Dear Editor (Andre),

I have enclosed a revised manuscript addressing the concerns thoughtfully raised by the reviewers. The changes made are enumerated point by point below. Reviewer comments are in black text and responses are in italicized blue text. Thank you for the time spent reviewing this manuscript! Let me know if I need to adjust the method of submission (e.g. remove line numbers or provide .pngs for figures).

Regards,

Cody

REVIEWER #2

In this manuscript, the author investigated the effects of different combinations of time-variation in natural mortality (M) and catchability (q) on important management parameters. Specifically, author explained the consequences of undertaking these types of assessments on management advice, emphasizing large differences in management advice underscore the importance of evidence-based approaches to incorporating time-variation in population parameters in stock assessment. Author used the eastern Bering Sea snow crab assessment to demonstrate the effect of time-variation in population parameters on management advice.   
  
Author did a good job of bringing out the idea of time-variation in vital population parameters and demonstrated that time-variation indeed improved model fit to data but with an unfortunate outcome of conflicting, sometime untenable, management advice. Author provided some advice to overcome these problems. Are there any improvements that can be made in the analytical approach and presentation to convey the message clear? Although further improvement in the analysis is difficult with currently available data, nevertheless the presentation can be vastly improved.  I will address those in the General and Specific comments sections below.  
  
The General Comments:  
  
My first impression is that the presentation needs more work to convey the message clearly to wider audience.   
  
(a) The author uses a general phrase "population processes" in several places while focusing only on catchability (q) and natural mortality (M). I suggest further clarifying this phrase focusing on present analysis. This is because in the first sentence of the abstract, author mentioned that the population processes can be mutually confounded, which is true; but the author has not looked at growth and other environmental parameters that confound with M and q estimation.

*I added a sentence in the discussion recommending that other processes be considered in a similar manner. If you’re okay with snow crab being a ‘case study’, the extension of the ‘case study’ is to only look at the processes M and q, so I kept the reference to ‘population processes’.*  
  
(b)  I agree with the author's statement in the abstract that evidence-based approaches are needed to incorporating time-variation in M and q. But the manuscript did not provide undisputable evidence for changes in M and q by time occur in nature. In my opinion, the analysis only demonstrated that data fitted better with time variation in parameters than when they were treated invariants. So, I suggest including this caveat from the analysis to the abstract.

*This was the entire point of the analysis…the abstract already contains this idea: “The North Pacific Fisheries Management Council did not adopt any of the models with time-variation in M or q for management because the estimated variability in M and q was difficult to explain with any known mechanism…”*  
  
(c)  Writeup needs editing/improvement to convey the results clearly to wider audience. I have noted them in the Specific comment section below.  
  
Specific Comments:  
  
Abstract  
  
This section is concise and conveys summary of the manuscript.  
  
A caveat can be added at the end of the abstract to convey a message "evidence-based time-variation in q and M could not be demonstrated with the available data on snow crab."

*I did not add this based on this information already being present in the abstract.*  
  
Introduction  
  
This section satisfactorily introduces the subject and scope of the work. However,  
  
Line# 23: Please revise 2014 to 2019 in Stawitz et al. (2014).

Done.  
  
Line# 32: Please revise 2014 to 2013 in Taylor and Methot (2014).

Done.  
  
Line# 60: Please revise "…quotas since 2004…" to "…quotas since 2005…"

Done.  
  
Lines# 68-69: "The most recent large pseudo cohort was spawned around 2010 and recruited to the survey gear in 2015. By 2018, it was the largest pseudo cohort ever observed in the eastern Bering Sea."  
Please justify elapsed time between spawning and recruitment with any supporting reference(s).  
Done.

Lines: 69-72: "However, in 2019, it was much smaller than expected given estimated growth, natural mortality, and fishing removals (Figure 2). This strongly implies time-variation in some population process and catchability or natural mortality are key suspects given previous research."  
Why didn't you list movement from the survey area? Plus, please provide references to previous research.

*Movement from the survey area is a mechanism by which catchability can change.*  
  
Methods  
  
This section satisfactorily explains the methodology used in the analysis. However,  
I have few comments/questions in this section:  
  
Lines# 109-110: "A smoothing penalty is added to the negative log likelihood of each model in the form of the squared norm of the second difference of each vector of additionally estimated estimated parameters multiplied by a user defined weighting factor."

First, please remove one "estimated" from the repetition.

*Done.*

Second, why do you need to add a smoothing penalty to likelihoods?

You need to explain the reasons behind adding this; otherwise, some readers may not understand your reason.

*Added the text, “to facilitate model convergence”.*

Finally, why didn't you undertake a sensitivity analysis on user defined weighting factor? Otherwise, the chosen weights might have some effects on the results of time-varying parameters.

*The goal was to make them small enough to allow convergence but no larger. That is described in the text. There is nothing to present except a table that says ‘weight’ and ‘converged?’, which would have numbers and yes/no in it. There would be no point in looking at the output from any of the models that did not converge. More sensitivities with larger weights would not add to the message that without a clear mechanism, it is hard to select among models with different hypotheses about time variation (regardless of the smoothness of the estimated process). It might be useful to write another paper about how to select smoothing penalties, but doing so here would distract from the central message. I now note that this should be explored in the discussion.*

*That said, a single ‘sensitivity’ was performed to check the differences among model outputs for the ‘minimum viable’ penalty from one model in another. It resulted in rather large changes in management quantities.*   
  
Lines# 122-135: This paragraph needs some revision to convey the message to readers from outside the US.

*I’m not sure what the reviewer would have me add. I state clearly what the OFL is and how it is calculated. The other reviewer had no concerns here.*  
  
Line# 129: "…average of the final 7 years of the process…"  
  
You can be more specific by stating averages of the final 7 years of M, q, and bycatch mortality????

*I think the preceding clause ‘for models in which time-variation was implemented’ is sufficient to direct the reader as to what is occurring? Perhaps I’m misunderstanding…*  
  
Lines# 134-135: Please add a reference as: (Amendment 24, NMFS; NPFMC, 2007).

Done.  
  
Lines# 135-136: Equation (1):  
  
Please modify the middle formula inequality part to  
     ….  if 0.25 < MMB/MMB35 <= 1

Done  
  
Results  
  
This section satisfactorily explains the results of the work. I have few comments though:  
  
Lines# 142-147: I have already mentioned about providing the purpose of adding smoothing penalties. Here you are mentioning about " The smallest viable smoothing penalties tested…" This is puzzling me because I have not seen any specific smoothing penalty formulas in this manuscript to evaluate your point.

*I noted in the methods that the ‘smoothing penalties were chosen by trial and error, with the aim of making them as small as possible while still avoiding convergence issues’. The form was noted in the same paragraph—second differences on the vector of estimated devs, I’m not sure what the utility of writing an equation out there would be.*  
  
Line# 175: Please revise "…estimated average catchabilities similar to…"  "…estimated average catchabilities were similar to…"

The sentence is “Models in which catchability was time-varying estimated catchability similar to…”. The subject is ‘models’, the verb is ‘estimated’. Changing it to the suggested sentence does not work here.   
  
Discussion  
This section adequately discusses the results with findings from published literatures and suggestions for further work.  
  
I agree with the first sentence of the last paragraph that ---Given the outcomes of this analysis, it is not clear if the estimation of time-variation in confounded processes is a problem we can model our way out of with the currently available data for snow crab.  
  
References  
  
The references are complete and referred to in the text. However,  
  
Lines# 248-249: page numbers not complete.

It is part of this issue and has no page numbers. DOI now included.  
  
Line# 263: Please expand Hilborn, R. et al. 2021.

Too many authors…though I’m not clear what the rule for Fish Res is…  
  
Line# 271: Please expand Johnson, K.F. et al. 2014.

Same as above.  
  
Lines# 322-323: Page numbers not provided.

Fixed.  
  
Appendix A  
  
I could not find any detectable errors in the set of population dynamics model formulas.  
  
I suggest adding the smoothing functions in this section. Also define the weights used in several likelihood formulas. Various weighs can be defined in a table.

I now note that the weights can be found in the .CTL file in the github repository with the text: “Weightings of all likelihoods and penalties can be found in the .CTL file in the github repository linked at the end of this paper.”   
  
Line# 352: Then et al. (2015) reference is missing in the reference list.  
Added.

Tables  
Two tables are adequate for the main text.  
Table 2. Please explain in the title how did you get a mean M for the varying M scenario.  
Text added noting it is the average value over the entire time series of estimated M.

Figures  
My document shows that there are 22 Figures to this manuscript. However, Figures 9 to 22 are not referenced to in the text. Am I missing something here?

I removed the figures from the manuscript and note that the reader can find them in the github repo as ‘supplementary materials’.  
  
Figure 1 title: Are you sure that retained catch data are reported from observer data? I think they are reported from fish ticket data. Perhaps some aspect may be confirmed by observer sample!! Please check.

Fixed. (Thanks!)

REVIEWER #1

Dear Dr. Szuwalski,

Thank you for your recent submission to *Fisheries Research* titled “Estimating time-variation in confounded processes in population dynamics modeling: a case study for snow crab in the eastern Bering Sea”. The manuscript summarizes modelling efforts for the Eastern Bering Sea snow crab with respect to including time-varying processes in the stock assessment model. Investigated processes include natural mortality and catchability, along with the combination of both processes. I found the manuscript easy to read and “complete” with respect to the analysis. I am suggesting that it be accepted with minor revisions because I would like to have a bit more interpretation of what is going on behind the scenes in the model and the management process that rejected the inclusion of the time-varying processes.

The paragraph that starts on line 56 was difficult to follow. I realize all of the information in it is about Eastern Bering Sea snow crab but it did not flow well for me. I think the topic sentence should be more general and just about snow crab rather than something specific about time-variation in snow crab because then I thought that the entire rest of the paragraph would be about biology and research associated with that biology. Also, the terminal sentence about not being able to age them and the definition of a pseudocohort should come before its use rather than afterwards in parentheses.

I split this paragraph in two and added some text to hopefully make it clear that the point was to 1) link snow crab to potential time-variation described in the previous paragraph (topic sentence), 2) say snow crab is the case study I’m using, and 3) give background on snow crab in the EBS.

I did not like the use of code for labeling in the paragraph that starts on line 101. I think that it is more explicit to say a “vector of deviations” (i.e., labelled as ‘dev\_vector’ in AutoDifferentiating Model Builder; Fournier et al., 2012)”. In the next sentence you anthropomorphize the vector by saying that it estimates. When I believe that sentence should read “The vector includes an …”.

Changed.

Can you please include some text regarding the increase in the scaling parameter on the time-varying processes across models? You mention it at the end of the results but then I did not remember you addressing it in the Discussion section. Basically, I am assuming that the model can kill fish many ways and it does not have a conscious with respect to assigning changes in the data to one process or another. What data do you think is needed to flush this out? Or, is it a lack of specification in the model? I think it the former but would be happy to read text about either.

I think you mean the ‘smoothing’ parameter and I’m not really sure based on the analysis in the paper what to say. I don’t think it would be particularly proper (or useful) to speculate on how I could identify appropriate smoothing penalties with the analyses in this paper, but I now note at the end it would be useful to understand how to do this (and I’m working on another paper that explores this idea some).

Discussion section is hyper-focused on time-varying processes rather than explaining or putting into context the resultant rejection of the time-varying models in the assessment process. I would like more context and explanation of the reasoning why the models were rejected and potentially why other models are rejected in favor of things that are “incorrect” simplifications. Feel free to reference other stocks. Also, you could talk about the min-max solutions as far as the least-worse outcome when we know that things will never be ideal.

The second paragraph of the discussion directly addresses this point--improvements of fit are expected if you allow the model more freedom (regardless of the process) and if you let the wrong process vary, management advice is very different. There was no clear mechanism for the time-variation, so there was no satisfying way to choose among the hypotheses. I added a sentence emphasizing this at the end of this paragraph.

Thank you for a well-written, easy to review manuscript. It was a pleasure to read. I just think that it needs a bit more in Discussion about why the models were not used if they provided better fits to the data. Below you will find minor suggestions in addition to the larger ones listed above.

Kind regards,

Kelli

Global changes

“e.g.” to “e.g.,”

Fixed.

” should be “Eastern Bering Sea”

No, I don’t think so. “Bering Sea” is the name of the location—‘eastern’ describes a part of it. It’s not sufficiently well-known to warrant capitalization (I think).

“status quo model” should really be “status-quo model”

I didn’t change this one. Hyphen makes sense for two adjectives, but ‘status quo’ is not two adjectives. I’m not sure what the right answer is here.

Whenever the word data comes after “size composition” it should be “size-composition data”

ok

All dashes between numbers should be en-dashes rather than just a regular dash, where an en-dash is a mid-sized dash approximately equal to the width of the letter n and used for ranges between numbers and dates.

Not sure what this referring to in the manuscript.

You define OFL and then use it subsequently as “The OFL”. The article (e.g., the, an, a) is not needed unless you are using the acronym as a modifier, which you are not.

ok

Should you use “MMB-per-recruit” rather than “mature male biomass-per-recruit”?

Changed.

Parameters are only italicized in equations where they should also be italicized in the text in my opinion.

Changed.

Remove the quotes from all subsequent uses of model names. Sometimes you use single quotes and sometimes you use double quotes but the single quote should only be used in the first instance.

I did not do this because it made it harder for me to read because it was harder to delineate where the model name ended in the sentence.

You consistently anthropomorphize the models. Statements such as “Models … estimated” when the model doesn’t really estimate anything. Instead a parameters was estimated would be more correct.

I am fine with this as it is. I don’t particularly like the passiveness of ‘was estimated’ and the model is what did the estimating (I sure can’t calculate all those derivatives…).

All instances of “Fish Res.” in the References section should be “Fish. Res.”

Changed.

Information printed on the screen is NOT read by screen readers and is therefore inaccessible to those with disabilities and should be clearly laid out in all captions.

Full stops at the end of figure captions.

Fixed.

Line-specific changes

6: “mortality (M)” to “mortality, M,”

7: “catchability (q))” to “catchability, q)” to get rid of the double bracket.

All those commas looked strange, did not change.

7: remove comma before “but”

11: Comma after “Here”

20: “natural mortality, catchability,” to “natural mortality (M), catchability (q),”

Changed.

27: “in convincingly” to “when”

Changed.

39-40: remove “(e.g. … )”

41: change “:” to a comma

66: “(Figure 1 & Figure 2)” “(Figures 1 and 2)”

Changed.

68: I think “recruited to” should be “first selected by”

Changed.

73: remove the comma

78: Add a comma after “Here”

90: there should be a space before “mm”

Changed.

96: Change “Survey biomass indicies” to “Survey indices (mt)” or whatever unit of biomass that you are using

97: add the units associated with fishery and bycatch biomass

101: Change “:” to a comma and remove the sentence starting with “Each” as it is redundant with other information

108: Should it be “Eastern Bering Sea” rather than just “Bering Sea”?

Changed.

116: Add “models with” after “and”

Added.

117: Remove the sentence starting with “A retrospective pattern “ because you have already used the term two times in the manuscript, and thus, you already expected the reader to know what you were talking about.

I left the sentence in there to be safe.

119: I think the sentence would flow better if you used commas rather than parentheses for the which clause

129: The integer 7 should be written out in words

Done.

Eq (1): Why does only the first definition include bycatch because presumably bycatch occurs no matter what MMB is

Added the word ‘only’.

136: change “MMB is the projected mature male biomass” to “MMB” and change “mature male biomass” to MMB in the subsequent sentences. Remove “he” before “*FOFL*”

Done.

140: Why did you use 0.1? Perhaps add a reference.

Add that the council set this.

150: It is unclear to me what confidence intervals you are talking about, can you please specify which model they were estimated in?

Clarified that they are the input CIs that come from the calculation of the survey index.

165 and 167: remove the comma before “but”

168: remove “model”

Removed.

169: change to “estimates of natural mortality were much higher than estimates from status quo in recent years. Estimates from”

170: remove “very”

Done.

173: Did you call the survey the NMFS survey before, if so I don’t remember you it and the defined name of the survey makes me think that there are additional surveys rather than just one

Changed.

179: change “natural mortality” to “*M*” and “mature male biomass” to “*MMB*”

Changed.

188: THANK YOU for doing this additional analysis, what great foresight.

193: remove the comma

215:216: I think we meant to suggest that it should be a minimum default given realistic amounts of available data in Johnson et al. (2014).

Thanks, fixed.

246: looks like the link was botched and the underscores should be single where the manuscript looks like they were doubled, i.e., “\_\_” rather than “\_”

Hm. I’m not sure what’s up here. They are single.

249: no page numbers

Fixed.

271: I think all authors need to be listed or at least a first x number of authors, not sure

I’m not sure either…will let editor sort us out.

330: Fix the abbreviation of Reviews in Fisheries Science

Fixed.

Appendix: TLDR but I am assuming it is fine

Table 1. Acronyms used in the table need to be defined in the caption. The table would be easier to read if the decimals lined up going down a column. Number of decimals should be consistent, i.e., the trailing zeros should be present, e.g., “0.00” not “0”.

Talk to Rmarkdown, because it’s not listening to me.

Table 2. MMB is not defined nor are other acronyms used in the column headers. I recommend changing the header of average recruitment to “avg rec” and defining it in the caption rather than “avg\_\_rec”. It is unclear if you are reporting differences or actual values because you say “Changes in”.

Changed to ‘Reported management quantities…”

Figure 1: Change “inthe” to “in the”. Change “1000t” to “1000 mt” in the y-axis labels. Define MMB in the caption. You do not talk about the year in which quotas were introduced in the text.

Fixed. Yes, the year for quotas was discussed…but it was in the paragraph you found hard to read! Hopefully the confusion I caused is fixed now.

Figure 2: Use actual lines in the legend rather than a rectangular gradient because you use integers rather than a continuous representation of these values, i.e., you always report them for a given point in the year rather than fractions of years as implied with a gradient legend. I think you can shorten the survey name to just “survey” because there is only one survey. Which model is the dashed green line from?

Lines and text fixed.

Figure 3. Need to define the model names used in the legend in the caption. Use a full stop after “survey”. Change the y-axis label to “(1000 mt)” rather than “(1000t)”. What are the points and vertical bars from?

Added mt and full stop and text describing CIs.

Figure 4. Reference caption from Figure 3 for more information. Define MMB and change units in y-axis label. Change “mature biomass at survey” to “mature male biomass (MMB) of the survey”. Use en-dash for year range.

Added text.

Figure 5. Define what the points and vertical lines mean. Do you really need the male label when all you are plotting are males? Colors are not defined. Change units in y-axis label. Define Mohn’s rho.

Added text for points and lines. Removed ‘males’. Colors are consistent through the figures.

Figure 6: “Model predicted” should be “Model-predicted”. Same comments as some of the other figure captions.

Changed to “Predicted mature male biomass…”

Figure 7. Only the lower panel is labeled, all three should be labelled as defined in the caption.

Not sure what this means.

Figure 8. Use actual math notation with a superscript in the y-axis label. Change “time varying” to “time-varying”

Done.

Figure 9. Cannot you just plot one line and say that all models used the same information?

I could…but this gets the same point across. These are supplementary figures that I’ll just put in the repo.

Figure 10. Change y-axis label to have “mt” rather than “t”. The caption should explain the point versus line.

Done.

Figure 14. Do not use the same colors for males and females that were used for models in previous figures.

Figures 15-18. No y-axis label.

Text added.

Figure 22: “time varying” should be “time-varying”. Fix y-axis label.

Done.