August 17, 2019

Dear Dr. Adkison,

Thank you for the opportunity to revise our manuscript UMCF-2019-0014 “Evaluating Population Recovery Characteristics for a Long-lived Protected Species: A Case History of Gulf Sturgeon in the Apalachicola River”. We have extensively revised the manuscript based on input from editorial team and reviewers. Below we provide the original reviewer comment in italics, and then provide a detailed response to these comments with reference to the location of the revision in our revised manuscript. We look forward to working with you to finalize revisions to this manuscript.

Reviewer 1

*1) It is difficult to evaluate the quality of the model, and its potential for management use, without seeing the uncertainty of the data and how well the model fit the data.*

This comment is echoed by the AE. To clarify, we have not developed a new model nor have we estimated new life history or population parameters for use in this model. To address reviewer concerns, we have revised Table 1 and Figure 1 to better represent data uncertainty and demonstrate model fit to available data. The revised Figure 1 now includes point estimates of abundance available from other studies that were used in tuning input parameters used in this model to match observed population characteristics. We have also added a detailed paragraph (lines 322-345) in the Results section discussing overall model uncertainty to leading parameters *recK* and *N0* as well as *Mai*and *k*.

*Where did the F values mentioned in lines 172-173 come from, and what are their variances?*

These F values are in the range estimated by Ahrens and Pine (2014), clarification included in revised manuscript.

*A range for the 1985 abundance estimate is shown in Table 1 (I assume this is a confidence interval?) – was this used in the estimation process and simulations?*

Table 1 has been revised to clarify the credible and confidence intervals on abundance.

*Lines 191-194 discuss fitting the data, and some text and plots showing how well the model fit the data would be helpful. The graphs should also show confidence bounds on the estimated abundances.*

Figure 1 has been revised to include estimates of population size including confidence bounds from other studies which provides insight into how well this model fits available data.

*2) Even if a particular management action was found to increase produce a recovery, there is the question of whether how well the current monitoring program and assessment procedures could estimate population size and this recovery. The projection model assumes perfect knowledge of future abundances. The authors should reference some of the management strategy evaluation literature (see Punt et al. 2014), which considers how sampling and population assessment errors affect estimates of stock abundance.*

We agree that there is substantial uncertainty as to whether the current monitoring and assessment programs could assess recovery. At present there is not standardized monitoring program for Gulf sturgeon range wide, and objectives of monitoring efforts within individual rivers are poorly defined. We have refrained from framing this exercise as an MSE because we do not have the buy-in from management agencies to frame the management objectives. This exercise does meet the basic steps described in Punt et al. 2016 and our hope is that resource managers will be interested in refining this model for use in formal policy evaluation. We have included text highlighting how a formal MSE could be the next step in this process (lines 357-361) and also added further emphasis on the need to incorporate monitoring and assessment programs that could evaluate whether population benchmarks were met.

*3) If one of the goals for delisting the population is “**population abundance that could sustain a fishery”, then the authors should consider how this relates to the carrying capacity. The manuscript reads as if rebuilding to carrying capacity is the metric for success. However, sustainably harvested stocks will be below carrying capacity; for common biomass dynamic models, maximum sustained yield occurs at half of carrying capacity.*

The discussion of whether rebuilding to carrying capacity is a metric for success is important and we have added text addressing this point (see new section on population productivity lines 316-321). In estimating Umsy and Bmsy we have used the current carrying capacity (with the dam in place) in these calculations. Deciding which metrics should or should not be included as recovery goals is a management question answered by the two federal agencies tasked with managing this species and this is highlighted in the revised text (368-372).

*4) I found the description of the methods (and particularly the simulations) to be difficult to follow. Some examples are below:*

*a) Lines 146 – 155) The steps involved in calculating the stock-recruitment parameters from the compensation ratio should be explained more clearly (and the compensation ratio should be defined when it is first mentioned).*

We have provided a definition of the compensation ratio.

*Equation 1 has a typo – please see Eq. 9 of the Pine et al. (2013) paper the authors cite.*

The typo has been corrected.

*The β parameter is not simply carrying capacity, but is a density-dependent parameter. In this formulation, carrying capacity would be a function of both α and β. In general, the description of this stockrecruitment model in Pine 2013 is much clearer, so I suggest using that as a guide.*

We agree and apologize for that obvious mistake. We have ensured all mention of the ** parameter is referenced as the density-dependent parameter, whereas all discussion of carrying capacity now specifically relates to asymptotic abundance for the entire population (i.e. classic ecological definition of carrying capacity)

*More importantly, given that “each pre-recruit stanza had equal relative mortality and habitat capacity, which implies each recruitment stanza was equally long and with similar bottlenecks”, I wonder if anything is gained by separating recruitment into stanzas.*

It is necessary to separate recruitment into stanzas to reflect a pre-stocking stanza (wild fish are competing among themselves) and a post-stocking stanza, where wild fish may need to compete with stocked fish. We have added text (lines 158-161) at this point of the manuscript to make this clear.

*b) Lines 168-170) state that “We initialized our population model (initial population size (N0)) with parameter estimates representing the initial, pre-exploitation population of Gulf sturgeon”, but lines 194-195 (in describing the baseline model, or scenario 1) state that “We then used this model to compare differences in time until population benchmarks are met using a range of estimated (N0) values”. So was a single value of N0 used for Scenario 1, or a range of values? If lines 194-195 refer to scenario 2 (as suggested by Table 3), it is confusing not to include them in the paragraph describing paragraph 2. Also confusing is referring to the initial population size as “parameter estimates” if they were set to a particular value (even if this was an estimate from some other model) because estimating the initial population size as a parameter is a common feature of population models.*

We have revised the text to clarify that these were parameter values rather than estimates, given this is a simulation model and the initial population size was not statistically estimated in this exercise. Otherwise, the text formerly in lines 168-170 (“We initialized our population model…”) is correct. We have replaced the text mentioned from lines 194-195 with “Outcomes of this scenario were the basis against which all other scenarios were compared” (lines 201-217), which is what should have been said initially. There were no scenarios where N0 were adjusted, so no further text (specifically in Scenario 2, where depletion level was varied by manually adjusting pre-closure exploitation rates (U)) was changed.

*c) Lines 171-174) “Removals during this period are simply the initialization of the population model (transient dynamics) that results from adjustment of the population to the applied rate of mortality.” Initialization of population models usually refers to defining the starting abundance and model parameters necessary for forward simulation, and would not be affected by mortality rates beyond the initial year. Transient dynamics usually refers to short-term fluctuations in a deterministic model before equilibrium is reached. The initialization of the model can affect the transient dynamics, but these are two different concepts. This sentence seems to be just saying that the population abundances are a function of the specified starting values and mortality rates.*

We have revised this section for clarity and have used the suggested sentence. Thank you for the suggested text.

*d) Line 188 – 190, 200) – “A baseline population simulation (Scenario 1) was created to estimate simple projections of population size and establish a reference from which to compare other models.” “Simulation” is singular, but “projections” is plural. If scenario 1 is a set of simulations, then how do they differ from each other, and how many simulations were done? The caption in Figure 1 suggests that the spread shown in Figure 1 was obtained from simulations at the upper and lower 95% confidence bounds of the estimate of N0 – it would be useful to put this in the text when first describing scenario 1. In line 200, what does ‘median population trajectory’ mean? What was the set from which the median was calculated?*

We have clarified that Scenario 1 is a single simulation, although we have also properly described how we calculated our ‘confidence limits’ (lines 200-2010). We have further clarified at the end of this paragraph that confidence limits are for visual reference only as they are not true confidence limits (as the reviewer points out).

*e) Lines 199-200) The fishing mortality rates were initially introduced as, I assume, instantaneous rate (i.e., F rates), but are now being introduced as annual exploitation rates. Over what time period do these manipulations in mortality occur? Table 3 refers to U = 1.0, so if this is an annual rate it would remove the entire population. Umsy is referred to in the text, but not in Table 1, whereas Fmsy is referred to in Table 1 but not the text (and neither are defined). There seems to be some confusion on annual and instantaneous rates of mortality.*

Fishing mortality rates are actually apical exploitation rates where *U* is multiplied by an age-specific vulnerability to get total exploitation. Therefore, if apical exploitation rate is 1.0, 100% of the vulnerable population will be removed each year, but the some of the entire population would remain. We have made this clear on lines 221-224. We have removed *FMSY* from Table 1, as it is not a parameter and was accidentally left in through the development of the paper. We have also defined *UMSY* and *BMSY* in the text (lines 261-264)

*5) The authors note that SPR is not a measure of the size of the spawning stock, and for that reason they many want to consider whether it is useful as a metric for population recovery. SPR values are used as an indicator of fishing intensity, and the SPR calculation the author employ (which is often call the transitional SPR; see Mace et al 1996) is an indication of fishing intensity experienced in the past by the cohorts in the population. However, the cessation of fishing does not necessarily imply that the population has increased to meet recovery goals.*

We completely agree with the reviewer and have added text to this effect, suggesting neither SPR nor total abundance are truly reflective of low extirpation risk (lines 471-482). We have also added the Mace et al. 1996 reference.

*Minor comments*

*Line 81) Change to “. . . before population recovery is considered . . .”*

Change made.

*Line 244) Scenario 1 SPR not shown in Table 3.*

We have added this line to Table 3

*Lines 251-252) Repetitive, repeats material from lines 243-244.*

Duplication removed.

*Line 272) t1 is defined as 1901 in line 171, but fish stocking is modeled as beginning in 1985. More generally, why are you simulating what fish stocking rates could have happened in 1985? By now, it should be known how many fish were stocked in the past.*

The scenarios are all hypothetical and the years are referenced in terms of “if this action had been taken in 1985 what would the population have looked like in 2023” the reference year identified by the recovery plan.

*Line 279) “Important”, not “import”*

Correction made.

*Lines 303-304 “. . . and BMSY at current carrying capacity levels only about 1859kg (<200 age-4+ fish).” Where was this number in the Results section, and how was it computed?*

We have added a section on population productivity, which discusses these findings (lines 316-321). We have also added text to Methods (lines 261-264) where we discuss how MSY was numerically solved.

*Line 338) “effective fishing mortality (catch/abundance)” This is an exploitation rate (for most people, I think the phrase “fishing mortality” would refer to an instantaneous rate).*

*Lines 337-339) “landings that occur following population collapse likely increase the risk of depensatory mortality where the effective fishing mortality (catch/abundance) is high even if landings are low relative to historic.” Depensation means that the pre-capita rates of recruitment (or population growth) decline as the stock becomes smaller, and requires sigmoidal stock recruitment curves or production functions. High fishing rates at low stock sizes does not define depensation. Evidence for depensation was not presented in the paper.*

This has been revised to remove discussion of depensation.

*Line 361) “States”, not “stats”*

Correction made.

*Line 376) How much ability to managers have to reduce natural mortality? It seems like things like red tides, weather anomalies, and hurricanes would be beyond the control of a fishery manager.*

*Table 3) SPR values should not be above 1. For the scenario 3 runs, I suspect that the natural mortality values for the unfished population was not consistent with the value for the fished population.*

This is consistent with the scenario we have described. In Methods (lines 227-229), we state the natural mortality may have changed since closure of the fishery, and this is the basis for this scenario evaluation. As such, we change mortality after 1985 and examine how this impacts abundance and SPR; text in Results (lines 282-285) reiterates this point. From that perspective (as the reviewer correctly states), SPR can be greater than 1.0.

*Figure 6) There is no description of an analysis related to age structure in the Methods section. What is κ? The caption (which discusses a surface plot and cells) does not match the figure.*

This figure has been deleted.

*References*

*Mace, P., Botsford, L., Collie, J., Gabriel, W., Goodyear, P., Powers, J., Restrepo, V., Rosenberg, A., Sissenwine, M., Thompson, G., and Witzig, J. (1996). Scientific review of definitions of overfishing in US fishery management plans. Report Supplemental to NOAA Technical Memorandum, NMFS–F/SPO–17. 20 pp.*

*Punt, A.E., Butterworth, D.S., de Moor, C.L., De Oliveira, J.A. and Haddon, M., 2016. Management strategy evaluation: best practices. Fish and Fisheries, 17(2), pp.303-334.*

*Reviewer: 2*

*Comments to the Author*

*This is an interesting, well written article modeling population recovery of a long-lived species, Gulf Sturgeon, under different management actions. The modeling approach and most of the assumptions associated with it are laudable. The modeling predicts recovery time to a 2023 benchmark, at least to contemporary carrying capacity.*

*While the modeling approach is valid and noteworthy, really nothing presented in the manuscript is novel. The exception being it is for a single Gulf sturgeon population in the Apalachicola River.*

The model is not new, but the application as it relates to potential management actions that could be taken to promote recovery given the large infusion of funding from *Deepwater Horizon* related settlements.

*Moreover, existing literature on sturgeons is lacking within the manuscript (for example, population modeling conducted by Gross et al. (2002) and Velez-Espino and Koops (2009)). The situation for this population is very similar to lake sturgeon in the Great Lakes where the adult population was decimated in the late 1800s, and population recovery was impeded by anthropogenic stressors such as damming of rivers, pollution and habitat alteration (e.g., Auer 1999) which probably could be applied to Atlantic sturgeon (Hilton et al. 2016) and most other NA sturgeons. Similarly, sensitivities to adult mortality (Boreman 1997) and safe exploitation rates have been well established for sturgeons (e.g., Bruch 1999; Rieman and Bermesderfer 1990). Finally, I am not sold on the parameters used for stocking. It can be an effective tool for sturgeon (e.g., Schram et al 1999) however, age class stocked could have a drastic effect on survival (e.g., McDougall et al 2014), therefore 90% survival may be a little high. It also could be construed as the optimal solution by senior managers. More reference to existing sturgeon literature would strengthen this case study.*

We have performed a more in-depth literature review and added additional citations to help support our study, especially in the Discussion section. Many of the additional citations are sourced from Lake sturgeon which are also long-lived, but are not anadromous. We have cited key Lake Sturgeon papers identified by this reviewer and have focused additional citations and information on Atlantic Sturgeon literature, which are a conspecific to Gulf sturgeon.

*Specific comments*

*L180 - there are studies that quantify the effect of dams on sturgeon (e.g., Haxton et al. 2015) which may corroborate to refute the parameters used.*

We recognize the effects of dam on sturgeon extensively in this paper particularly as related to carrying capacity. We have provided the additional references as recommended by the reviewer.

*L279 - should be 'important'*

Revision has been made.

*L332 - in conjunction with other anthropogenic stressors. There would be a reduced reproductive potential with the construction of the Jim Woodruff Lock and Dam.*

This is accounted for by adjusting the carrying capacity for the river.

*L358 - include scientific name for white sturgeon*

Addition has been made.

*L367 - but not to other sturgeon populations.*

This has been revised to indicate this is similar to other sturgeon populations.

*L391 - 392 - it would not be as simple as providing passage. Spawning habitat has to be still available upstream of the dam (i.e., not flooded out), and young (i.e., drift phase) has a chance to survive (i.e., sufficient length of river for larval drift enhancing survival e.g.,* *Auer and Baker (2002); Braaten et al. (2008); Mailhot et al (2011)).*

This section has been revised to highlight that not only does passage have to occur, but spawning habitat has to exist and conditions suitable for larval and juvenile phases.

*L406 - include the scientific name for Atlantic sturgeon*

This addition has been made.