Reviewers' comments:

Please make the changes suggested by the reviewers, particularly those that improve the description of the methods. Consider running the additional simulations as suggested by the reviewer to make the paper more complete or explain why they are not necessary or not possible.

We have addressed all of the suggestions provided by the reviewers and have run all of the additional simulations.

Reviewer #2: This paper describes a growth model for tag-recapture data of Antarctic toothfish that decomposes growth into a persistent and transient component and then uses a simulation study to validate this approach.

The subject of this paper falls within the general scope of the paper. The statistical model used in this paper appears scientifically sound, and the approach of identifying a persistent and transient component in growth models should be of importance for those interested in estimating fish growth and indeed in fish stock assessments generally. I also commend the authors for using a simulation approach to validate the method they propose. However, the paper reads less rigorously, particularly in the introduction and the method section regarding the simulation. A well-organized description of the simulation process will help reader to understand the simulation logic and I provide suggestions for how the paper could be modified.

Major comments

1. The introduction is generally difficult to read. To help the reader, I suggest you introduce the general concept of persistent and transient growth e.g. in line 56 before describing these aspects in details.

Done.

The sentences in lines 53-55 ("Similarly …") and lines 60-63 ('For example…") do not link with the previous sentences. Also, you have to more specific with 'size-specific harvest' (line 61), as a fishery may target small rather than large fish which could lead to a dominance of fast-growing animals in the fish population.

We have altered these sentences, hopefully linking them together better. We have also changed the sentence starting line 61.

The introduction mentions mixed-effect and hierarchical models, but leaves the reader to ponder why they were describe here (e.g. were they used in this study?).

We have added a sentence at the end of the para to address this.

2. Please present the model fits from all four growth models and combine them to one figure (not separately as in current Figures 1 and 3) so that the reader can easily compare the impact of the different model structures.

Done, see Fig. 3

3. My main points of concerns are about the simulation:

a) The logic and flow of this paper would be much helped if the evaluation of the method through simulation was reported before the actual application of the method.

The order of the simulation/application has been swapped.

b) The description of the simulation is unclear and insufficient to exactly understand the logic of the simulations and how they were performed. Please describe in detail what the general approach was and what it tests? I could also not understand some details, such as in lines 240-243: do you mean that for each sample, you randomly selected two out of the three parameters age at release, age at recapture and time at liberty, then randomly select estimates for those two parameters and finally calculate the parameter value for the third parameter?

We have added more detail here and also changed the simulation structure slightly (see comments below) so the question surrounding lines 240-243 no longer applies.

c) I remain unconvinced that the simulation fully tests the proposed method in the intended way. As far as I understood, you may have tested some statistical properties of the growth model, but not whether the model does correctly and unambiguously describe the actual processes of growth.

By deriving samples from the real data set as it was done in this study, the underlying processes that have led to the structure of these real data remain unknown. To conduct an simulation evaluation correctly, you need to fully simulate all processes when creating simulation data. This must be done to avoid false-positive conclusions about growth processes, where e.g. the preferred model includes variation in k and z, but the data was only created with variation in k. I suggest you conduct four simulation scenarios where data are created with (1) nor variation, (2) with variation in k, (3) variation in z, and (4) variation in k and z. Each of these scenarios is then evaluated with all four model structures. The same approach could also be used if you want to further expand the evaluation and investigate e.g. the impact of annual/spatial variation in growth between animals as touched on in the Discussion. These simulations can be conditioned with toothfish-like parameters, but of course proof of generality of the method would be beneficial.

The age at tagging and time liberty have been sampled from proper distributions with parameters estimated from the toothfish data now, rather than the resampling procedure used before. The four scenarios described above have been created, and the four estimation models fitted to each of the four simulation scenarios 200 times (i.e. 4 x 4 x 200). We also did an additional experiment where we simulated from just one of the scenarios (variation in k and z) with 50, 100, 250 and 500 individuals. We estimated each of these with the four different models (i.e. 4 x 4 x 200). These scenarios take over 2 weeks in total to run so we did not run the four different samples sizes (50, 100, 250 and 500) for any of the other simulation scenarios. Doing these additional simulation runs required that the figures be changed substantially. We feel that these additional simulations has improved the paper substantially and thank the reviewer for the helpful input.

4. What is the consequence of fixing <PSI> to values other than 0 (line 167)? Are the conclusions the same?

We tried estimating this parameter but generally found that models did not converge. When models did converge, psi was close to zero. Therefore, we fixed psi at zero in all other model runs. We have added additional sentences about these results and suggest that the exploration of this parameter be left for another paper.

Minor comments:

5. Please provide all parameter values that you use for the simulation in a table, especially from non-publicly available working group papers such as Dunn et al. (2006).

Added Table 2 which contains these values. We merged some of the other tables from the previous MS into this table as well to reduce the number of tables in the paper.

6. Lines 137-138 is repeated in lines 146-149. Be more specific on what other time intervals you tested.

Repeats removed.

7. Avoid judgemental descriptions (e.g. 'reasonable' in line 216), and replace with e.g. 'to evaluate potential bias and precision of the models'.

Done.

8. I don't consider that papers in review should be cited in a publication, but leave this issue to the editor of this journal to decide.

Will leave this for editor to decide. Two of these papers are no longer in review and have been updated in the MS.

9. There are also a number of grammatical and word errors in the text which should be easily picked up in the editorial process.

I picked up a couple of these.

Reviewer #3: The paper presents a generalized method to estimate, from tag-recapture data, both persistent and transient variation between individuals in growth rates. The proposed method seems to work well for toothfish, and for simulated data based on toothfish. The paper is interesting and well-written, and the topic is timely. I recommend publication.

153: Conventional ageing methods. Do you assume ageing is done without error? Is that reasonable for this species?

We’ve added a sentence here that this is done without ageing error and that this is a reasonable assumption.

166: Did you try any sensitivities with some other value of psi? Was this parameter important?

We tried estimating this parameter but generally found that models did not converge. When models did converge, psi was close to zero. Therefore, we fixed psi at zero in all other model runs. We have added additional sentences about these results and suggest that the exploration of this parameter be left for another paper.

174: "Standard errors are then estimated via the information matrix and delta-method." Some more detail on this would be useful.

Details added.

190: Delete extra "in the"

Done.

208: Where these parameters the same for males and females?

Yes, we re-worded slightly to try and make this clear.

209-210: You already said that psi was fixed at zero for the case study.

We’ve removed the second instance of this statement.

265-267: Estimate of L0. Judging from figure 1, there were very few data points for fish younger than 3. Was this true for the data-set used in Dunn as well? Differences between the results in this study and Dunn et al. are probably caused by differences in sample size of small fish rather than any difference in methodology.

Have added a sentence to reflect this, thank you.