**Reviewer: 1**

This is an excellent brief of standard method application and new method development. Both new and experienced professionals will be benefitted from reading this contribution in using standard terminology with simple English. The modified age-bias plots and the straightforward approach for partial age validation based on strong-year-class recruitment are broadly applicable.

My minor comments can be addressed at the step of production edits.

1. Title: consider to use “strong-year-class recruitment” to replace “recruitment variation”.

**RESPONSE:** We changed the title to “Age, year-class strength variability, and partial age validation of Coregonus kiyi from Lake Superior”.

2. Lines 83-84. >160 mm? Also, you mean fish >160 mm are rare, right?

**RESPONSE:** Rewrote this sentence to be more clear.

3. Line 92. Consider to use a new sentence to express when you used multiple images for a given section and why.

**RESPONSE:** Rewrote this sentence to be more clear.

4. Lines 418-419. Figure caption for figure 4. Remove (). Add “on” before “otolith-derived age estimates”.

**RESPONSE: Done.**

5. The captions for Figures 2 and 3. Not easy to read, and better rewrite. The title sentence should be just “A modified age bias plot for Lake Superior Kiyi”. Use a separated new sentence to explain the X axis, and the Y axis if you feel it is still needed. Then, continue to explain: the dashed horizontal line and the vertical bars. No need to explain the dashed horizon line twice. Use a new separated sentence to explain the indications of significant differences from zero.

**RESPONSE:** Done.

**Reviewer: 2**

Comments to the Author

I have reviewed the management brief titled “Age, recruitment variability, and partial age validation of Coregonus kiyi from Lake Superior” and find the manuscript to be concise, cogent and generally well written. I have a few specific comments that can easily be addressed to strengthen the manuscript.

Line 95: Here the authors state that a limited number of scales were aged because of the expected difference based on numerous other studies. While well-known differences in scales and otoliths have been shown for other species, including coregonines, I’m not sure this is sufficient enough a reason to age a limited number of fish from scales. As the authors state early on (line 49) the scale-otolith comparison has never been done. The assumption, or more specifically a null hypothesis, is that there is no difference between the two structures. Are scale-derived ages available for all fish aged from otoliths? If so I recommend using the full data set to your advantage and building a stronger case for otoliths over scales for another Great Lakes coregonine?

Why was only one reader chosen for scales? I would suggest providing data from both structures for both agers to be consistent in your evaluation of aging structures. This ties into the previous comment regarding approaching the structure comparison more objectively.

**RESPONSE:** We did not age scales from all fish for which otoliths were aged because much previous work on related *Coregonus* species from Lake Superior and other locations worldwide have shown scale ages to be consistently less than otolith ages. This has also been shown for other species worldwide and appears to be pretty universal. We cite some of these studies in the discussion. We assumed that the difference between these structures would be strongly different and, thus, we would not need a large sample size to detect this difference (while not formal, this is not unlike a power analysis prior to a study to determine the required sample size to detect an effect). Thus, with time and budgetary constraints, we put minimal resources into answering this question and directed more resources into assessing questions related to otolith-derived age estimates. We feel that our approach for comparing ages between scales and otoliths for one region of Lake Superior is adequate to address the question of how estimated ages from these two structures compared because the difference between these structures was so strong. The results are compelling, convincing, and consistent with previously published results. Thus, we think that the question about how scale and otolith ages compare for Lake Superior Kiyi has been answered with the amount of resources and effort we used in this study.

Lines 136-138: I think the approach of using length-frequency analysis here to determine ycs from the previous year is reasonable, however there are some assumptions being made regarding Kiyi ages and the first discernable mode being considered age-1. This is mainly because no fish of age 2 or 3 were aged. Are these age classes lumped into the first “age-1” mode? Somewhere in the discussion the authors should discuss possible bias in using the length frequency approach and assuming all these are age-1 fish.

**RESPONSE:** In Lake Superior and historically in the Lake Michigan, Kiyi were approximately 100 mm about one year after hatching. To support this we cite Deason and Hile’s work on Lake Michigan. This length is a bit less than, but similar to that observed for cisco and bloater, two related species. More age work has been done on these species, so that work could be cited. We could also cite unpublished data from our laboratory, but we prefer to not cite unpublished data. However, we don’t feel that these citations strengthen our argument and, thus, have not included them.

Furthermore, we are confident that age-2 or 3 fish are NOT lumped in with the age-1 mode because we can follow the growth of the age-1 mode in subsequent years on the series of length frequency histograms (current Figure 6). In addition, large increments on the otoliths from age-1 to age-2 and age-2 to age-3 suggest “good growth” such that these two older ages would most likely be separated from younger ages on a length frequency.

Given this, we think that the bias that this reviewer suggests is very unlikely. Thus, we have chosen not to modify the discussion as suggested.

Line 143-144: Here biomass and density estimates are presented. Are these ranges per trawl tow? Or by depth? Please specify.

**RESPONSE:** Changed sentence to say, Biomass and density estimates from individual trawl tows ranged from 0-12 kg/ha and 0-253 fish/ha, respectively.

Line 160: Please state minimum age estimated from otoliths as well.

**RESPONSE:** Done.

Line 205: I suggest rewording the lead-in sentence here to reflect that strong year classes are inferred from length frequency analysis. Something like “Distinct modes in length frequency distributions from 2003-2014 provided evidence for strong year classes, and partially validate our otolith-derived age estimates.”

**RESPONSE:** Done.

**Associate Editor Comments to the Author:**

This paper examined the accuracy of age estimates between otoliths and scales from a deepwater coregonine fish, the Kiyi, in the Great Lakes. The paper also attempts to use age-length keys derived from other species to conduct a ‘partial validation’ of the age estimates from the otoliths. Kiyi are not present throughout all of the Great Lakes, and unlike their closely-related congeners, the Bloater, are not as well-studied.

We received two reviews that, like myself, found merit in the paper. Reviewer #1 points out a few suggested edits, including some re-writes of the figure captions. Reviewer #2 points out some holes in the Methods that should be addressed with more rationale. I also found some problems with the captions and especially the Methods that I believe should be addressed to make this paper stronger. Several sentences could be bolstered with the inclusion of additional citations (especially recent).

I was a little surprised that more recent genetic papers have not been cited that indicate that the deepwater ciscoes of the Great Lakes are members of an assemblage of species that is often referred to as the cisco species flock, given that they interbreed significantly. Importantly, Koelz (1929) originally identified them as species (as the authors correctly note), but since then, the inter-breeding and ecological polymorphisms of these ciscoes has become apparent. How were the Kiyi identified? Could ‘hybridization’ with other ciscoes have occurred and their offspring be present in your samples?

RESPONSE: While interbreeding among deepwater cisco species in Lake Superior could and likely does occur to some extent, we do not think the “deepwater ciscoe hybrid swarm” concept described in Eshenroder et al. (2016) for Lakes Michigan and Huron applies to the current group of deepwater ciscoes present in Lake Superior. Kiyi generally are the easiest to distinguish from the other ciscoes and we are confident in our ability to identify Kiyi in the same way that they have been identified in the past. We did modify the manuscript to (a) acknowledge the concept of interbreeding for the decline of Kiyi in the other Great Lakes (related to your comment further below) and (b) describe the criteria we used to identify Kiyi.

An image of a sectioned otolith show annular and growth rings would be very welcomed and would increase the utility of the paper.

**RESPONSE:** We agree an image of a sectioned otolith would be nice to include and have added one as Figure 2.

In addition, I provide some detailed comments below.

Abstract: Please include ranges or SDs with means for ages.

**RESPONSE:** We can certainly provide this information if absolutely necessary. However, we don’t think that this is appropriate for the abstract, at least partly because none of these summary statistics relays much information about the distribution of ages due to the very uneven age distribution that results from the presence of strong and weak year-classes. We feel that the information that we do provide (maximum ages and the major modes (ages 5, 6, and 11)) relay more information about the distribution of ages then the mean, SD, or range would. Worse, we feel that the mean, SD, or range may paint a picture of the distribution that is not accurate (i.e., the mean may be an age that does not exist in high numbers in the distribution and providing the range suggests an even distribution across that range). Therefore, we have not included this request in the revised manuscript.

Lines 26-29: In a brief sentence or two, please indicate how strong year classes were determined, and thereby use them to validate otolith ages.

**RESPONSE:** Done.

31-33: Given the current state of knowledge, I don’t think Kiyi can accurately be referred to simply as a ‘species’, but rather, a member of a cisco species flock. See comments above.

**RESPONSE:** See our response to the first comment from the Associate Editor.

34-37: The hypothesis of alewives and smelt outcompeting Kiyi, thereby leading to its demise is only one of many, potentially interacting ecological occurrences that may have led to the decrease in Kiyi. I recommend the authors include some additional hypotheses in a few more sentences, and include additional citations. Only 1 old citation is used.

**RESPONSE:** There are no direct studies that we are aware of that have tried to evaluate the demise of Kiyi in the lower lakes. Christie (1974), with his thinking that the introduction of alewife and rainbow smelt is the reason for their demise, is most frequently cited. Overfishing is also mentioned in other Great Lakes fishery community reviews, but we know of no studies that have explicitly evaluated this for Kiyi. Eshenroder et al. (2016) does not shed any new light on the cause of the demise of Kiyi in Lakes Ontario, Michigan, and Huron, other than that several deepwater cisco species likely interbred and became an indistinguishable flock of species, that are presently known as bloater. We did modify the manuscript to explicitly identify that interbreeding may have been the cause in the other Great Lakes. If we can be provided other compelling citations we would include them.

That said, the focus of our paper is on ages and year-class strength variability in Kiyi, so it is probably not appropriate to spend too much text on the historical demise of Kiyi in the other Great Lakes. We are more inclined to eliminate this sentence then to augment it with hypotheses (for which references don’t exist, to our knowledge) that detract from our central thesis.

52-53: ‘fairly regular recruitment’—this phrase does not convey a lot of useful information. Please address.

**RESPONSE:** We feel that this is useful in the context of the next two sentences that juxtaposed “little annual variability” for the historical sample with “high inter-annual variation” of the modern sample. Regardless we include the parenthetical statement “i.e., recruitment was evident in most years” in the sentence in question.

56-58: ‘Lake Superior cisco recruitment’—noun train, please edit. (Cisco should not be capitalized in this occurrence, because it is referring to >1 species).

**RESPONSE:** Done, but note that Cisco is still capitalized here because reference is to only a single cisco species, *C. artedi*.

84: Sentence ending on this line—would benefit from the inclusion of a citation, personal observation, or personal communication.

**RESPONSE:** We have removed the reference to “historical collections.” We did consider historical collections when designing the study, but it is sufficient for the publication to just say that fish <160 mm proved to be rare in our samples.

85-86: Scale sampling area—differs from the preferred area for salmonids (of which coregonines are a part). Why did the authors sample this particular area? Please include appropriate citations.

Were the otoliths sampled on the boat or in a laboratory setting? –this would be useful information for others who are wondering how difficult it is to harvest otoliths in the field.

**RESPONSE:** This is the location scales have been removed from coregonines in our laboratory for 60 years and follows standard procedures for many other coregonines. While we consider this standard procedure that does not need reference, we added Hogman (1968) as requested (three other references related to coregonines are: NAJFM 28:625-635; NAJFM 28:1928-1940; Journal of Great Lakes Research 27:386-389). For this study, as written, length, weight, and sex were determined and scales and otoliths were removed in the laboratory. [We commonly (and easily) remove otoliths on the ship. We don’t think this needs to be stated here.]

87-95: Please include citations that guided the methods for the otolith preparation.

**RESPONSE:** We consider these general processes to be standard procedures. As with most studies, our procedures were slightly modified to what was most efficient for Kiyi. We believe that there is precedent in the literature (we reviewed several “otolith papers” in NAJFM in the last 15 years, including the two listed above) for describing the process used and not citing a specific source (as no one source has the exact procedure).

107-108: Please indicate how many otoliths were thrown out because a consensus could not be achieved.

**RESPONSE:** We view this as a result which is then reported in the second sentence of the age analyses section of the results. As stated there, the number excluded was 22.

108-110: Sentence is potentially misleading and not necessary. Reads as though only 1 reader was blind to biological information (did the other one have this information?), when earlier I thought it was indicated that both reads were blind….

**RESPONSE:** This sentence was referring to the one reader of the scales, not the two readers of the otoliths. We changed sentence to clarify this.

115: ‘significant bias’ = statistically significant? Please indicate.

**RESPONSE:** We assume that the word “significant” in a scientific publication means “statistically significant,” especially when in the context of a “test” as it is used here. We find the use of “statistically significant” to be redundant and clunky and would prefer not to include its use in the manuscript. Nevertheless we added the word “statistically” before “significant” here. We also assumed that this comment was not specific to this one line, so we added the word “statistically” in front of “significance” in four other places in the manuscript.

151: Were any of the 7.6% vateritic?

**RESPONSE:** No, we did not observe any vateritic otoliths.

Fig. 2. Please simplify and clarify the figure caption. If it is a modified age-bias plot, then indicate as much at the beginning of the first sentence vs. parenthetically later on. The bit about horizontal lines is confusing because it reads similar: ‘short horizontal lines’ vs. ‘the horizontal dashed line’. Consider replacing sample size numbers on the x-axis with an inserted bar chart. Please summarize the key conclusion from this figure. Please clarify the means vs. the ages that agree; this is also confusing.

Fig. 3. Similar comments here.

**RESPONSE:** Caption was clarified. We removed the “modified age-bias plot” parenthetical statement as we felt it was more important to clearly describe what was being plotted. We removed reference to the “horizontal dashed line” and just refer to “zero” (which was where the horizontal dashed line was). We clarified how to determine which ages were statistically in agreement by referencing the confidence intervals relative to zero.

We did not create a new bar chart with the sample sizes. We feel that presenting the sample sizes in this plot is important as it helps to understand the relative widths of the confidence intervals (i.e., some of the wide confidence intervals are due to small n). However, the samples sizes here are not important in the larger context of the paper. The sample sizes that are important in the larger context are provided in Figure 5 (which was Figure 4 in the original submission). If the sample sizes on these figures cause considerable angst to the reader, then we would prefer to just remove them rather than make another figure with just those sample sizes.

We also slightly edited these figures. We changed the zero line to a solid, but light gray line (rather than black dashed line) and made the mean and confidence interval “lines” narrower. We added the word “age” into the x- and y-axis labels of Figure 3 (was Figure 2 in the original submission).

Fig. 4. What is the source of the age-length key being referred to in the caption? This caption would benefit from this information so that the figure can stand alone from the text. This also raises my curiosity for size-at-age for Kiyi.

**RESPONSE:** We feel that this is currently well-described in the manuscript. The figure legend notes that the ALK was based on consensus otolith-derived age estimates. In the main text, we note the same thing and include the specific R function used to apply the ALK. Specifically, as written in the last paragraph of the age analyses section: “An age-length key (Fridriksson 1934; Ketchen 1949) was constructed from consensus otolith-derived age estimates. The age-length key was then used to assign specific ages to all Kiyi captured in 2014 using the method described by Isermann and Knight (2005) as implemented in the alkIndivAge function from the FSA package.” The level of information in the main text does not seem appropriate for the figure caption.

We had considered adding size-at-age information to this manuscript. However, because there are so few ages, due to the periodic production of strong year-classes, present in our 2014 samples, the size-at-age information is sparse (i.e., only a few mean lengths-at-age can be estimated when only a few ages are present), complicated by minor errors in ageing, and not particularly informative. Thus, we chose not to present that information. The interested reader can garner some information about size-at-age from current Figure 6.

Fig. 5. I am not clear on what the authors are trying to convey with regards to the numeric labels for 2004, 2006, and 2010.

**RESPONSE:** We feel that this is described in the figure caption (i.e., “The numeric labels in 2004, 2006, and 2010 are the age of fish in those modes in 2014”). In other words, those numbers indicate the age that the age-1 fish in each of those sample years would be in our 2014 sample. These ages are referred to in several places in the manuscript and are highlighted here to tie together our 2014 sample with the historic length frequencies, and the strong year-classes determined from the historic length frequencies and the ages we identified in fish sampled in 2014. Without a more specific statement about what is not clear to the editor, we cannot further clarify this point.