**Editor’s comments:**

-- Please ensure that you follow the specific suggestions by one of the reviewers to state some of the conclusions by Frank et al in a more 'neutral' tone.

**We followed the suggestions of Reviewer 1 and have removed reference to Frank et al. (2016) when referring to the non-collapse hypothesis and we have made sure that conclusions of non-collapse hypothesis are stated in a neutral tone throughout MS.**

-- Please reorganize your manuscript to adhere to the structure of a Research Article (IMRAD), i.e. do not combine the Methods and Results. I regret asking for this change, as it is very labour intensive, but the benefits will outweigh the cost. Dr. Mueter comments regarding this point: "Personally, I prefer a separation of methods and results. [...] Separating methods from results in this case could help clarify what are new analyses versus a review of published results and would also help clarify some of the methodological issues raised by reviewers." In addition, separating this combined section will avoid the impression that this manuscript is there to prove Frank et al. wrong, avoid the irritation of having a Methods section full of references to other researchers' results and allow other researchers to understand and replicate your work more easily.

**We have separated the Methods and Results section as recommended (Methods: L106-176; Results:L178-549)**

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

In addition to the above, please address the following formatting points by the in-house editors in your revision:

(i) Please include postal codes for all author affiliations.

**Postal codes have been added for all authors**

(ii) Please note that the number of references you cite is relatively large (125) and could actually be considered excessive in relation to the length of your manuscript. Please check whether all references are necessary and avoid 'strings' of citations (more than 3 cites in a row), as long as this does not compromise your scientific message. Such strings make the text very tedious to read and give the impression of a literature review rather than the selectivity expected in a scientific article. A common length of a literature cited section for a research article of this length would be 50-75 citations.

**We have reduced citations by 24 (total: 101)**

(iii) However, as you use the free open-source software R, it would be appropriate to include a reference to the version of R that you used (i.e. R Core Team YYYY). It is important to at least note the version number.

**We have added R citation (L161)**

(iv) In the Literature Cited, please include the city of publication for books and reports, and also include page ranges when citing specific sections.

**We have added city of publication and page ranges in bibliography.**

(iv) On the title page, please remove the following text from the author-contribution footnote: "and others have contributed equally. Authors have been listed in a alphabetical order for each contribution level.” This part is too vague to be informative and we (the publisher) do not see how the degree of contribution of different authors (apart from the first author(s)) is of interest to the readership.

**We have removed detail about alphabetical order of co-authors.**

(v) In your next, and hopefully final revision, in addition to the low resolution figures in the manuscript, please provide also high resolution figures preferably in editable source files (pdf, eps, ai, psd or tif (LZW compression) for optimal print production. You should provide individual files for paneled figures (for maps or contour plots) to avoid loss of image quality.

**All figures have been attached in high resolution (eps)**

**Reviewer # 1**

First, the center of gravity (COG) analysis using the VAST package in R requires a more in depth explanation of what was actually done and how you assessed model fit. The paper describes the analysis as essentially inputting survey data into a virtual black box that outputted COG estimates. The geostatistical delta GLMM developed by Jim Thorson is a relatively sophisticated model that has not been widely used, and therefore requires a more in depth explanation of the model structure, how it was parameterized, and how model fit was assessed. None of this information is reported in the document even though it’s an original analysis, and therefore it is not possible to replicate the analysis. Based on the paper format, I would be satisfied with the detailed descriptions of the model structure and diagnostics being reported in the supplementary materials, with just a few extra sentences in the manuscript that summarize key points. There are published examples in the literature that can be used to gauge the level of detail needed to adequately describe what was done (Thorson keeps a running list of publications on his GitHub that have used geostatistical GLMMs with or without VAST). As an aside, it seems a bit overkill to calculate COG from a geostatistical GLMM using VAST since the abundance estimates are not used and no catchability or environmental covariates are included in the model to improve COG predictions. A less complicated approach would be to calculate COG and inertia (cf. Woillez et al. 2007) from the survey data using the SI.cgi() function from the Rgeostats package in R. If your goal is to demonstrate that observed shifts in capelin are not biased by changes in the survey domain, a more convincing approach could be to calculate COG for the survey itself by weighting each sample equally (although I’m not sure VAST will allow this – if not, use the SI.cgi() function in R), then comparing interannual changes in COG for the survey with capelin distributions (could do it for both capelin presence samples using a binomial spatiotemporal GLMM in VAST & those weighted by capelin CPUE using the full delta GLMM). Just a suggestion… personally, I like using VAST too.

**We added more details on COG analysis in Methods section (L140-162) and have included diagnostics in the Supplementary section (Figs. A1-A3).**

The reader would benefit from additional figures included in the manuscript. One figure per section isn’t necessary, but it’s not clear why so few figures are used to summarize these analyses (refer to detailed comments below). I have suggested revisions to each of the figures and Table 1 to improve clarity and aid a reader with little background knowledge of this system.

**We have taken this recommendation onboard and included 2 additional figures (Figs 5, 6) and we have improved Table 1 for clarity.**

In general, consistently refer to the collapse hypothesis and non-collapse hypothesis throughout document, rather than the “hypothesis of a non-collapsed capelin stock.” It is also unnecessary to cite Frank et al. (2016) when referring to the non-collapse hypothesis – it’s very clear to the reader who you are referring to.

**We have changed wording in MS for consistency (collapse and non-collapse hypotheses) and have removed the citation to Frank et al. (2016) when referring to the non-collapse hypothesis.**

Abstract

L22-24 Not necessary to specify the indices that were used to show the stock collapsed. Recommend just saying the DFO determined the stock suffered an order of mag decline…

**We changed as suggested to “Fisheries and Oceans Canada determined that the Newfoundland capelin stock…” (L24-26)**

L27 Rather than say “there is also literature suggesting”, change to “an alternative hypothesis has been proposed that suggests…”

**We changed sentence as suggested to “ …, an alternative hypothesis has proposed that rather than collapsing …” (L28-29)**

L36 Suggested revision “The weight of evidence approach led us to reject the non-collapse hypothesis and conclude that the Newfoundland capelin stock did collapse in 1990-91 with minimal recovery over the subsequent 25 years.”

**We revised final sentence as suggested (L36-38).**

Intro

L63-64 There is evidence that top-down processes have also impacted capelin recruitment in the Barents Sea (Hjermann et al. 2004, 2010)

**We have included top-down forces in sentence and included suggested references “…ecosystem changes, both bottom-up [..] and top-down forces, were behind…” (L62-64)**

L73 Define NAFO

**We defined NAFO (North Atlantic Fisheries Organization) (L73)**

L73-74 Topic sentence is unrelated to paragraph, recommend moving and starting with 2nd sentence

**We deleted opening sentence as suggested. The paragraph now starts with “ Fisheries and Oceans Canada (DFO) concluded that the NAFO (North Atlantic Fisheries Organization) Divisions 2J3KL capelin stock (hereafter Div. 2J3KL capelin stock) experienced an order of magnitude decline in 1990…” (L73-76)**

L80-81 Do you need 9 references to say there was a regime shift?

**We have reduced the number of references to 6 in this sentence (L79-80).**

L89 Delete “Some authors do not feel” and explicitly state “Frank et al. (2016) have presented an alternative hypothesis that the… stock did not collapse in the 1990s.” And then delete all (Frank et al. 2016 citations from rest of paragraph

**We changed sentence as suggested “Frank et al. 2016 have presented an alternative hypothesis…” (L87-88) and deleted all remaining references to Frank et al. 2016 in paragraph.**

L90 Delete “This non-collapse…” and replace with “The non-collapse…”

**We replaced “This” with “The non-collapse…” (L88)**

Methods and Results

L110-111 Replace “weighting evidence in” with “summarizing support”

**We changed “weighting evidence in” with “summarizing support” as suggested (L172).**

L137-138 Since you quantify the shift in COG using bottom trawl data in next section, maybe move this comment “southward shift in stock distribution…” to that section where you evaluate the COG results

**We removed reference to “southward shift in stock distribution” in Offshore capelin abundance: acoustic surveys section. The southward shift in distribution is now referred to only in the following COG section (L260)**

L146 Looks like it dropped two orders of magnitude

**We changed to “…decreased two orders of magnitude” (L203)**

L149 Insert range of biomass estimates from recent increase “acoustic survey from XX to XX mt,”

**We have inserted range of biomass estimates (660-980 thousand tonnes in 2013-2015) (L205)**

L154-155 What does “qualitatively estimated densities” actually mean?

**Changed sentence to “The non-collapse hypothesis uses observations of overwintering maturing capelin…” L211.**

L195 Seems like “impossibility” is a big jump from “unlikely” in L189, revise to “highly improbable”

**We have revised sentence to “…the sheer number of fish required to compensate for the ‘missing’ 3-6 Mt of capelin offshore (~40,000 to 170,000 kg km-2) is staggering and was never detected by the seasonal inshore acoustic surveys” (L248-50)**

L201-203 Combine and revise these two sentences as a single topic sentence

**We revised the opening two lines of this paragraph: “The westerly, inshore shift in the center of capelin concentration in 1996-2010 compared to 1985-1995 was considered support for the non-collapse hypothesis.** **However, the center of concentration of capelin from 1985-1995 was inshore in three high abundance years …” L 254-257**

L217-219 See general comments above VAST

**Based on Reviewer 1’s comments, we re-calculated the centre of gravity (CG) and inertia (cf. Woillez et al. 2007) for the years 1983-2016 from the survey data using the Rgeostats package. The CG was calculated by year and its mean coordinate was weighted by abundance. The CG was relative within each year regardless of the probability of catching capelin. Details are in Methods section (L140-162), Figure 4, and supplementary figures of model diagnostics (Figs. A1-A3).**

L224-225 Move “(i.e. Campelen 1800 shrimp trawl)” to after “one gear type”

**We are now including capelin presence/absence data from both the Engel and Campelen trawls so this sentence has been deleted (L158-62).**

L225 Need to specify the model used (e.g. what error distributions and link functions were used in the delta model) and provide some info on model specs (e.g. number of spatial knots, did you use a spatial mesh or grid, are the intercepts for each year fixed or was autocorrelation included from year-to-year, did you assume spatial relationship was isotropic and allow for anisotropy). For example’s, go to Thorson’s Github page and search for running list of recently published papers that have used VAST. It’s an original analysis, therefore you should also include how you assessed model fit – much of this can go in supplementary materials, but it has to be reported somewhere

**We have included diagnostics for our new CG model in the supplementary material (Figs. A1-A3). Details on model are in Methods (L140-62).**

L237 Should add a comment that capelin were still vertically distributed above the acoustic deadzone, otherwise this opens up uncertainty that the decrease in acoustic estimates may have been partially attributed to sampling bias related to reduced capelin availability to the acoustics

**We added “Since 1991, capelin diel vertical migratory behavior has changed with deeper distributions below the cold intermediate layer (~ 200 m) and in closer association with the bottom (peak density ~ 20 m off bottom) with capelin rarely found in the upper 50 m of the water column (Mowbray 2002).” (L270-73)**

L271-272 Rephrase; it’s not clear what “earlier observed maturation of capelin” means. Is it earlier in the year, or is a greater proportion of younger fish (i.e. age-2) maturing a year early? Can you use a citation from Iceland or Barents that early maturation occurs when abundance is low?

**We merged 3 sentences to clarify the topic sentence of this paragraph “While a delay in the timing of spawning should not have a direct impact on the acoustic abundance index derived from the spring acoustic survey, earlier maturation could alter the age structure of the stock with inter-annual variability in the proportion of age-2s starting their adult migration patterns and would, therefore, be unavailable to the survey. The proportion of mature age-2 capelin has increased since 1991 (4% pre-1991 compared to 37-79% post-1991) (Mowbray 2014; DFO 2018).” (L302-307).**

**We have now included literature on capelin growth rates during the ‘busts’ in the Barents Sea. Gjøsæter (1998) reviewed the literature and there is evidence that capelin growth may be stock abundance dependent with faster immature growth when the stock size was small. We may be seeing the same phenomenon in NL (L307- 312).**

L277-281 Break up long sentence

**We have broken this sentence in two at the “which” (L312-315; 315-317)**

L285-288 These proportions should be reported in a figure, and L286-287 is awkward

**We have included a figure on the change in commercial fishery age composition from 1980-2017 (Fig. 5). We deleted second half of this sentence: “Post-1991, the contribution of mature age-2 capelin increased to almost half of commercial inshore catches (Fig. 5). However, the age structure of offshore catches are not consistent with the non-collapse hypothesis…” (L321-325)**

L292 Was the increase in % of age-2 fish maturing offshore actually observed?

**Yes, DFO technicians age and maturity stage capelin from offshore samples collected during the May acoustic survey. While the May acoustic survey does not target migrating capelin, maturing capelin (age-2 and 3) are sampled and processed. The age of maturity has changed since 1991 (see DFO 2018).**

L299-301 Awkward sentence, consider revising

**We’ve changed order of first 2 sentences of this paragraph and changed a couple words in second sentence: “The two inshore indices collected by DFO during the 1980s and 1990s were an aerial abundance index and inshore commercial catch rates. The non-collapse hypothesis implied that strong correlations between the two independent inshore indices post-1991 in addition to no obvious decrease in these indices between the late 1980s and the early 1990s was supportive of a stable stock.” (L334-338)**

L355 Replace “cannot” with “should not”

**We replaced “cannot” with “should not” (L369)**

L343-346 Data should be presented in a figure

**We have presented capelin larval data in Fig. 6.**

L394-396 Rephrase topic sentence. Suggest “… support the non-collapse hypothesis as Atlantic cod condition is associated with spatial overlap of capelin. Atlantic cod condition was worse in Div. 2J3K where capelin were no longer observed post-1991 compared to cod in the south (Div. 3L) where the capelin population had contracted.”

**We have changed final paragraph in the Atlantic cod section as suggested (L429-433).**

L408 “invasion” suggests a spatial movement that was not described

**Changed wording from “invasion” to “occurrence” (L445)**

L410-412 Again, would be helpful to show these data rather than refer reader to Stenson 2012. Also, recommend restructuring sentence order. Suggest “…In the Northwest Atlantic, Arctic cod and Atlantic herring were available as prey to harp seals during the early 1990s (Lilly et al. 1994, Bourne et al. 2015, DFO 2017), as indicated by the large proportional increase of these alternate prey in harp seal diets during this period…”

**While we acknowledge that having all figures on hand would be useful, in order to keep our paper tightly focused on capelin, we have decided to only include new capelin figures in this paper rather than re-print figures from source material. We have changed sentence as suggested (L446-449).**

L462 Replace “warmer and the” with “warmer, the”

**We have changed sentence as suggested (L508).**

L463 Delete “, respectively”;

**“Respectively” was deleted (L499).**

L463-465 Rephrase, suggest “In addition, the lack of capelin in seabird diets during the 1990s in Labrador (Bryant & Jones 1999, Baillie & Jones 2004) is inconsistent with the non-collapse hypothesis.”

**We have rephrased as suggested (L500-502).**

L466 Topic sentence should be revised to state something along the lines that increases in seabird populations post-1991 were related to factors unrelated to availability of capelin as prey. Then give examples of closures of gillnet fisheries and pollution/hunting

**We have revised topic sentence as suggested: “Increases in seabird populations post-1991 were related to factors unrelated to availability of capelin as prey. For example, the population increase of common murres post-1991…” (L503-510)**

L468 Delete “They considered” and begin sentence with “The removal”

**We replaced “They considered” with “The removal…” (L506).**

L481 Suggest replacing “local abundance” with “availability” since predator diets do not necessarily reflect prey abundance

**We have replaced “local abundance” with “availability” (L517)**

L482-498 Your argument appears to be that CPR data does not accurately represent changes in copepod abundance, but takes a while for you to get to make that point. This section would benefit from some editing to more clearly state why CPR data is not a good indicator of copepod abundance (pretty sure the studies from the Barents used net samples from extensive survey effort, not surface measurements from vessels of opportunity, to demonstrate a negative relationship between zoop abundance and capelin – that seems important to note)

**We have edited this section as suggested: “Given the magnitude of the capelin collapse, a significant increase in their main copepod prey, Calanus finmarchicus, might have been expected (e.g., Gjøsæter et al. 2002). Unlike the extensive zooplankton sampling effort in the Barents Sea, zooplankton on the NL shelf (Divs. 3KL) have only been sampled on a regular basis using plankton nets since 1999 (Fisheries and Oceans Canada’s Atlantic Zonal Monitoring Programme). The Continuous Plankton Recorder (CPR) dataset was used to test the non-collapse hypothesis as it commenced in the late 1950s. However, the usefulness of CPR data in tracking changes in copepod abundance in the Northwest Atlantic has been questioned.” L519-526**

L488-490 Frank et al. 2016 said they only included years with 11-12 months of observations, but you refer to only 2 years with 8 or more months of data between 1960-1978, and later indicate there’s a gap in CPR data from 1979-1990 on L496 – does that mean there was no data between 1979-1990, hence only 2 years between 1960-1990?

**Frank et al. don’t provide details on how many years met their minimum requirement of data. Years with gaps in data were included in their analysis using a mean value. Based on our literature search, there are significant gaps in this dataset for this region and it’s likely they only used 2 years of data from 1960-1978 and potentially up to 6 years of data from 1979-1986.**

L504-512 Long sentence is awkward, break up and revise

**We have revised sentence: “Elsewhere in the literature, 1991, not 1996, has been identified as climatologically important decadal signal due to its strong cold anomaly […]. The cold anomaly was associated with a regime shift in the North Atlantic ecosystem…” (L539-547).**

L513-517 Revise this statement. This study’s analyses were intended to support the hypothesis that the capelin population collapsed after 1991, not to identify processes that led to the collapse. It’s reasonable to say this study’s findings show evidence of a collapse that coincides with a regime shift, and are consistent with other studies that have attributed ecosystem-wide changes to the cold water anomaly… or something along those lines

**We have revised final summary statement as suggested: “In summary, this study’s findings show evidence of a capelin collapse that coincides with a regime shift, which is consistent with other studies that have attributed ecosystem-wide changes to the cold water anomaly.” (L548-550)**

Discussion

L522 It’s not clear from your analysis that “bottom-up, climate-driven” processes led to the collapse – not disputing it, but it’s not what this study examined nor was the topic discussed until the prior section that needs to be revised. For this paragraph, focus on the weight of evidence that supports the collapse hypothesis. Maybe in the last sentence you could insert a comment that the collapse coincided with the regime shift.

**We have revised Discussion as suggested: “The weight of evidence led us to conclude that the Div. 2J3KL capelin stock suffered a population collapse in 1990-91 with minimal recovery in the subsequent 25+ years.” (L553-54). We have decided not to include a sentence on the regime shift in the final sentence of Discussion.**

Table 1 – The table should be re-organized to list each section heading from the Methods and Results – above all, it should fit on to a single page since this summarizes your findings. Only list individual indices under a section if there are differences in the direction of support (e.g. its unnecessary to list all 4 offshore acoustic surveys). Not crazy about your choice of arrows to symbolize direction of support since they typically represent increasing or decreasing trends, not positive or negative support. + & - symbols seem more intuitive (bold for “support in favor/against”, non-bold or smaller font size for “consistent/inconsistent”), along with different symbols for “equivocal” (=) and “cannot be used…” (x)

**We have revised Table 1 as suggested:**

* **We reduced the number of items in the table to match those of the headings in the Results, adding an extra item “Analysis of ‘missing’ capelin distributed inshore”, as we feel this is an important piece of evidence in support of the collapse hypothesis**
* **We changed the symbols used as suggested by the reviewer**

Fig. 1 – Cruise tracks are indiscernible as gray-scale, recommend using colors with greater contrast for surveys that overlap (e.g. Fall Canada = med-dark color, Fall Russia = light color, Spring Canada = dark color, Spring Russia = med-light color, Trinity Bay = darkest color) and do not use red and green in same plot. Extend the x-axis to 44°W and add labels for Grand Banks, Scotian shelf, and Flemish cap since these features are repeatedly referred to in the text and their locations may not be known to reader’s unfamiliar with this region

**We have changed colours on map to viridis ramp. The contrast of these colours will be discernable in grey scale. We have labelled the Grand Bank, Flemish Cap. We have extended map to 44°W.**

Fig. 2 – This figure is difficult to interpret, especially in grayscale. I recognize why the points for each series are not connected by a line (that’s my personal preference as well when there’s a break between years), but 5 sets of points makes a busy figure. At a minimum, increase the symbol size, consider alternating open and closed symbols for overlapping series (suggest using one color & shape for Canadian offshore surveys, another for Russian surveys, and then use open symbols for spring and closed symbols for fall surveys. Use a third color/shape with closed symbol for Trinity Bay), state what the errors bar represent and comment why errors are not available for the other data series.

**Like Fig. 1, we have changed the colour scheme to viridis ramp. Figs. 1, 2 use the same colour scheme for the corresponding survey areas. We have connected the points for each series as suggested. We have also increased the symbol size, alternated open/closed symbols, and use a different shape and colour for inshore Trinity Bay survey. We have included a sentence on what the error bars represent (“The vertical lines indicate 95% confidence intervals of the index. A Monte Carlo approach to calculate confidence intervals could only be used on more recent data (1988 onwards) (Mowbray 2014)”) (L610-612)**

Fig. 3 – For barplots B & C, change color scheme to avoid red and green in same figure. It’s difficult to interpret age-compositions in grayscale, so recommend using a color palette with greater contrast (e.g. gray80, gray60, gray40, gray20, and black should work just fine). Can also reduce the y-axis labels to intervals of 0.2 so it’s less busy

**We have changed figure as suggested using the viridis ramp colour scheme. We have reduced y axis to fewer intervals so it’s less busy.**

Fig. 4 – I like the idea of showing areas not covered or poorly covered by the surveys, but its misleading because it does not show interannual changes in the survey domain which is more important when reporting shifts in COG. If your goal is to demonstrate that observed shifts in capelin are not biased by changes in the survey domain, a more convincing approach could be to calculate COG for the survey itself by weighting each sample equally (although I’m not sure VAST will allow this – if not, use the SI.cgi() function in R), then comparing interannual changes in COG for the survey with capelin distributions (could do it for both capelin presence samples using a binomial spatiotemporal GLMM in VAST & those weighted by capelin CPUE using the full delta GLMM). To visualize your results, recommend adding inset plots of northings and eastings for survey & capelin COG in a 2x2 (or 2x3 if you include capelin presence COG) multipanel plot within or under the map to clearly show how movement from north to south was greater than east to west, and that interannual shifts in capelin COG were not greater than corresponding shifts in survey COG.

**We have addressed Reviewer’s 1 comments on this analysis by using a different approach based on Woillez et al. (2007) (detailed in Methods L140-162) and have included a supplementary section with figures that provide diagnostics of the model (Figs. A1-A3).**

References

Hjermann DØ, Bogstad B, Dingsør GE, Gjøsæter H, Ottersen G, Eikeset AM, Stenseth NC (2010) Trophic interactions affecting a key ecosystem component: a multistage analysis of the recruitment of the Barents Sea capelin (Mallotus villosus). Can J Fish Aquat Sci 67:1363–1375

Hjermann DØ, Stenseth NC, Ottersen G (2004) Indirect climatic forcing of the Barents Sea capelin: a cohort effect. Mar Ecol Prog Ser 273:229–238

Woillez M, Poulard J-C, Rivoirard J, Petitgas P, Bez N (2007) Indices for capturing spatial patterns and their evolution in time, with application to European hake (Merluccius merluccius) in the Bay of Biscay. ICES J Mar Sci 64:537–550

**We have included suggested references in bibliography.**

**Reviewer # 2**

In most of the lines of evidence discussed in the paper there is room for a certain discretion. One example would be the physical variability. Frank presented a PCA and argued that a major shift in environmental change happened about five years after the “collapse”. Given the parameters included in that PCA, that conclusion was probably valid. However, other authors have argued that the most markable shift in oceanographic features happened earlier, about at the same time as the capelin disappeared from the surveys. It is not obvious to a reader which parameters should be considered in this case, and why one analysis is better than another, even though a large majority of experts seems to favor that a dramatic regime shift happened around 1991. Also, in other lines of evidence it is not obvious why the present interpretation of the data is truer than that presented in Frank.

**We have revised the Physical Variability section based on comments from Reviewer 1 and 2. While Frank’s conclusion may be valid, based on the data used in their PCA analysis, we have presented a review of the literature that supports the hypothesis that the capelin collapse was associated with the cold anomaly and regime shift in the early 1990s, which corresponded with the collapse of other commercial and non-commercial species (L537-550).**

The argument discussed at line 541-549 seems to me to be one of the strongest in favor of the collapse hypothesis: It is difficult to accept that 3-6 Mt of capelin distributed inshore at mean densities of >40 tonnes per square km, would have gone unnoticed for so many years. The lines of evidence discussed may not be totally convincing when seen in isolation, but it seems to me that the authors in totality have a “good case”.

**We agree that the analysis of inshore capelin biomass required to make up for 3-6 Mt of missing offshore capelin is the strongest line of evidence that capelin did collapse, and when it’s presented with seasonal inshore acoustic data, centre of gravity analysis, offshore acoustic data, larval data, and predator diet data, there is strong support for the capelin collapse hypothesis. Also, please note that we have corrected an error in our presentation of the capelin densities from the inshore survey. Originally, we included the biomass estimates (~10,000 kg in June) rather than the density estimates (biomass/ area of survey). This is now presented correctly in MS and Figure 3a.**

The paper is well organized, and the line of arguments easy to follow. The references to relevant literature may seem a bit overwhelming in some cases but are probably needed. The authors seem to have a good overview of the relevant research and literature in their field, both for “their” capelin stock and capelin stocks elsewhere.

**We have reduced the references by 24 (so far).**

Some detailed comments:

Line 10: Progamme = Programme

**We have corrected misspelling (L11)**

Line 57: Lavigne 1996 missing in reference list

**We have deleted Lavigne 1996 from MS**

Line 571: Miller and Carscadden 1983 missing in reference list

**We have included** **Miller and Carscadden 1983 in reference list (L807-88)**

Line 753: Lilly GR (1987) not cited

**We have deleted Lilly 1987 from reference list.**

Line 807: Mowbray F (2012) not cited

**We have deleted Mowbray 2012 from reference list.**

Line 949: Winter GH, Carscadden JE (1978) not cited

**We have deleted Winter GH, Carscadden JE (1978) from reference list.**

**Reviewer # 3**

General Comments:

1. The Introduction is a nice description of the population-level changes of capelin in the north Atlantic, but could expand on some of the reasons for these changes. For example, the authors equate “ecosystem changes” with “bottom-up forces” (Lines 63-64). I think top-down forces are also included in ecosystem changes. Also, there is also no description of what bottom-up forces are at play. I can think of many, but which of those are important here?

**We have included top-down forces in the description of ecosystem changes (L62-65). We have included a description of bottom-up forces (temperature and North Atlantic Oscillation) (L63)**

2. Following on that comment, I suggest stating the conclusion “up front” in the Introduction. I needed to read the manuscript at least twice to understand the overall argument and conclusion because I was caught up in the details the first read, and those subsections often seemed disjointed. Having the “end game” laid out at the beginning would, I think, help guide the reader through the authors’ arguments detailed in the subsections. Maybe refer to Table 1 at the beginning so that the reader can follow the argument in text and symbol.

**We have included a final sentence in Introduction that outlined the method we used in our analysis (triangulation) and that our analyses supported the collapse hypothesis, with reference to Table 1:** “…**collapse hypothesis versus the non-collapse hypothesis using multiple, independent data sets and diverse statistical methods (e.g., triangulation, sensu Munafò & Davey Smith 2018). The collapse hypothesis was strongly supported by the combination of previously-published results and additional analyses based on independent data sets (Table 1).” L99-103**

3. Overall the authors provide convincing evidence to support the capelin collapse. My comments are not necessarily meant to counter that argument, but are meant to help clarify the authors’ arguments and provide alternative perspectives that may be used in further rebuttals.

**Thank you for your comments; they have proved useful to us in tightening our argument.**

4. The authors provide convincing evidence of a capelin collapse. But to play devil’s advocate:

1. I have been on the Barent’s Sea on a vessel with echosounders running and at night the capelin “disappear” from the echosounder because they have risen to the surface and are in the near-surface blind zone of the transducers. I am not familiar with the acoustic surveys reported here and I wasn’t able to find a protocol document for the surveys (i.e., 24-hour or day only or night only survey), but this behavior needs to be accounted for.

**We have included additional information on the acoustic surveys. Acoustic surveys are conducted 24 hours a day (L120). While capelin may be missed in both the bottom dead-zone (4m) and transducer blind-zone (16 m) near the surface, the diel vertical migratory behavior of capelin has been deeper since 1991 and capelin were rarely sampled in the upper 50 m regardless of time of day (Mowbray 2002) (L277-82).**

5. Table 1 needs to be cited well before the Discussion. I suggest citing in the concluding statements for each Methods and Results subsection.

**We have cited Table 1 at the end of each Results subsection**

Specific Comments:

1. Line 58. Expanding on “important” would help strengthen the justification for the study. Do the authors mean ecologically, economically, combination of those, others?

**We have expanded on “important”: “The three largest and most economically valuable capelin stocks in the North Atlantic…” (L57)**

2. Line 89. I’m not sure “authors” is the best description of those who support the “non-collapse” hypothesis. Maybe a rewording to something like “A counter-hypothesis has been proposed that the Div. 2J3KL capelin stock has not collapsed (Frank et al. 2016).” Or make it explicit “Frank et al. (2016) do not feel ...” or my wording ...” by Frank et al. (2016).

**Based on comments from Reviewer 1 and 2, this sentence has been revised: “Frank et al. 2016 have presented an alternative hypothesis that the capelin stock did not collapse in the 1990s. The non-collapse hypothesis postulates that the stock did not collapse in 1990-91…” (L88-91)**

3. Figure 1. The bathymetry isopleths are a bit light. It would be helpful to have them a bit darker.

**We have darkened the bathymetry lines in Figure 1.**

4. Line 106. Figure 1 doesn’t really show the “test of the hypotheses”. I’m not sure Figure 1 should be cited here. Maybe after “independent data sets”?

**We have deleted reference to Figure 1 in this line.**

5. Line 185 and line 543. Does the argument rely on “uniformly distributed” capelin? As a back-of-the-envelope I understand why you make the assumption of uniformly distributed, but does your argument fall apart if the capelin are schooling (which they do)?

**We have stated that our estimate is a minimum estimate of capelin density uniformly distributed rather than taking into account capelin aggregations (L132-38). However, when we compare our estimate to the seasonal biomass data collected in months outside of the spawning aggregations, when capelin are disperse due to feeding behavior and may be more uniformly distributed, the biomass estimates of capelin is only <1 kg km-2 compared to the required ~41,000 kg km-2 to explain the missing 3-6 Mt (Fig. 3a).**

6. Line 197. The fact that there were 100s of echosounders does not support any argument without reporting what those echosounders recorded or the captains reported.

**We have revised sentence: “…remained undetected in the inshore strata since 1990 by DFO and fish harvesters,…” (L579)**

7. Lines 202-214. Were any “calibration/comparison” tows done when the bottom trawl gear changed? If so, does the comparison provide any indication of a major change in catchability? I was left wondering what level of “improved catchability” there was and how this influenced interpretation of the capelin catch data.

**Conversion factors between the two gear types do not exist for the majority of species, including capelin. We are unsure of what level there has been improved catchability of capelin, except that the Campelen shrimp trawl (with smaller mesh) catches more small fish so the tows are now shorter duration and slower speed compared to the Engel trawl (L158-162)**

8. Lines 225-226. The authors’ geostatistical analysis can not say anything about 1990-1991 because the data start in 1995. The analysis is valid for subsequent time periods. Referring to comment 7, were there any corrections used in the capelin index for pre-1995 that may shed light on this?

**There is no conversion factor available for the catchability of capelin between these 2 trawls. However, based on comments from Reviewer 1, we have used an alternative analysis based on Woillez et al. 2007 (L140-162) that allowed us to include the entire fall bottom trawl survey data which allows us to say definitively that capelin centre of gravity was offshore in 1991.**

9. Is it possible to do the center-of-gravity analysis on the acoustic data?

**The centre of gravity analysis only works with the fall bottom trawl data due the spatial coverage of the data. The spring acoustic data is only in Div. 3L. Furthermore, we wanted to test the non-collapse hypothesis using the same data that Frank et al. used.**

10. Line 287. Replace “if” with “of”.

**This sentence has been changed as suggested by Reviewer 1.**

11. Lines 337-347. The first sentence of the paragraph states larval production did not decrease, but the last sentence states a substantial decrease in larval productivity. This paragraph should be rewritten to avoid appearing self-contradictory.

**We have clarified the first sentence of this paragraph: “Capelin larval production in August in Trinity Bay did not change appreciably pre and post-1991, which was used as support for the non-collapse hypothesis.” (L371-78); and we clarified throughout the paragraph that we were comparing late-larval densities in August in years pre-1991 to September in years post-1991 due to persistently later spawning in capelin post-1991.**

12. Physical Variability section starting on line 499. The authors need to provide some physiological arguments for cold vs. warm temperatures and what and how those link to the population. As it reads, I’m completely confused as to what the purpose of this section is.

**Based on comments from all 3 reviewers, we have re-written portions of this paragraph (L537-550)**

13. Ahhh. After reading the first sentence of the Discussion, I understand the “Physical Variability” section. If I understand, that section is a major portion of their argument, and is maybe even the “coup de grace”. If this is the case, much more elaboration is needed for that section. I really was confused by that section when reading it.

**Based on comments from reviewers, we have deleted the reference to bottom-up processes in the first sentence of Discussion and clarified the Physical Variability section.**