Characterization of relations between autumn outflow and survival, recruitment, and habitat quality for delta smelt (*Hypomesus transpacificus*)

Collaborative Adaptive Management Team Work Element 3-1-3

Response to reviews

Executive Summary

The Scoping Document outlines a staged structure in which estimates of survival and recruitment would be linked to covariates in several ways: correlations with all suggested covariates, correlations with fall outflow only, correlations with covariates except fall outflow, and correlations of fall outflow to the other covariates. It appeared to the Panel the plan of proposal was in part disconnected from the Scoping Document goals.

> We do not feel that this characterization of the scope of work is accurate. The scope of work does not outline a staged structure. Instead, it simply focuses on the types of information that might be provided by the analyses. It notes, "Analyses ideally will allow inference to topics including but not limited to the following . . . ," where *the following* includes the extent to which environmental variables hypothesized to be strongly associated with autumn survival and recruitment of delta smelt are correlated, the strength of association between fall outflow and autumn survival and recruitment of delta smelt, the strength of association between environmental covariates and autumn survival and recruitment of delta smelt if fall outflow is not included in the model, and the strength of association between fall outflow and other environmental variables that are strongly associated with autumn survival and recruitment of delta smelt.

The weak points of the proposal include potential mismatch between the spatial/temporal scales of the models and the smaller scales that likely control Delta smelt movement and survival, imprecise descriptions of habitat quality measures and overly simplified statistical and simulation models of Delta smelt movement and survival.

> Unfortunately, the available data largely were not collected at the resolution at which some hypotheses suggest the drivers of movement and survival operate. Additionally, we explained in the proposal that our final selection of measures of habitat quality (covariates) will be conducted in consultation with members of the Delta Smelt Scoping Team (DSST), Collaborative Adaptive Management Team (CAMT), Ken Newman's research group, and others with expertise on Delta smelt and the upper San Francisco Estuary.

Descriptions of how simulation and statistical models will be linked and how biases will be evaluated are not articulated in the proposal. Finally, the mismatch between the coarse scales of the survey data compared to the small scale that recent research indicates controls Delta smelt detection and movement may limit inferences from the proposed work.

We will use simulation to construct models with known relations between physical covariates, such as fall outflow and turbidity, and vital rates, such as survival or movement. These relations are defined mathematically as functions. For example, survival in a given month is a function of the outflow in that month and a coefficient that defines the strength of the relation between survival and fall outflow. The simulation models are then used to calculate abundances over time. A given abundance is used to initialize the model, and subsequent abundances are calculated as a function of the vital rates.

The simulation model also can be used to simulate catches of delta smelt. Simulated catches are affected by both the proportion of fish captured and stochastic measurement error. Because the dynamics of both the population and the observation process are known in the simulation model, we can simulate catches given different levels of data quality (i.e., observation error).

We subsequently will construct a separate estimation model to estimate the relations between physical covariates and vital rates given the simulated catch data. The model uses statistical methods to estimate the most likely coefficients of the environmental covariates given the simulated catch data. This process can be carried out many times to understand the probability of estimating the true underlying coefficient values (those used in the simulation model) from the simulated catch data. This process also can be repeated given different levels of measurement error to understand how data quality may affect inference from the estimation model.

Estimates of the model coefficients that are consistently below or above the true coefficient value indicate bias in the ability of the estimation model to reflect the underlying relations between physical drivers and population vital rates. Alternatively, the estimated coefficients could have an extremely large range of values relative to the true coefficient value. The latter outcome indicates that the data may not provide information on the relation between the vital rates and the physical covariates, and thus the estimated coefficients are not precise. If the bias is low and the precision is high, then one might expect the estimated coefficients to accurately reflect the true relationships.

As the panel indicated, there may be an interest in understanding the extent to which the proposed model reflects movement at finer temporal resolution. It may be possible to use the Rose et al. (2013) individual-based model as the simulation model. The estimation model would be our proposed abundance model, in which estimates of movement parameters at coarser temporal resolution could be compared to the finer-resolution movement rules used in the Rose et al. (2013) model.

The Panel suggests the research plan could be improved if the proposal goals were prioritized to better align with the Scoping Document and to better assess the effects of data limitations and model assumptions on estimates of survival and abundance and their relationships to covariates.

Again, we do not feel that the proposal is misaligned with the scope of work. Data limitations are addressed in the previous response.

The Panel suggests the basis framework of modeling survival and abundance of Delta smelt as defined by equations 1-10 in the proposal are for the most part useful, with the exception of the movement model which currently assumes that regional scale signaling of habitat quality drives movements between regions. The Panel suggests consideration of movement in terms of tidal "surfing" processes that are controlled by local environmental conditions. This is a challenging endeavor that is likely to involve considerable effort and input from the Delta smelt research community.

> The DSST has suggested that we leverage the ongoing work of the entrainment investigators rather than expanding our scope of work to include fine-resolution modeling of behavior or movement.

Spatial and temporal scales relevant to Delta smelt will need to be identified and appropriate covariates developed. For example, instead of correlating region- scale abundance of Delta smelt with regional seasonal flows averaged across seasons, the patterns of catch probabilities at individual stations may need to be related to the value of covariates at the time of the trawls. The panel understands that such an examination of abiotic values at this level of temporal and spatial resolution would likely require significant additional work. The panel believes that such an approach would expand the utility of the regional-scale model.

> We cannot estimate detection (catch) probabilities at the station level with the available data because multiple trawls were not conducted within each time step. We are limited to assuming that the statistical distribution of detection probability is applicable to each station within the region.

Besides suggesting an evaluation of the appropriate scales and functional forms of the statistical model, the Panel suggests expanding the simulation modeling to explore effects of modeling different scales and functional forms for survival movement and recruitment. The Panel sees a comparison of statistical and simulation models with different dimensions and scales as a potential way to evaluate the statistical model biases. For example, fitting a lower dimension statistical model to simulated data generated by higher dimension individual based model of Delta smelt (e.g. Rose et al 2013) would provide information on bias resulting from misspecification and simplification of the statistical model.

> The Panel has concerns about the ability of the model to capture movement patterns that are on shorter time scales than the observed data. In general, the Panel feels that the dynamics of the model may be too simple to reflect the actual underlying dynamics of delta smelt in the estuary.
>
> The state process of movement can be modeled in myriad ways. Daily movements can be modeled as a function of tidal signals, and transitions (movements) among patches can be simulated at fine temporal resolution. However, it is difficult to estimate the parameters at fine temporal resolution because data were collected at coarser temporal resolution, and there are many mechanisms by which observed density at one point in time can lead to observed density at another point in time. Multiple combinations of parameters could yield the same observations. Without data on the movements of individual fish (e.g., from

satellite tags), it is virtually impossible to elucidate movement dynamics on the basis of a statistical model.

If the mechanisms of movement can be structured through an expert elicitation, then models of movement at fine temporal resolution can be developed. However, such a model is unlikely to provide much insight into movement dynamics given that assumptions about movement must be made a priori during the process of model development.

It might be feasible to work with Rose's individual-based model to explore whether movement dynamics can be estimated with the current sampling methods for Delta smelt. It also might be feasible to examine what level of sampling, such as sampling at stations or tagging individual fish, would be necessary to estimate movement parameters if processes such as tidal surfing affect movement.

Notwithstanding that alternative models of movement could be developed, the DSST has indicated that it does not wish to support a substantial expansion of the scope of work.

Currently the research plan includes a limited comparison of the statistical model to simulated data. The Panel suggests the project would benefit from an expansion of this type of comparison. Furthermore, the Panel suggests the work be expanded in consultation with regional experts studying the biology and behavior of Delta smelt as well as with experts on Delta smelt modeling.

See response to the third comment above.

We agree about the benefits of consultation with other experts. We already are collaborating regularly with the entrainment investigators and with Ken Newman's research team.

The Panel suggests the occupancy modeling be given low priority. It was not made clear how this work would lead to characterizing Delta smelt abundance or the relationship of abundance with environmental covariates related to fall outflow.

If we find that the inference that can be obtained from the simulation model of abundance is limited, the occupancy model will allow us to work with the fall midwater trawl catch data in a more robust manner. That is, the fall midwater trawl data may be sufficient for estimating occurrence (presence / absence) and colonization of or persistence in relatively small regions, but not for estimating abundance. A drawback of the occupancy model is the inability to test whether there are relations between survival and outflow. Instead, the occupancy model would focus on relations between spatial patterns in habitat use and outflow.

The Panel suggests the Investigator Team evaluates the value of including predators in the study due to the scarcity of Delta smelt in diets of predators.

The DSST has asked us to retain the proposed work on predation. We respond to comments on the proposed methods below.

Brief summaries of responses to the 5 Panel charge questions are listed below.

1. Are goals and objectives presented in the proposal clearly articulated, internally consistent, and are they adequate to address the questions outlined in the Scoping Document?

The Panel agrees the goals and objectives of the proposal are relatively well articulated and internally consistent but deviate somewhat from the questions outlined in the scoping document.

We commented above on the panel's interpretation of the scope of work.

The Scoping Document emphasizes the interconnectedness of fall outflow, other environmental covariates, and survival/recruitment of Delta smelt, whereas the proposal focuses on a few other processes (movement, occupancy, colonization/extinction), but with much less development of how environmental conditions, X2 and especially fall outflow, are correlated with survival and abundance. The Panel suggests that neither the data nor the analysis outlined may be adequate to answer the questions in the Scoping Document and emphasis of parts of the proposal should be revaluated.

We agree that the data may not be adequate to achieve all of the objectives in the scope of work. We proposed an analysis of occupancy because we are concerned that the data may not allow strong inference on survival and recruitment. We proposed to build simulation models to determine how accurately our model can estimate the coefficients relating covariates to survival and movement given different probabilities of detection and sample sizes.

2. Does the proposed work have strong potential to contribute to an improved understanding of Delta smelt survival through the fall and winter-spring recruitment?

The Panel was in agreement that the approach outlined by the investigators is sensible and fairly general regarding treatment of the count data. However, there were doubts about whether the data are sufficient to model accurate and reasonably precise survival estimates.

We have similar doubts. See response to the previous comment by the panel.

3. Does the proposed work using zero-inflated models, region designation, and occupancy modeling represent a feasible and robust approach to addressing questions such as detection probability, occupancy, and collinearity among covariates?

The proposed methods all appear adequate and represent well accepted modeling techniques for the most part, however, there are some concerns based on mismatch between modeling assumptions and this particular study system. The Panel has questions about the utility of the occupancy modeling in relating the fall outflow with Delta smelt recruitment and survival.

We did not propose to use occupancy modeling to address recruitment and survival. We proposed occupancy modeling as a means of examining whether environmental covariates are associated with spatially explicit probabilities of presence, to extend definitions of habitat, and to account for imperfect detection.

4. Is the proposed expert elicitation approach precedented, sound, feasible and realistic; and does it have the potential to contribute to understanding of the Delta smelt population?

The method proposed is a modified version of the roundtable discussion combined with the anonymity (presumably) of the Delphi Technique.

This characterization of the proposed method is not quite accurate; see below.

The proposed method does take advantage of expert opinion; however, it is possible that a strong personality in the discussion can affect the outcome so it will be up to the organizer to assure that no single opinion outweighs the "minority" opinions. The elicitation needs to include members of other research groups studying Delta smelt. These should include those developing life cycle models and those studying the behavior and movement of Delta smelt in their local environment.

We agree; see below.

5. Is the proposal explicit in what data will be used and how limitations of the data will be addressed in relation to the questions being asked?

In consideration of the objectives stated in the scoping document, the proposal is clear in what data will be used; those data which are provided over the historical period from 1990 to 2014, collected in conjunction with trawls. However, there seems to be little *a priori* consideration of the limitations of the data and their ability to account for the majority of the variation observed in the existing data. Indeed, there appears to be little allocated time in the proposal budget to investigate the limitations of the data or development of covariates for inclusion in the analyses. The work allocation should be modified to allow for such analysis.

As described above, the potential limitations of the data are one of the principal reasons why we proposed to conduct an analysis of occupancy. Also, as described above, we proposed simulation modeling to evaluate how accurately our model can estimate the coefficients relating covariates to survival and movement given different probabilities of detection and sample sizes.

We agree with the panel that we did not account sufficiently for the effort that likely will be necessary to fully understand and format the available data on response variables and covariates, and will structure our timetable and financial budget accordingly.

Responses to questions

1. Are goals, objectives, hypotheses and questions clearly articulated and internally consistent and are they adequate to address the questions outlined in the Scoping Document?

General comments

The Panel agrees the goals and objectives of the proposal are clearly articulated and internally consistent but deviate somewhat from the questions outlined in the scoping document.

> As explained above and below, we believe that the panel may have mischaracterized the scope of work.

The scoping document (SD) emphasizes the interconnectedness of fall outflow, other environmental covariates, and survival/recruitment of Delta smelt, whereas the proposal focus on a few other processes (movement, occupancy, colonization/extinction), but with much less development of how environmental conditions, X2 and especially fall outflow, are correlated.

The Panel notes the Scoping Document requesting that the proposal provide information to infer how covariates strongly correlated with survival and recruitment. Four specific correlation analyses [were] identified as questions:

Scoping Document question 1) correlations of all variables with survival and recruitment,

Scoping Document question 2) correlation of fall outflow with survival and recruitment,

Scoping Document question 3) correlations of variables except fall outflow with survival and recruitment,

Scoping Document question 4) correlations of fall outflow and the other variables. Survival and recruitment will be estimated using catch data using the three seasonal fish surveys (summer townet, fall midwater trawl and spring kodiak trawl surveys).

> We feel that this characterization of the scope of work is not accurate. The only correlation analyses identified in the scope of work are analyses of the extent to which covariates are correlated. The scope of work does not necessarily identify four analyses, but four topics on which the analyses are expected to provide inference.

The Panel had doubts that the available data are adequate to make the predictions and correlations needed to identify which environmental variables are strongly correlated with annual changes in fall survival and recruitment. A number of issues were raised concerning adequacy.

> We share these doubts, and these concerns in part motivated our proposal to analyze occupancy. We proposed simulation modeling to evaluate how accurately our model can estimate the coefficients relating covariates to survival and movement given different

probabilities of detection and sample sizes.

Spatial coverage and temporal variability of data

To understand the survival and recruitment dynamics and their environmental covariates, the authors propose to use a suite of local and global (regional) covariates as potential predictors of variation in various state (occupancy and abundance) and process parameters (survival, movement, colonization and extinction). Inferences about survival will come mainly from a model that combines elements of state-space and Dail-Madsen modeling. Recruitment will be assessed primarily through a stock-assessment approach. One concern of the Panel is the potential mismatch between the underlying scale-dependent dynamics controlling Delta smelt survival and recruitment and the spatial and temporal scales of the data.

We agree that the available data on response variables and covariates may not be sufficient to draw strong inference to drivers of survival and recruitment.

Given the substantial data needs that are required by the models, the Panel felt that they were not sufficiently informed about the types and structure of the available data. This was a problem because there was a sense that the data were not of sufficient quality or quantity to perform the modeling exercise. It was not clear from the conference call whether these data were "in hand" by the authors – a discussion of the ability of the model to handle the existing data would have been valuable (see simulation issues). The relevance of this is that the data and modeling will necessarily limit the inferences that can be made. Some outstanding issues then are to what extent the data are capable of informing the model and to what extent the data structure will limit inference.

We agree, but it is impossible to fully vet data to which we do not yet have access.

Simulations

The proposed simulation study is an excellent idea, and can shed light on the ability of the current data to address questions of concern to the Delta Smelt Scoping Team. The purpose of the simulation is to determine how accurately the model can estimate coefficients relating environmental covariates to survival and movement given different probabilities of detection and sample sizes. A number of issues were raised regarding the simulations and their applications.

Selection of flows – Ultimately, the simulation will rely upon historical flow / outflow records to predict the distribution and abundance of Delta smelt.

The simulation will draw from records of flow or outflow and from data from other sources or on other covariates.

This may be appropriate to "test" the simulation. However, the appropriateness of the simulation going forward, considering the effects of climate change and associated impacts on flow, may not make it an appropriate model without adjustments to the predictive protocols. Recent research suggests that large-scale climate oscillations (AMO, PDO, and ENSO) do not influence continental weather patterns in the manner that has been recorded over the past one hundred years. In some areas, additional flooding might be expected. However, it appears that historically low and high flow patterns will now be much reduced. As a result, the low outflow simulation may not be appropriate to predict the distribution of Delta smelt. The assumption of a continuous connection between high quality habitat patches may not be valid under lower than "historical low" flows.

> The maximum flows into the ecosystem well may be reduced if snowpack decreases and snowmelt occurs earlier in the year. However, given the extent to which the ecosystem is managed, it may be difficult to project relations between future climate and hydrology. Additionally, we were not charged with projecting future survival and recruitment, but explaining observations of past survival and recruitment.

The Panel suggests that the proposal Investigator Team considers an even lower flow scenario that reflects this potential flow regime in the Delta. The flow regime is dependent upon releases from upstream reservoirs but these are likely to become impoverished as agricultural demands increase and as low flow / drought conditions continue or become the new "historical" condition. The discontinuity in patch connections may dramatically alter predictions of the model. This comment is emphasized by a large number of simulations (for example, Maurer 2007 and Xenopoulos et al. 2005) which predict dramatic reductions in flow and loss of freshwater fish species (up to 37% losses from the Sacramento, in some scenarios) by 2070 if climate- change induced changes in river flow patterns in the Sierra Nevada continue. An alternative would be to run simulations with varying degrees of connectivity between habitat patches.

> Again, we were not charged with projecting future survival and recruitment, but explaining observations of past survival and recruitment. If we are able to explain substantial variation in past survival and recruitment, then it may be worthwhile to expand the proposed work to future projections.

Movement models

The goal of this simulation is to assess the bias in assumptions of capture probability. The Panel suggests that it is equally important is to assess biases in the movement model. In this regard the Panel has concerns for the adequacy of the movement model.

Recent studies suggest movements are driven by local (channel width and tidal cycle) scales of asymmetry in turbidity that induces intermittent "tidal surfing" (Bennett and Burau 2014). The occurrence and strength of this movement mechanism varies within the Delta and is likely to vary over the season (Feyrer 2013). Tidal surfing may concentrate Delta smelt in the X2 region in the summer (Kimmerer et al. 2014) but allow upstream migration in the fall. The Panel does not see how the wide-scale horizontal attraction model of movement equations (5) and (6) can be related to this emerging understanding of the effect of local process on movement. Note that equation (6) is incomplete: W_{rt} and $kappa_{rt}$ are not defined.

> W_{rt} is the vector of physical driver variables affecting habitat quality and $kappa_{rt}$ are the coefficients that define the strength of the relations between the physical covariates and

habitat quality.

The Panel believes that development of a movement model that reflect the local scale dynamics is beyond the scope of the study; the Panel recommends consideration of some revision of the model that takes into account the recent studies of Delta smelt movement dynamics. The challenge is to revise the current movement model to capture some issues affecting fall flows on fish survival and recruitment.

> The panel seems to be suggesting both that a more-detailed movement model is desirable and that a detailed movement model is outside the scope of work. We are uncertain what the panel is asking of us. Nevertheless, we proposed to model survival, movement, and recruitment as separate response variables. Both movement and survival then affect abundance. The technical feasibility of modeling movement at fine temporal resolution is addressed above, as is the DSST's preference that we not substantially expand the scope of work.

The Delta is an interconnected system with major inputs flows from the Sacramento and San Joaquin rivers and exports through the pumps in the southern Delta. While these flows set the mean water movement, the Delta is dominated by tidal flow which, coupled with the complex geometry, creates tidally varying regimes of velocity, turbidity and salinity. For example during the fall flush, the rate of movement of turbidity from the Sacramento River depends on the rate of export of water from the sound Delta pumps. Delta smelt movement is influenced by the tidally varying distributions of turbidity gradients over scales of a few meters (Bennett and Bauer 2014).

The Panel believes that to understand the effects of fall flows on Delta smelt populations, details of the interactions of these flow elements will need to be addressed.

> See comments above.

The proposed movement model assumes that movement is driven by the horizontal differential of region-wide measures of habitat quality and distance of separation. In the model, dispersal probability is proportional to the Euclidean distance between regions (centroid-to-centroid distances) (equation 7). This approach may be adequate for modeling movement in a relatively linear system, such as the Hudson River. Euclidean distance between regional centroids is unlikely to accurately represent true distances (i.e., the distance 'as the fish swims'). Are regional centroids defined geographically without regard to sampling intensity, or will a weighting scheme be applied (e.g., number and location of stations in the region)?

> We are not proposing to weight values of the variables.

In particular, dispersal probabilities from the North and South regions may not be adequately represented by Euclidean distances between their centroids. The distance between the Northern and Southern centroids is likely to underestimate the actual distance that a fish would need to swim to move between these two regions. If the bias is constant among regions, than the effect would not be concerning. However, the bias in the use of Euclidean distances between regional centroids is likely to be inconsistent among regions. A more direct approach would be to estimate distances between all adjacent sampling stations (e.g., using their GPS coordinates) and combining these to estimate dispersal probabilities from any given station in a particular region to the nearest station in the adjacent region. Other similar approaches could be considered.

> It is not clear to us how the number of stations might affect the modeled state process of movement. Additionally, we cannot use stations as the sampling unit because surveys were not replicated within each time step.

Another approach to improve estimation of dispersal probabilities is to model the relationship between dispersal probability and time. Such an approach would acknowledge the biology of the Delta smelt and would lend additional realism to the model. For example, fish may be more likely to move prior to or around the time of spawning (spring); thus, dispersal probabilities in December should be higher than those in September. This type of response may be modeled as time- (month) dependent or age- (fish size) dependent.

> We agree that these relations could be modeled, and are happy to do so if the DSST concurs.

The Panel encourages the Investigator Team to consult with groups doing this work and incorporate suitable functional models and identify relevant spatial scales for their simulation to explore the bias in abundance and survival estimates. The Panel sees this as critical to address the Scoping Document questions because fall flows affect the turbidity distribution and tidal dynamics. Ideally the work could be facilitated by collaboration with the other research teams being funded as part of CAMT Delta smelt research.

> We agree about the benefits of collaboration, and are working with Ken Newman's research group and the entrainment team. We also have access to extensive expertise via the DSST and CAMT.

To assess whether the movement model is adequate to estimate survival, the Panel recommends a model comparison exercise where pseudo-data generated with a more complex individual based model, e.g. Rose et al. (2013), are analyzed with the abundance model outlined in the proposal. In this way ability of the analysis framework and selections of spatial/temporal scales are able to recover meaningful scale adjusted estimates of the underlying movement and survival processes used to generate the IBM data.

> See response to the third comment in the Executive Summary.

Linking model outcomes to Scoping Document

It is not clear how the Investigator Team proposes to integrate the individual model outcomes (abundance, movement, survival, occupancy, detectability, and expert elicitation) to address questions in the Scoping Document.

> In general, we are not proposing to integrate the outcomes mathematically. The expert

elicitation will generate a distribution of values related to predation that may serve as a covariate in multiple models.

Equation 4 will be used to determine factors affecting survival in fall (SD question 1), and equations 11 and 13 (page 17) will be used to characterize habitat quality for Delta smelt in the fall (SD question 3). However, it is unclear how SD question 2 concerning the relationship between abundance of Delta smelt in December and larval production in the spring (recruitment) will be addressed.

The proposal spends very little time addressing the recruitment analysis (mainly on pages 15 and 16). However, this appears to be an important part of the Scoping Document (questions 2 and 3). Although an understanding of correlations between predictors is a large part of the Scoping Document (questions 1 and 4) it is barely covered in the proposal.

Assessing correlations among covariates is highly relevant to the scope of work, but mathematically is relatively straightforward.

Objective 4 (page 4) implies that the team will 'compare model-based estimates of abundance in December with recruitment in the spring', but the proposal does not describe in detail how recruitment in the spring will be measured or modeled.

There is no consensus on a single response variable that best represents recruitment. The DSST and CAMT wish to understand whether autumn survival of delta smelt and environmental attributes in autumn are associated with recruitment of delta smelt in spring, but also wish to avoid confounding analyses of recruitment with survival of larvae, which could reflect other environmental attributes or environmental attributes in spring (and is outside the scope of work).

Therefore, we proposed to conduct a stock-recruitment analysis. Several metrics derived from the modeled abundance in of delta smelt in December could be used as the stock, such as the estimates of abundance in December and a spatial diversity metric reflecting the distribution of delta smelt in December.

As noted in the proposal, our discussions with the DSST identified three potential metrics of recruitment. The first is abundance estimates from the spring kodiak trawl. In this case, the underlying state variable is winter abundance (i.e., abundance of potential spawners). The second potential response variable is the length of delta smelt, which also would be derived from fish captured by the spring kodiak trawl. Use of this response variable reflects implicit hypotheses that length is related to body weight, body weight is related to fecundity, and fecundity is related to recruitment. The third potential response variable is abundance as derived from the 20 mm survey. In this case, the underlying state variable is abundance of larvae (i.e., production).

The text on page 20 suggests that occupancy model outputs for December will be compared with distribution of recruits in April; however, the output from occupancy models includes detection probabilities and estimates of the proportion of sites occupied, not abundances. It is unclear how occupancy model outputs will be used to estimate "abundance."

> The comparison refers to spatial distributions only. The outputs of occupancy models cannot be used to estimate abundance.

2. Does the proposed work have strong potential to contribute to an improved understanding of Delta smelt survival through the fall and winter-spring recruitment?

General comments

The Panel was in agreement that the approach outlined by the investigators is sensible and fairly general regarding treatment of the count data. However, there were doubts about whether the data will allow for a complex model that leads to believable and reasonably precise survival estimates.

> We agree, which is why we proposed an analysis of occupancy.

Very little is known about the environmental drivers that contribute to survivorship in Delta smelt. Indeed, based on the literature presented, the stock exhibits a wide degree of phenotypic variation in demographic and behavioral characteristics. Previous work has indicated that regional (spatial) designation has little predictive power in understanding the likelihood of capturing Delta smelt.

> Likelihood of capture (= detection) has not previously been estimated. Instead, previous work may have indicated that apparent occupancy does not have a strong spatial dependence

Modeling studies suggest that survivorship and recruitment are affected by multiple factors, working at small scales (Rose et al 2013). However, the Panel believes that the multifaceted effort proposed by the proposal authors is likely to make a meaningful contribution to understanding the mortality dynamics of the Delta smelt stock. The major contribution may be in identifying the limitations of the data to make such inferences and make substantial management recommendations to improve sampling efforts, such that the mortality estimates can be improved in the future.

The value of the approach will be contingent on the power of covariates to describe survival and the statistical power to detect meaningfully variations. At a minimum, the Investigator Team efforts to understand how sensitive mortality is to each covariate will be informative and the simulations will hopefully illustrate the power (or lack of power) based on current sampling design. An understanding of the relative sensitivity of state variables to environmental drivers and the functional forms of the linkages will provide valuable information to help guide monitoring practices. This represents a valuable contribution to the management of Delta smelt, especially given the many ambiguities in understanding the behavior and demographic characteristics of the species.

In particular, the modelers may need to constrain the movement model based on recent information on fish movement (see movement models above). In addition, the degree to which capture probability in the model conflates use and detection processes in the real world, as well as the implications of this, is unclear. Simulations and use of posterior predictive checks may help in identifying this and other issues associated with the goodness of fit.

In particular, the modelers may need to constrain the movement model based on recent information on fish movement (see movement models above). In addition, the degree to which capture probability in the model conflates use and detection processes in the real world, as well as the implications of this, is unclear. Simulations and use of posterior predictive checks may help in identifying this and other issues associated with the goodness of fit.

> See comments above about movement models

The proposed work will likely contribute to an improved understanding of Delta smelt dynamics particularly given that the Delta Smelt Scoping Team requested that the researchers use currently existing data only. This directly affects the level of uncertainty in the outcome of the proposed modeling work. The Investigator Team proposes to apply new modeling approaches and as such, this represents a positive, important step forward.

> The use of Bayesian methods, particularly in the simulation, allows comparisons of the degree of uncertainty in covariate effects as a function of sample size

The idea of consulting and working with the members of the Delta Smelt Scoping Team and the Collaborative Adaptive Management Team to frame hypotheses about attributes of habitat quality is excellent and should be pursued.

As an aside, the Panel had an internal discussion on definitions of detectability and catchability which are included here. To detect a species, at least one individual must occur at the site, and, as expected, the probability of detection increases when a greater number of individuals are present. Thus, detectability is a function of the number of organisms that are vulnerable to capture and the probability of capture. For a given level of abundance, researchers can maximize detectability by ensuring high catchability which entails the deployment of efficient sampling gear during times and in locations where the species is available (Bayley and Peterson 2001). Catchability, or the proportion of a fish stock captured with a single unit of effort (Gulland 1983), is the product of availability and gear efficiency (Kimura and Somerton 2006). Availability refers to the proportion of the fish stock that occurs in locations where the gear is deployed, and gear efficiency is the proportion of fish captured from those that occurred within the sampled area (Kimura and Somerton 2006). Although often assumed constant, catchability is temporally and spatially variable because availability and efficiency vary. For example, gear efficiency may be affected by environmental factors that alter gear performance and by fish behavior, as well as, by the selectivity of the gear and the vulnerability of individual fish. It is important to clearly differentiate detectability, availability, catchability, and gear efficiency when discussing sampling concerns for Delta smelt (Williams and Fabrizio 2011).

> Consistent with the literature on occupancy modeling, we are defining detection probability solely as the probability of observing an individual given it is present. Gear

efficiency affects detection probability. Abundance may affect detection probability. One also can include spatial and temporal terms, which generally serve as proxies for spatial and temporal variation in environmental attributes, as covariates in a model of detection probability.

Abundance modeling

The research team will model Delta smelt abundance and survival during fall using catch data from a time series of fishery-independent surveys. Two surveys provide such information: the summer townet (STN) survey (August), and the fall midwater trawl survey (FMWT, September through December). In addition to using different gears, the two surveys sample at different fixed stations, and the number of stations sampled varies: the FMWT survey samples 100 stations annually compared with 31 stations in the STN survey. How will abundance data (catch per effort) from the STN and the FMWT surveys be used in the same analysis? For example, the STN data will be used to initialize the abundance model (p. 10), but the two surveys are characterized by differences in the spatial intensity of sampling.

Although the spatial intensity of sampling differs, we are assuming that each sampling method provides information on regional abundance. Thus, it is not critical that sampling intensity is identical in space, but rather that the samples taken at the stations represent the abundance in the region during the month of sampling.

In the observation process model, the number of individuals observed is modeled as a binomial random variable, with regional abundances estimated from the FMWT survey and the probability of capture estimated as a function of covariates. In this model, in addition to consideration of the main effect of each covariate (which were not specified here, but presumably include flow, temperature, salinity, turbidity, etc.), the Panel recommends consideration of the interactions between covariates. For example, the probability of capture may increase with increasing temperature in low turbidity, but may not change with temperature at high turbidity. This will allow for more nuanced understanding of covariate effects.

Nothing in the current proposal precludes examination of interactions. However, when data on response variables are relatively limited, as is the case here, it is best practice to constrain the number of covariates.

The section on ‘Modeling multiple cohorts with the abundance model’ proposes to estimate overall cohort survival using regional estimates of abundances in December and September ($N_{r, Dec}$ and $N_{r, Sep}$). In the previous section, monthly regional abundances ($N_{r, Dec}$ and $N_{r, Sep}$) were said to be derived from the FMWT survey. These estimates rely strongly on the assumptions that all fish in a given region are available during each survey in a region (one assumption in N-mixture models). Clearly, estimates of detection probability will be dependent on both use and detection, and what is unclear is how well the N- mixture approach will work in this setting, both because of additional heterogeneity and because the product of the probability of use and detection is likely to be exceedingly small. Given the potential sparseness of the data, it is also unclear how much power there will be to detect variation in detection probability.

The simulation model could be useful for understanding the scope of inferences that can be drawn from the existing data.

Regarding the estimation of annual recruitment of Delta smelt (page 14), the 'stock- recruitment analysis' proposed by the Investigator Team is quite vague. Which models will be considered for this analysis? The traditional stock-recruitment models such as those due to Ricker or Beverton and Holt may not be useful to address stock-recruitment dynamics in Delta smelt. These models, which describe compensatory mortality of pre-recruits, often account for only a small portion (5-15%) of the total variation in recruitment of fishes (see recent papers by Cury et al. 2014 and Szuwalski et al. 2014). The assumptions and limitations of the Ricker and Beverton-Holt models should be fully considered prior to their application. In particular, the assumption that 'stock' and 'recruitment' are appropriately measured using FMWT abundances in winter (stock) and the number of potential spawners in the spring trawl survey (recruitment) requires careful consideration. Examination of these two abundance estimates appears to provide a good estimate of survival, but it is not clear how the stock-recruitment relationship would be described. The panel recommends using the abundance of larvae in the spring from the 20 mm survey as an index of 'recruitment'. The measure of 'stock' size in stock-recruitment models should represent reproductive output (total number of eggs produced is best, but failing this, other estimates such as spawning stock biomass or mature female biomass have been used). Finally, fitting the stock-recruitment models of Beverton-Holt and Ricker to fisheries data requires an adequate range of values for 'stock' size. Given the declining abundance of Delta smelt, it is not likely that a sufficiently broad range of values for stock size will be available. Due to low abundances of Delta smelt in recent decades, simpler models (e.g., a linear model) may be sufficient to describe the relationship between spawners and recruits. At these lower 'stock' sizes, recruitment may, indeed, depend on stock size (Szuwalski et al. 2014). Unfortunately, at low population sizes (such as those observed for Delta smelt), depensatory mortality may need to be considered, assuming depensation mechanisms can be described for this species.

We concur with the recommendation that abundance of larvae (production), derived from the 20 mm survey, be used as a measure of recruitment. One of the reasons for using data that extend back to the 1970s is to draw inference from periods during which abundance was higher, and the probability of density-dependent effects greater. We agree that the recent history is unlikely to provide much information on density dependence unless the habitat quality is so low that abundance reflects density dependence. Therefore, linear models likely will suffice. When abundances are low relative to carrying capacity, the Beverton-Holt function is linear. Therefore, the Beverton-Holt function can be used to reflect both linear dynamic (by setting the carrying capacity to extremely large values) and density-dependent dynamics.

3. Does the proposed work using zero-inflated models, region designation, and occupancy modeling represent a feasible and robust approach to addressing questions such as detection probability, occupancy, and collinearity among covariates?

General comments

The proposed methods all appear adequate and represent well accepted modeling techniques for the most part, however, there are some concerns based on mismatch between modeling assumptions and this particular study system. The Panel has questions about the utility of the occupancy modeling in relating the fall outflow with Delta smelt recruitment and survival.

> Estimates of occupancy do not directly inform estimates of survival.

Regional designations

The region designation is based on previously established regions within the Delta and the authors' plan, because of constraints in the spatial distribution of the samples, to delineate only four regions.

> Data likely are not sufficient to delineate more than four regions for models of abundance. For our assessment of habitat quality, we expect to divide the Delta into a greater number of regions among which there is spatial and temporal variability in the presence of delta smelt.

The regions were delineated based on previous work but it is not clear how consistent the spatial covariates (at local or regional scales) are within each region. We understand that the authors do not have the abundance data for Delta smelt at the time of writing this proposal, but it is difficult to justify the regional designations without a discussion about the coherence of the catch-per-unit effort or some metric of abundance of Delta smelt or the coherence of the habitat or environmental covariates. This raises the question "to what extent will the results of modeling work be sensitive to these designations?". The Panel recommends that the Investigator Team performs a *post hoc* sensitivity analysis to address the issue of regional designations.

> It is possible that abundance varies in space within a given region, but is temporally consistent. In this case, the spatial samples will not accurately represent the abundance in the region. It also is possible that the environmental covariates that we assume to be consistent across the region actually vary in space, and that such variation affects the distribution or population dynamics of delta smelt.

One outstanding issue is how the covariates (local and regional) will be synthesized. The Investigator Team mentions the range in spatial and temporal scale that the covariates exhibit. For example, some covariates exhibit high-frequency and fine spatial-scale variability (salinity, turbidity, water temperature) and others perhaps less so (volume of tidal marsh, habitat quality). It is not clear to what extent a covariate that exhibits high- frequency and fine spatial-scale variability is representative of a large spatial region. These are not insurmountable or novel issues but it would be informative to understand how the Investigator Team intends to deal with these.

> Our options for modeling these types of covariates are somewhat limited. If there is considerable environmental variability at high frequency and fine spatial resolution, the relation between environmental variability and detection probability only can be assessed if that variability was measured simultaneous with surveys. Alternatively, the mean, minimum, or maximum environmental variability can be estimated over longer periods of

time (e.g., one month) and then incorporated into the functions we proposed to describe the population dynamics.

On page 6, the Investigator Team states that the data are not likely to be sufficient to delineate more than 4 regions for models of abundance. Yet, for estimation of habitat quality (using occupancy models), more than 4 regions may be considered. In terms of abundance estimation, because the STN estimates will be used to initiate the abundance model, the spatial resolution will be limited to that which is supported by the 31 survey sites; 4 regions seems reasonable here. The spatial resolution of the data for occupancy modeling will allow delineation of more than 4 regions because the FMWT survey comprises 100 (index) sampling sites. Given the likelihood that detection probabilities will be low, and the fact that spatial replicates will be used for occupancy sampling, the Panel recommends that 'patches' include no less than three sampling sites.

We agree.

Delineation of these patches on the basis of hydrology seems reasonable and may result in 'patches' with more than three sampling sites.

We agree.

It would be helpful to know how sampling sites from the FMWT survey align with different regional delineations being considered.

We can overlay the stations on the boundaries of the regions used by Newman et al. and the other spatial delineations that were mentioned in the proposal. We have not yet defined regions on the basis of hydrology.

Collinearity among covariates

The issues of collinearity among candidate covariates were discussed on the conference call and will be treated in two alternative ways – one of which was a departure from the methods averred in the proposal: covariates will be treated outside of the model to inspect for collinearity and only those that are considered to add to the explanatory power of models would be included into candidate models. Collinear factors were to be treated in a data reduction statistical treatment using non-metric multidimensional scaling and incorporated into models.

We will use correlation analyses to assess collinearity among the covariates. Variables typically are considered to be correlated strongly if variance inflation factors are > 10.0 or correlation coefficients are > 0.60 (Neter et al. 1996). If we find strong correlations, then we will use ordination to construct orthogonal projections across the covariate data to capture the collinearity in the covariates. The best-known ordination method is principal components analysis, which transforms correlated variables into a set of linearly uncorrelated variables. Other ordination methods often are more appropriate for ecological data (Pielou 1984), and we will evaluate several alternatives if we find evidence of such collinearity.

The Investigator Team will consider the effect of multiple covariates on occupancy and detectability. During the review Panel's interaction with the Investigator Team, one of the researchers indicated that centering and standardizing of covariates will be considered. Although useful, the inclusion of multiple covariates and especially inclusion of interaction terms, will likely cause collinearity of candidate predictor variables, regardless of whether the variable is re-scaled.

> We did not suggest standardization as a means to address collinearity. Standardization of covariates is best practice regardless of collinearity.

Collinearity can be addressed in three ways, the first is to assess the pairwise correlation of predictor variables and the second is to examine variance inflation factor, *post hoc*. If necessary, the Investigator Team will use a multivariate analysis for data reduction. Such reduction includes using a Principal Component Analysis (PCA) or Non-metric multidimensional scaling (NMS), whereby multiple covariates are transformed into a single factor.

The Panel recommends the following steps:

1. Focus on a parsimonious set of management-related covariates.

> Not all covariates we are likely to examine can be affected by management, but we feel it is worthwhile to explore the extent to which management could affect covariates that are associated with any response variable.

2. If data reduction is necessary, the collinear variables should be limited to sets of variables that are interpretable (e.g., a PC vector based on multiple covariates that measure an overall regional climate signal, or an overall signal for 'flow').

3. If these approaches fail, the Panel recommends elimination of covariates because interpretation of the effect of the derived covariate (e.g. the 1st principal component) will be difficult.

The Panel also recommends thinking about correlated covariates (especially when correlation is between fall outflow and some other covariate of less management concern) in a decision support context and multimodal inference context. What would happen if the response was driven solely by one covariate vs. the other or at intermediate levels?

> Correlated covariates can be included in separate models and information theory applied to assess the level of support in the data for each model.

Occupancy modeling and detection probability

The proposed study will use spatial replicates to estimate probability of detection and occupancy; presumably, the spatial replicates are the multiple stations sampled within each patch.

Yes, that is correct.

Application of occupancy models requires the assumption of closure at the patch level which spatially encompasses multiple stations. Although a recent paper by Kendall and others relaxes the closure assumption, "the power to detect the violation of closure is high if detection probability is reasonably high" (Kendall et al. 2013). Detection probabilities for Delta smelt are unknown, but are not likely to be high, thus, it may be difficult to know if the closure assumption will be acceptable for these data. The relaxation of the closure assumption allows organisms to move seasonally between patches, which could describe Delta smelt movements in the Delta. However, spatial replication will need to be fairly high (a minimum of 3 possibly more sampling stations per 'patch') to provide 18 useful estimates of occupancy and detectability (see Table 6.1 in MacKenzie et al. 2006). How the stations are assigned to a patch is a critical step in the modeling process. That is, how are the 'patches' operationally defined in this study? The Panel recommends discussions with the agencies/scientists that collected the survey data to determine the best possible assignment of stations to patches, and patches to regions. The level of resolution of the patches appears to be a critical question that will affect interpretation of results as well as the selection and use of covariates.

As noted in the proposal, there are multiple options for defining regions (patches). For the sake of complementarity, we suggested that one of our delineations be that used by Newman et al. to construct a life-cycle model for delta smelt. Another option is to divide each of the four regions differentiated by Newman into two regions. Alternatively, regions could be delineated on the basis of hydrology and geomorphology.

Time-of-day of sampling is a relevant factor for occupancy models because Delta smelt may avoid areas with high light levels. Given this behavior, Delta smelt are not likely to be present in areas when light levels are high (sunny afternoons, clear water); therefore, time of day is a potentially important covariate for detectability. However, because the effect of 'time of day' may vary depending on cloud cover, the Panel recommends consideration of other covariates such as the interaction of depth and turbidity, which may be more informative.

We agree that time of day or light level is a relevant covariate for estimates of detection probability, but it is not clear to us how the interaction between turbidity and depth would provide information on daylight and cloud cover. It may be possible to use archived data from weather stations (whether measured on the ground or remotely sensed) to assess cloud cover at the time of sampling. However, the latter likely would be quite labor intensive.

On page 18, the research team describes the use of the summer townet data (August sampling) to determine initial occupancy states. However, there are only 31 stations sampled by the STN survey, so the level of resolution will be different from that of the FMWT, which samples at 100 stations.

We agree that the distribution of summer townet stations is a consideration in delineating regions. If we require each patch to encompass ≥ 3 stations, then the maximum number of patches will be 10.

The ‘patches’ for the STN survey data will be geographically larger than those for the FMWT. It is not clear how the less spatially intensive STN can be used to initialize occupancy states for the FMWT data.

> The size of the patches will not differ, although the number of samples within each patch taken by the summer townet survey and the fall midwater trawl survey will differ.

It is unlikely that colonization/extinction rates will be able to be estimated at the station level – particularly because ‘stations’ will be used as the spatial replicates (of the ‘patch’).

> We proposed to estimate colonization and extinction at the region level, not the station level.

The proposed use of cross-validation of the occupancy model parameters may not be viable – the samples appear to be too limited.

> We can partition the stations into training and validation sets because we are essentially running the model on data for each year with the assumption that the coefficients do not change among years.

Multi-state occupancy models may be worth considering to better understand detectability and occupancy of Delta smelt. Although multi-state occupancy models are discussed, it is not clear if the research team intends to apply such models. One difficulty will be determining the abundance ‘bins’ (low, medium, high abundance) to describe the multiple states. It may be useful to consider non-linear thresholds for the bins – for example, low (0- 1 fish/tow), medium (2-10 fish/tow); high (greater than 10 fish/tow).

> Whether we apply multi-state occupancy models will depend on the outcome of the proposed analyses of abundance and occupancy. If we can model occupancy, but not abundance (e.g., if we discover that the precision of the abundance data is insufficient for the proposed abundance models), then it may be worthwhile to explore multi-state occupancy models. Additionally, we can explore different options for defining abundance states.

As mentioned above, the Panel is concerned about patterns of use within regions and how this process will interact with underlying variation in detection rates. These concerns result from the use of spatial replicates for both occupancy and abundance models. In the occupancy modeling, it may be possible to define regions at smaller spatial scales, potentially removing some of the heterogeneity in use that may occur in the larger regions. However, this may also lead to only a few replicates per region. In addition, both the use of the spatial replicates (Kendall & White, 2009) and the low number of regions and replicates raise concerns about bias in estimates of occupancy, colonization and extinction.

> We have similar concerns, but the available data do not give us other options.

Abundance distributions

Multiple options were considered for modeling abundance (p. 10) including the Poisson, negative binomial, zero-inflated, and mixture models. One type of model that is commonly used in fisheries is the delta-lognormal model (as well as other delta models). The approach of the delta model seems to be described on page 15, but is not identified as such.

> We are not familiar with the term delta-lognormal model, so did not label it as such.

It is important to recognize the multiple approaches that can be used in modeling abundance data from survey data, but there is no indication of which approach will be applied by the research team. In the section titled 'Research Challenges', the discussion in the text appears to favor use of the negative binomial and zero-inflated negative binomial models for estimating abundance. Will each of these approaches be used or will the determination of approach be made using an objective criterion or some other method? The Investigator Team highlights the trade-off between methods that account for over- dispersion versus those that model processes that can lead to over-dispersion (e.g., the two-component Poisson). However, the team does not specify which approach will be used. Questions concerning the number of parameters and the principle of parsimony could be addressed here, as well as the expected limitations concerning the ability to generalize results from two-component Poisson models to survey data collected in future years. The Panel recommends exploring the sensitivity of model outputs to different assumptions of abundance.

> We likely will begin with a distribution that has relatively high flexibility (e.g., the zero-inflated negative binomial). We then will attempt to use mixture methods to fit models that explain the zero-inflated portion of the distribution (i.e., to identify environmental conditions associated with zero catches) and the over-dispersed portion of the distribution (i.e., to identify environmental conditions associated with high catches).

4. Is the proposed expert elicitation approach precedented, sound, feasible and realistic; and does it have the potential to contribute to understanding of the Delta smelt population?

When it is not possible to obtain more field data to create habitat or response models of specific organisms, experts are consulted and, through a process of joint or independent conversation, create habitat or response models based upon a combined expert opinion.

> Conversation is not used to create a model. Instead, estimates that may be informed by conversation are used by the investigators to create a distribution of values of a parameter.
>
> Methods for the elicitation process itself, and for developing parameter distributions on the basis of an elicitation process, have been studied extensively by both statisticians and social scientists. The methods we are proposing are considerably more structured and rigorous than traditional application of the Delphi method. We are eliciting values of variables, not knowledge about habitat attributes or behavioral choice in a more general sense. The elicitation process does not yield one estimate but a distribution of estimates. We will provide more detail on the process in our full response.

Although the list is not complete, there does appear to be an error in the suggested list of predators to model. The proposal suggests that a Delphi evaluation will be created for *Ptychocheilus oregonensis* (northern pikeminnow) which, unless new records exist, is not likely to appear in this particular drainage. That species is found in the Columbia River drainage. Perhaps the authors meant *P. grandis* (Sacramento pikeminnow)?

> Yes, we did, and apologize for the error

More importantly, the Panel questions whether this is an appropriate exercise as part of the study? The gut-contents data that have been published or presented to Panel reviewing the Long-term Operations Biological Opinions (LOBO) seem to indicate that predator pressure on Delta smelt is quite small.

> We are uncertain that analysis of gut contents provides considerable information on predator pressure. Delta smelt may be evacuated from predators' guts fairly rapidly. Additionally, many individuals with whom we have consulted to date feel it is possible that predation is affecting survival and reproduction of delta smelt. Given the limited data on predation, we would prefer to evaluate the hypothesis rather than to reject it outright.

While the expert elicitation approach may provide information on covariates and their ranges, it is not clear how the process can be applied to evaluate and potentially revise the model functional forms according to input from experts.

> We did not propose to use elicitation to derive functional forms, only distributions of input variables. We will test whether relations among variables, whether derived on the basis of empirical data or on the basis of elicited values, are explained by a range of functional forms.

5. Is the proposal explicit in what data will be used and how limitations of the data will be addressed in relation to the questions being asked?

General comments

In consideration of the objectives stated in the Scoping Document, the proposal is clear in what data will be used; those data which are provided over the historical period from 1990 to 2014, collected in conjunction with mid-water trawls. However, there seems to be little a priori consideration of the limitations of the data and their ability to account for the majority of the variation observed in the existing data. Indeed, there appears to be little allocated time in the proposal budget to investigate the limitations of the data or development of covariates for inclusion in the analyses. The budget should be modified to allow for such analysis.

> See comments above and below. We will allocate more time and money to data development and formatting.

Descriptions of covariates

The description of covariates was limited, particularly for the station-specific covariates measured by the fisheries surveys (STN, FMWT). The spatial scale of the covariates is also not described. For example, do the single-point-in-time measures of salinity represent conditions at the sampling site only or will salinity measures be averaged at the patch level? How is 'flow' measured and considered in the models and at which spatial scale? Additional covariates (e.g., habitat type such as river, channel, etc.) may apply to multiple adjacent patches, but the scales of measurement and application need clarification. Studies have indicated that Delta smelt distributions and abundance are correlated with turbidity but not with annualized covariates of water quality or flow (Latour 2015).

> Values of the covariates must be defined at either the level of the stations (where fish are captured) or at the regional level. Even if a covariate is global, it must be defined at the regional level.

Data for abundance and occupancy modeling

The proposal states which data sources will be considered and suggests the use of simulations to determine the efficacy of abundance modeling (a sound suggestion). The Investigator Team may want to consider a similar approach for the occupancy modeling to establish whether parameters may be potentially biased.

> It would be worthwhile to construct a simulation model during development of the occupancy model. The simulation model could evaluate two aspects of the existing data. The first is the extent to which occurrence (presence / absence) data in the catch accurately reflect habitat. A strong relation between occurrence and habitat quality was assumed by Feyrer et al. (2011) and Manley et al. (2015); neither of the studies accounted for imperfect detection. The second aspect is the extent to which estimated coefficients may be biased if detection probability is heterogeneous among stations within a region. The bias could be evaluated given either that the heterogeneity is accurately modeled or that the heterogeneity is not accurately modeled.

The team may want to develop a backup modeling approach if neither abundance nor occupancy modeling can be undertaken with the available data.

> If we are unable to estimate occupancy, then it is unlikely that the available data will allow inference to survival and reproduction.

The proposal indicates that if the assumptions that accompany use of spatial replicates for occupancy modeling are not met, then estimates of abundance may be biased (page 24). However, it is not clear how estimates of abundance will be derived from application of occupancy models.

> The occupancy models will not be used to derive estimates of abundance. However, the outputs of the models may suggest covariates to include in analyses of abundance.

The proposal lacks a discussion of the necessary assumptions of occupancy models and how these may affect inferences from the proposed application.

> Occupancy models have five core assumptions. First, occupancy of a species at a given site or sampling location does not change among surveys (i.e., the closure assumption). Second, either there is no heterogeneity in probability of occupancy among sites, or heterogeneity is effectively modeled. Third, either there is no heterogeneity in probability of detection of a given species among surveys and sites, or heterogeneity is effectively modeled. Heterogeneity in detection probability may result from observer error, variation among observers, or abiotic or biotic environmental attributes that affect observers' ability to detect a species. Fourth, detections of the species are independent across surveys and locations. Fifth, occupancy models assume that species are not falsely detected.
>
> Conducting surveys within a relatively short period of time reduces the probability of violating the closure assumption. Because replicate samples were not collected at each station during each time step, we will use spatial replicates (e.g., samples from multiple stations within a region during each time step).
>
> We have proposed to test whether probabilities of occupancy and detection are heterogeneous and, if so, to model heterogeneity as functions of biotic and abiotic covariates.
>
> Detection of the same individual at multiple locations does not violate the assumption that detections are independent. To meet the assumption of independence of detections, the observer's actions must not affect movements of individuals.
>
> The probability of false detections of Delta smelt is extremely low.

The STN survey, which will provide catch data from which to estimate abundance of Delta smelt in August of each year, is characterized by inconsistent sampling frequencies since its inception. Prior to 2003, 2 to 5 surveys were completed each year, whereas, 6 surveys were completed from 2003 to 2014. What is the effect of variable sample sizes and how might this affect understanding of annual survival of Delta smelt across 'wet' and 'dry' years? Perhaps a simulation study using the 2003-2014 data can provide information on the utility of reduced sampling for estimation of Delta smelt abundance in August.

> The variable number of samples in the summer townet survey may affect the precision of the estimate of abundance in August that we use to initialize the abundance model. The precision of the underlying state variable of abundance in August may increase as the number of samples in a given year increases. We could conduct a sensitivity analysis to explore how the number of samples in the summer townet survey affects estimates of coefficients in the abundance model.

The information provided about the 20 mm survey is insufficient, but it appears that the duration of this survey will limit the analysis of stock and recruitment to the period during which this survey was operational, 1995 – 2014. Information regarding the number of stations, the type of sampling design (fixed station survey, random stratified survey, or other), and the spatial distribution of stations would help to understand how this survey can be used to inform the recruitment analysis. The 20 mm survey includes replicate tows, but it is not clear if these are independent replicates, or how the catch and other data from replicate tows will be treated (whether they are independent or not).

> The implementers of the 20 mm survey assumed that the replicate tows are independent, which allows them to average the catches per unit effort and generate the 20 mm index of abundance.

Several references are made to the use of Bayesian estimation, but these references are not clearly linked with a specific model. Also not clear is how the N-mixture models will be used and which questions these models will address. The section on 'other models of dependence between abundance and detection' is not well integrated with the remainder of the proposal.

> We noted that we will use maximum likelihood and Bayesian methods to fit all of our models. The subsection on other models of dependence between abundance and detection is within a section on research challenges that we were asked to include in the proposal. Our aim in this subsection was to explain that detection probability can be modeled as a function of abundance, but the ability to do so is constrained by data availability and the time available for analyses.

One of the limitations that should be acknowledged is that occupancy models may not be able to detect a relationship between occupancy (the proportion of sites occupied) and a given environmental covariate if the range of the measured covariate is narrow. For example, the full salinity or temperature gradient of the Delta may not be sampled by the STN or FMWT surveys. In any given year, the range of the 'flow' covariate(s) may also be restricted. The temporal scales of analyses (monthly flow vs annual flow) require further clarification.

> Correct, but we only can analyze the data that have been collected. There is little we can do to address the possibility that the full gradient of salinities or temperatures has not been sampled.

Flow data in occupancy models

How will global covariates (outflow, e.g.) be considered in occupancy models? Does outflow affect occupancy, detectability, or both? If both, how will outflow be characterized/measured? The temporal/spatial measurement scale of the covariate depends on how it is used in the occupancy model. For instance, the outflow covariate for detectability within a given patch (an area encompassing multiple stations) may be represented by the mean flow estimated at each station within the patch. The outflow covariate for occupancy may be represented by the mean flow across patches. Are station- specific estimates of 'flow' available for these calculations or will hydrodynamic model-based estimates of flow be used?

> This comment is a bit confusing. Because detection occurs at the station, the flow

covariate is a station-specific metric such as velocity or turbidity. We do not plan to model flow at the station level. Because occupancy will be estimated at the patch level, flow can affect probability of occupancy; flow covariates are measured at the extent of the patch. Values of global variables (i.e., those that operate across the ecosystem) still need to be specified at the extent of the patch.

Another concern with the outflow covariate for occupancy is that outflow may be confounded with year and cohort of Delta smelt; that is, a single outflow measure is available for each year (or group of months, say, September to December or September to April). Perhaps categories of outflow such as low/med/high, may be useful in modeling occupancy. Alternatively, annual estimates of occupancy (proportion of area used by Delta smelt) can then be compared with annual patterns in outflow conditions in the Delta. The recent work of Latour (2015) is highly relevant to identifying how flow will be represented in the analysis. This work indicated little support for the importance of flow.

This is the reason we proposed to use the data from the summer townet survey to estimate abundance of Delta smelt in August and to account for the possibility of different patterns of occupancy given different levels of initial abundance.

Tidal data in observation process

The Panel suspects that both tidal stage during each fish survey and the geomorphological characteristics of the station, e.g. channel area (depth, width), tow length of sampling, channel and tidal volume excursion, may be informative covariates in identifying catchability. These covariates will also be important to include in the occupancy modeling.

We agree. Our proposed methods allow us to evaluate the extent to which the data support numerous hypotheses about associations between detection probability and covariates.

One of the covariates proposed for consideration is 'surface area or volume of tidal marsh'; during the review Panel's discussion with the Investigator Team, one of the PIs indicated that area and volume will be determined from areas adjacent to the patches sampled for Delta smelt. Will area or volume be considered as a covariate for occupancy, detectability, or both?

We do not remember this part of the discussion, and are uncertain what is meant by "areas adjacent to the patches." Nevertheless, area or volume could serve as covariates for either detection or occupancy.

Allocation of time to prepare data

The Investigator Team expects to update the currently available survey catches for Delta smelt with information collected through 2014 . . . no time has been allocated for quality-assurance testing or quality control. We recommend allocation of sufficient time for these important tasks.

We have assumed that the data will be updated, their quality will be evaluated, and they will be provided to us. However, we agree that we need to allocate more time to working with those who are updating, evaluating, and providing the data to ensure that the metadata (e.g., the processes by which the data were collected and compiled) are clear and detailed. We also will allocation additional time to formatting the raw data for inclusion in models.