survey data that, to date, have not been thoroughly evaluated.

7.) Does the proposal address the most important potential data limitations relevant to questions about Delta Smelt entrainment and Fall outflow?
The Reviewer noted that existing surveys targeting delta smelt may be limiting since potentially important areas are not routinely sampled (e.g., the shoal-channel interface). I support the general philosophy that new, well designed field studies can be informative and should be implemented when possible to increase our understanding of the dynamics of natural resources. Despite potential limitations of existing data, they still hold value and are relevant to management deliberations, provided that interpretations of those data are offered with the necessary/appropriate caveats. I am sensitive to the issue that surveys in the Delta may be limiting – in reality, most all fish surveys worldwide have limitations – however, the FMWT and SKT represent two principle sources of delta smelt information and, accordingly, it is important to keep the moving science forward and in-line to support management objectives.

Reviewer #2

1.) Are goals, objectives, hypotheses and questions clearly articulated and internally consistent?
The Reviewer raised an important point regarding internal consistency of the proposed research questions and 'closing the circle' through a comparative analyses. This idea is a clear omission in the proposal that I embrace fully. Efforts will be made to stepwise compare predicted relative abundance indices and significance of covariates from analyses that do and do not account for key covariates that affect catchability and temporal/spatial autocorrelation.

2.) Will the proposed work contribute to our overall understanding of Delta Smelt abundance and distribution?
The potential limitations of the existing designs of the FMWT and SKT surveys have been acknowledged, however, it is important to continually evolve the inferences obtained from these data sources since they are fundamental underpinnings of management.

An outcome of the proposed work will be refined estimates of delta smelt relative abundance, however, it is also my intent to focus attention of developing a more refined understanding of delta smelt distribution. Perhaps this goal was not clearly evident in the proposal. By focusing some model formulations to more explicitly incorporate spatial covariates (e.g., station, region, and latitude/longitude), I believe the proposed work has the potential to shed new light on spatial patterns of delta smelt and potential changes in them over time. To date, very little

attention has been placed on modeling survey catches of relative abundance, so the proposed work along with the recently published manuscript by Latour (2015) are attempts to fill this information gap. As mentioned in my response to question 4 for Reviewer #1, I plan to delve more deeply into trying to understand patterns and potential drivers in false zero probabilities. I believe more fully understanding false zeros can aid our understanding of delta smelt habitat utilization as well as conditions under which they are sampled successfully and unsuccessfully.

3.) Are the budget and the schedule reasonable and adequate for the work proposed?
Specific journal outlets for work products were not articulated as I am open to suggestions. In general, I strive to place my research results in highly credible journals and I believe I have done so throughout my career. Candidate journals might include: Estuaries & Coasts, Marine and Coastal Fisheries, and Estuarine and Coastal Shelf Science to name a few.

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 could be used in the analysis?
The Reviewer more explicitly outlined ideas on how to address the previously mentioned idea of 'closing the circle' through suggestions of bootstrapping, power analyses, and subsampling strategies (random or predefined stations, years, etc.) to evaluate robustness in relative abundance indices and inferred significance of covariates. As indicated above, I fully embrace these ideas and will work to incorporate them into the project. Thanks to the Reviewer for raising these important ideas.

5.) Will the results from this proposal add value to other work that uses these survey, and if so, how much confidence can be added from the analyses?
It is possible that accounting for variable survey catchability and temporal/spatial autocorrelation will not have a huge effect on trends in delta smelt relative abundance, but as the Reviewer noted, we just do not know. Most all of the historical research on pelagic fishes in the Delta that utilized survey data has not focused on model catch-per-unit-effort. Latour (2015) was the first to rigorously treat the FMWT data with statistical models and I believe this project will be the first to do the same for the SKT data. I am confident that this work will bring to light new ideas about delta smelt and the longstanding sampling programs used to draw inferences about the status of this species.

6.) Are the proposed analyses suited to constructively inform management actions, such as those associated with the existing biological opinions?
The Reviewer's comment here basically falls under the broad umbrella of further understanding uncertainty, which is a fundamental and motivating principle of this project. In addition to the potential outcomes of the work described in the proposal and in this document, it is also entirely possible for this work to aid sampling strategies for delta smelt moving forward. For example, if temporal/spatial autocorrelation is substantial within the data sets, then to mitigate this issues, adjustments could be made to the order that stations are sampled, time of day of sampling activities, and possibly to the overall number of stations sampled per cruise. This project has the potential to inform refinement of survey activities designed to gain efficiencies.

7.) Does the proposal address the most important potential data limitations relevant to questions about Delta Smelt entrainment and Fall outflow?
The Reviewer restated thoughts raised in questions (2) and (6), so here I would simply point to my responses to those comments.

Reviewer #3

The Reviewer provided some overarching remarks that summarize aspects of sampling theory and the differences between design- and model-based approaches to statistical inference. Highlighted in the prose is the general need for randomness when selecting sampling locations and the Reviewer pointed out that both the FMWT and SKT survey are not based on designs that employ random site selection; they are fixed station sampling programs. This is an important limitation that was noted in the proposal and in other venues of discussion about the data underpinning discussions about the status of delta smelt. Only modifications to survey operations and/or additional field work can overcome this limitation, however, as mentioned previously, I believe the historical data still hold value and rigorous analyses of these data with appropriately acknowledged caveats/assumptions have the potential to continue informing management.

Two additional points raised here by the Reviewer are worth noting. Regarding attributing estimated effects from statistical modeling activities to q_e rather than q_a, this issue was acknowledged in the proposal and requires making the assumption that localized delta smelt abundance in spatiotemporal subunits (e.g., region by month combinations) is constant such that replicate tows within those subunits can be considered to be operating on the same underlying local population. These replicate tows will drive the estimation of effects due to tidal stage, time of day, and other covariates hypothesized to influence survey catchability. Given that regions within months are typically sampled over a few days, the closed population assumption may be reasonable. Lastly, the autocorrelation analysis is expected to define the minimum distance between samples required to achieve statistical independence, as noted by the Reviewer.

1.) Are goals, objectives, hypotheses and questions clearly articulated and internally consistent?
The Reviewer indicated the goals/objectives of the proposal were clearly stated. Additional discussion is provided above if necessary.

2.) Will the proposed work contribute to our overall understanding of Delta Smelt abundance and distribution?
A primary goal of the project is to shed more light on the relationship of hypothesized covariates and survey catches. Modeling results is this vein are intended to represent our best understanding of such relationships, short of initiating new data collection programs.

3.) Are the budget and the schedule reasonable and adequate for the work proposed?
Charges questions (3) and (4) were switched by the Reviewer. No concerns were raised on budget and work schedule.

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 could be used in the analysis?
Responses to comments about local availability and potential data limitations have been described previously.

5.) Will the results from this proposal add value to other work that uses these survey, and if so, how much confidence can be added from the analyses?
Responses to comments about local availability and utility of detecting and accounting for spatial autocorrelation have been described previously.

6.) Are the proposed analyses suited to constructively inform management actions, such as those associated with the existing biological opinions?
As the Reviewer noted, the proposed analyses may identify some covariates that influence survey catches which, in turn, may increase understanding of delta smelt population distribution. This is, in fact, a fundamental global objective of the proposed work.

7.) Does the proposal address the most important potential data limitations relevant to questions about Delta Smelt entrainment and Fall outflow?
Responses to comments about potential data limitations have been described previously.