Center for Independent Experts (CIE) Independent Peer Review Report

Atlantic Cod Research Track Peer Review

Virtual

July 31 – August 3, 2023

Dr. Noel Cadigan

Centre for Fisheries Ecosystems Research
Marine Institute of Memorial University of Newfoundland
St. John's, NL. Canada

September 2023

Executive Summary

The assessment documentation provided was extensive. The Woods Hole Assessment Model (WHAM) for the stock assessment of Western Gulf of Maine (WGOM), Georges Bank (GB), Southern New England (SNE) and Eastern Gulf of Maine (EGOM) cod stocks were competently applied by the assessment teams. The WHAM model is a very useful state-space stock assessment package for age-structured assessments. I was impressed with the flexibility of WHAM, including the ability to do short-term projections internally. This is a useful way to include uncertainty in the projections.

I mostly agreed (see below for differences) with the Panel conclusions for the WGOM, GB, SNE and EGOM cod stocks. The assessments for WGOM, EGOM, GB, and SNE cod are acceptable for use in subsequent management track processes, although there are some research recommendations that must be first addressed for EGOM, GB, and SNE cod before using the models in management track processes. However, I don't think that addressing these recommendations will require substantial new investigations. I also don't think advice will change substantially.

My conclusions for ToR 6 and 8 are a little different than the Panel conclusions. I conclude that ToR 6 was mostly met, whereas the Panel concluded that this ToR was met for all four stocks. I find that the WG did not consider the sensitivity of projections to assumptions, as required by ToR 6. This should be considered before a management track. For ToR 8, I conclude that this was partially met, whereas the Panel concluded that this ToR was met for all four stocks. However, specific backup alternatives were not reviewed by the Panel, so I suggest that if the proposed assessment approaches are rejected in a future management track assessment, then more investigations will be required to develop a credible backup approach.

The main deficiency of these assessments is limited length and age sampling for EGOM and SNE cod. However, recent surveys do not catch many cod in these regions so there are not many cod to sample. Nonetheless, it seems that stock-specific length and age sampling programs can be improved. There is a need to quantify uncertainty in fishery catches and discards, as well as to standardize and quantify the uncertainty and sampling distribution of age compositions. The lack of regular age sampling for SNE and WGOM cod, and for longline surveys in WGOM, means that a fully age-structured assessment model may not be best practice. A length-age structured model that can use length and age samples directly seems more appropriate, but this requires additional research.

The main research recommendations by the Panel dealt with addressing these deficiencies. I provided additional and longer-term recommendations.

Background

The Research Track Peer Review meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review stock assessments and models. The research track peer review is the cornerstone of the Northeast Region Coordinating Council stock assessment process. The results of this peer review are incorporated into future management track assessments, which serve as the basis for developing fishery management recommendations.

A research track assessment for Atlantic cod (*Gadus morhua*) was initiated in the fall of 2021, including the development of a Working Group (WG). Over the period of the research track, the WG met monthly over a series of virtual meetings. WG members also met through additional sub-groups meetings to discuss finer details of various research to support individual TORs, of which discussions and recommendations were brought before the WG for consensus. The WG also received input and contributions from non-WG members. This Review Panel membership is described in Appendix 3. The support of all these scientists and staff to the Review Panel process is gratefully acknowledged.

The WG developed a term of reference (ToR 9) to formalize the treatment of newly synthesized Atlantic cod stock structure information. Based on a review of the science on Atlantic cod stock structure and the available data to support stock assessment, the WG came to a consensus on the most appropriate spatial assessment units. The WG recommended that the cod research track process proceed with four spatial units for assessment: 1) eastern Gulf of Maine, 2) western Gulf of Maine (winter and spring spawners combined), 3) Georges Bank, and 4) Southern New England (including the Mid Atlantic Bight). The rationale for a 4-unit structure, relative to the historical 2-unit model, was better alignment between the scale of cod stock assessment units and biological stock structure that can be supported with available information. This change in cod stock structure meant revising fishery catch data, survey indices, and estimates of biological parameters, all of which required Panel review.

The purpose of the Atlantic Cod Research Track Peer Review was to provide an external peer review of the US Atlantic cod stock assessments. The National Marine Fisheries Service (NMFS) required three reviewers to participate in the panel review. The chair, who was in addition to the three reviewers, was appointed by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee. Reviewers were required to have working knowledge and recent experience in the use and application of index-based, age-based, and state-space stock assessment models, including familiarity with retrospective patterns, model diagnostics from various population models, and how catch advice is provided from stock assessment models. In addition, knowledge and experience with simulation analyses was an asset. Each reviewer was required to write an individual review report in accordance with the Performance Work Statement Office of Management and Budget Guidelines, and the Terms of Reference (TORs) below.

Role of reviewer

All assessment documents and most supporting materials were made available to the Panel via a web page (https://apps-nefsc.fisheries.noaa.gov/saw/sasi.php) more than two weeks before the meeting, on July 12, 2023. These documents are listed in Appendix 1. I reviewed the background documents I was provided and compiled a list of issues for clarification during the Review Panel meeting. I attended the entire Review Panel review meeting remotely via WebEx. I reviewed presentations and reports and participated in the discussion of these documents, in accordance with the SoW and ToRs (see Appendix 2). I drafted text for the Panel report, especially for Georges Bank cod. After the meeting I participated in email discussions to finalize the review panel summary report. This CIE report is structured according to my interpretation of the required format and content described in Annex 3 of Appendix 2. This includes providing a "Summary of Findings for each ToR in which the weaknesses and strengths are described, and Conclusions and Recommendations in accordance with the ToRs".

Summary of findings

I first provide summaries that apply to all four assessments, and then present stock-specific summaries where necessary.

Acronyms: GOM – Gulf of Maine; GB – Georges Bank; WGOM – Western GOM; EGOM – Eastern GOM; M – natural mortality rate.

The WG report provided chapters for each ToR. I really liked this. From a review perspective, it simplified navigating documents and made it easier for me to understand information rather than spending a lot of time searching for information.

I have provided references for some of my comments, and I include the references at the end of each ToR they apply to.

ToR 1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.

Three WPs were also provided for this ToR: WP1 Stakeholder Meeting Summary; WP2 Development of Ecosystem Indicators; WP3 Environmental Influences on Cod. An ecosystem profile was presented based on a literature review of papers published after the year 2000 and whose primary research area was the NE US.

I agree with the Panel's conclusion that this TOR has been met for the WGOM, GB, SNE and EGOM cod stocks. I am not an expert on cod stock productivity dynamics in this area, although I have been involved in Canadian cod stock assessments for many years – more in the Newfoundland and Labrador regions, and to a much lesser extent in the Gulf of St. Lawence and Scotian Shelf regions. The latter areas are probably more relevant to GOM and GB. I recognize it is practically impossible to review all of the literature that is possibly related to cod stock productivity dynamics in this area; however, I was aware of some papers that seemed relevant regarding the impacts of seal predation, which I summarize below. The papers suggest that grey seals may be contributing to increased M for stocks they are associated with. Predation may be more of a problem in the winter and/or during spawning periods. A stakeholder also reported that grey seals prey upon cod at a higher rate in the fall than in the spring. However, to the north, it seems the impacts of increasing harp seals on cod is not clear, and bottom-up processes may be more important.

Hammill, M.O., Stenson, G.B., Swain, D.P., and Benoît, H.P. 2014. Feeding by grey seals on endangered stocks of Atlantic cod and white hake. – ICES Journal of Marine Science, 71: 1332–1341.

In the Cabot Strait, when overwintering aggregations of cod were present, cod accounted for 68% (range 57–80%) of the male diet from stomachs, and 46% (range: 31–64%) of the diet determined from intestines. Among females, cod represented 14% (range: 0–34%) and 9% (range: 3–54%) of the diet from stomachs and intestines, respectively.

Swain, D.P. and Benoît, H.P., 2015. Extreme increases in natural mortality prevent recovery of collapsed fish populations in a Northwest Atlantic ecosystem. Marine Ecology Progress Series, 519, pp.165-182.

They concluded that the lack of recovery of several groundfish species in the southern Gulf of St. Lawrence is due to dramatic increases in M of larger individuals. These high levels of M reflect a 'predator pit' or predation-driven Allee effect, resulting from the severely depleted abundance and the high and rising abundance of grey seals. Recovery does not appear to be possible under current conditions in this ecosystem, even in the absence of fishing.

Swain, D.P., Benoît, H.P. and Hammill, M.O., 2015. Spatial distribution of fishes in a Northwest Atlantic ecosystem in relation to risk of predation by a marine mammal. Journal of Animal Ecology, 84(5), pp.1286-1298.

Distributions of cod, hake and skate were strongly related to risk of predation by seals, with distribution shifting into lower risk areas as predation risk increased. Non-prey species did not show similar changes in habitat use. Spatial variation in fish condition suggests that these low-risk areas are also less profitable for cod and skate in terms of food availability.

Rossi, S.P., Cox, S.P., Swain, D.P. and Benoît, H.P., 2017. A mixed-effects approach to estimating time-varying natural mortality in a statistical catch-at-age model: Application to stock assessments for eastern Georges Bank Atlantic Cod and Georges Bank Yellowtail Flounder. TRAC Working Paper 2017-08.

The degree to which grey seals influence M in GB fish is unknown, though satellite tagging has shown high proportions of grey seals foraging on Georges Bank and neighboring waters in winter and early spring (Breed et al. 2006).

den Heyer, C.E., Bowen, W.D., Dale, J., Gosselin, J.F., Hammill, M.O., Johnston, D.W., Lang, S.L., Murray, K.T., Stenson, G.B. and Wood, S.A., 2021. Contrasting trends in gray seal (*Halichoerus grypus*) pup production throughout the increasing northwest Atlantic metapopulation. Marine Mammal Science, 37(2), pp.611-630.

Since 2004, the rate of increase in pup production at Sable Island has slowed to about 5%–7% per year, while the newer colonies in southwest Nova Scotia and the northeastern United States are increasing rapidly. In 2016, the Muskeget Island (MA) breeding colony produced 3,900 (SE = 200) pups, making it the third largest breeding colony in the northwest Atlantic. This southward shift in production may reflect climate-mediated changes in population growth as well as re-establishment of colonies throughout the former range associated with increased protection. Where there are pups, there are adults, at least during the winter season.

Buren, A.D., Koen-Alonso, M. and Stenson, G.B., 2014. The role of harp seals, fisheries and food availability in driving the dynamics of northern cod. Marine Ecology Progress Series, 511, pp.265-284.

Biomass dynamics were best explained by a combination of fisheries removals and capelin availability, whereas seal consumption was found not to be an important driver of the northern cod stock.

Regular, P.M., Buren, A.D., Dwyer, K.S., Cadigan, N.G., Gregory, R.S., Koen-Alonso, M., Rideout, R.M., Robertson, G.J., Robertson, M.D., Stenson, G.B. and Wheeland, L.J., 2022. Indexing starvation mortality to assess its role in the population regulation of Northern cod. Fisheries Research, 247, p.106180.

Their results indicate that starvation-induced mortality represents an important component of M experienced by Northern cod, supporting the idea that insufficient food contributed to the collapse of the stock in the early 1990s and, subsequently, limited its recovery.

ToR 2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.

Three WPs were also provided for this ToR: WP5 Rec Discard Mortality, WP6 FDD Exploration, WP7 Cod LPUE. I agree with the Panels conclusion that this ToR has been met for all four cod stocks reviewed (WGOM, GB, SNE and EGOM). Extensive documentation was provided on commercial and recreational landings and discards. However, this is a ToR that a review panel has limited ability to assess, not having specific knowledge and experience with the fisheries in the NE US region. We simply must trust that all important sources of cod fishing mortality have been considered, but this seems like a safe assumption considering the wide and varied participants in the research track process.

I was not sure what "*Characterize the uncertainty*" meant for this ToR. The WG discussed uncertainty, which may be sufficient to characterize it. However, I recommend that uncertainties in landings and discards need to be quantified in a way that can be included in age/length based stock assessment models. In a few Canadian assessments this is done using bounds on catches, which involve a plausible range of landings and discards (i.e., bounds) that stakeholders and data providers agree with (e.g., DFO, 2017; DFO, 2020), such that it is considered implausible that catches could be outside the bounds. In time periods when the landings and discards "estimates" are very complete (i.e., a census) then lower and upper bounds could be the same as the estimates. The methods used to fit assessment data with landings bounds have been published (see e.g., Hammond and Trenkel, 2005; Bousquet et al., 2010; Cadigan, 2016; Van Beveren et al., 2017) and could be implemented easily with WHAM. Note that this is very different than fitting assessment models to different catch "streams", which is a more incomplete way to account for uncertainty. The choice of bounds will be subjective, but at least a range of choices could be investigated, which I suggest is better than not accounting for uncertainty at all.

The Panel noted that ToR 2 did not ask about the reliability/uncertainty in the size/age composition of the catch, which is an important contributor to uncertainty in assessment models, including state-space stock assessment models. I agree with the Panel who recommended research continue to standardize and quantify the uncertainty and sampling distribution of age compositions. The Panel provided some references for this. Thorson et al. (2023), in the context of state-space assessment models, recommended using design-, model-, or bootstrap estimators to identify the variance of all data inputs. I recall the NEFSC had a bootstrap procedure to estimate uncertainty of catch age-compositions, but this was not developed during the research track. Maybe there were good reasons for this, but I endorse the above recommendation from Thorson et al. (2023). I realize this is complex and I consider this to be a long-term recommendation.

There was insufficient age-sampling for some catch data, and none for recreational data which was the main source of catch in the SNE cod stock. The various procedures the WG used to produce catch age compositions from length compositions seemed reasonable, as were the decisions when to not estimate catch age composition. However, clearly a more direct (some would say honest) approach is to use an assessment model that can fit to both age and length composition data. This will remove the need to convert length compositions to age composition even for data sources that had no age samples collected. Again, as a longer-term research recommendation, I fully endorse the Panel's suggestion that age-length-based models that can use the length data directly as sampled should be investigated. This will also remove the need to fit multinomial regression models to predict age from length, for lengths with no ages. This will also remove the need to use stochastic von Bertlanffy growth models to predict age from length, which are known to be biased for large sizes.

The assessment report provided extensive documentation of the methods used to estimate the age composition of catches and discards. I appreciate the complexity of this, especially since sampling programs in this area were not designed for the four stocks now being assessed, and getting good sampling from a species that is infrequently caught in some areas is difficult anyway. The choices involved in estimating catch-at-age that the assessment teams made seemed reasonable, and we did not have time to evaluate the impacts of different choices and if there were better options. There was substantial text for ToR 2 about the sparseness of samples for some year-gears-regions. I assume this was to address the ToR in terms of "*Characterize the uncertainty in these sources of data*". This was good, but at the same time I became concerned about the quality of the age compositions and how they could be used in fitting assessment models. The bubble plots provided in the working paper (e.g., Figure 2.12) did not give me a sense of how well cohorts were tracked, and at what ages was the composition sampling better or worse. For this purpose, I like to use "SPAY" plots (e.g., https://rpubs.com/rajeevkumar/SPAY). These are just plots of standardized deviations in composition proportions over time and they can be useful to detect relatively strong and weak year classes. By comparing multiple data sources, we can get a high-level understanding of the consistency of the information across the data sources. There are other ways of plotting the composition information that also give useful information about cohort-consistency and changes in total mortality rates over time, and ICES has some good procedures for this.

I obtained the commercial catch-at-age estimates from a github site provided by the assessment team. The SPAY plots for these data (Figure 1, end of ToR) demonstrate that both strong and weak cohorts tracked fairly well at ages 2-9 for GB and WGOM cod. The sampling for these stocks was sufficient to track large or small years classes when they occurred. However, the catch composition cohort information for EGOM and especially SNE cod are sparser and noisier. Of course, if there is little cohort variation then these plots will be noisier because there is nothing to track. Nonetheless, the relatively weak 1975 EGOM cohort tracked that way through 5 years of sampling, and relatively stronger 1987 and 1988 cohorts tracked reasonably well. Note that the 1975 cohort was relatively strong in GB, and above-average in WGOM, which indicates different cohort dynamics in these areas. Figure 1 also demonstrates other differences in relative cohort strength between GB and WGOM, such as the 2011 cohort which was relatively strong in WGOM but not in GB, and the 1985 cohort which was relatively strong in GB but not in WGOM. There is little catch information about the sizes of recent cohorts in EGOM and SNE.

I also obtained estimated age compositions of the recreational catches (Figure 2). These do not track cohorts as well overall, especially in EGOM. However, some weak and strong years classes have tracked through the recreational age compositions for WGOM cod, and to a lesser extent for SNE.

Overall, I conclude from the SPAY plots that the GB and WGOM fishery age compositions are sufficient to support a cohort-based stock assessment model. These SPAY plots can also indicate possible better ages for plus groups, which I consider under ToR 4 below. There is less information for EGOM and SNE cod, but enough to attempt fitting cohort models. Surveys are also important for this purpose.

Commercial and recreational landings per unit effort (LPUE) were explored as potential indices of abundance to be used in cod assessments (WP 7). The LPUEs were ultimately not used for the EGON, WGOM, or GB assessments, so the panel spent little time on these analyses, because of the Panel's time constraints. I did not fully understand the standardization methods. However, the recreational LPUE was used in the assessment for SNE cod. The documentation was not clear about the selection of trips to compute LPUE, which we learned was only trips that caught at least one cod. I suggest that the Panel's research recommendation #2 for the management track should be a condition before this model could be used in a management track.

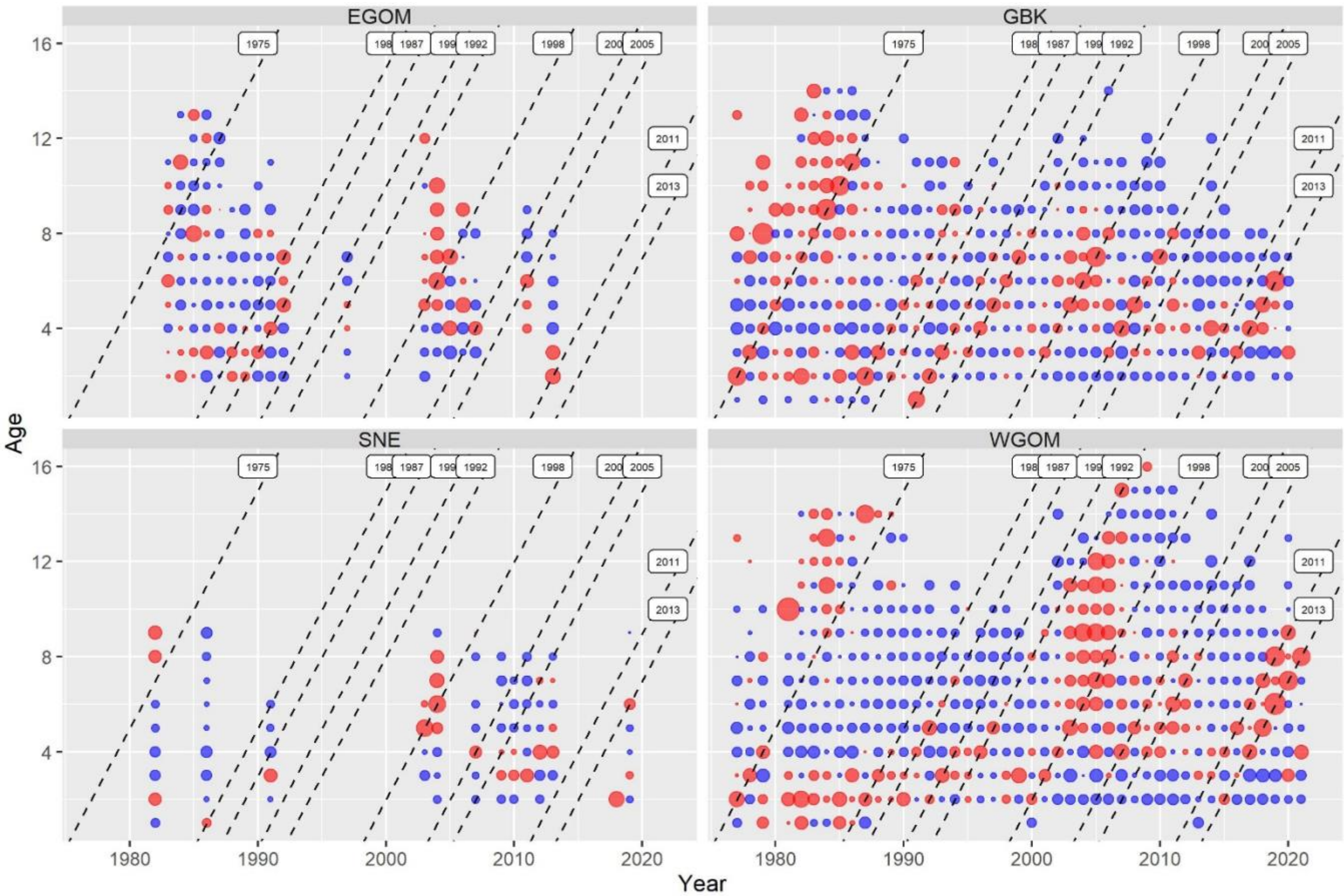

Figure 1: Standardized proportion-at-age (SPAY) for the estimates of fishery catches for each stock (panels). The area of a bubble is proportional to the absolute value of the standardized proportion. Red is positive and blue is negative. Dashed lines indicate illustrative cohorts, which are shown in the text boxes.

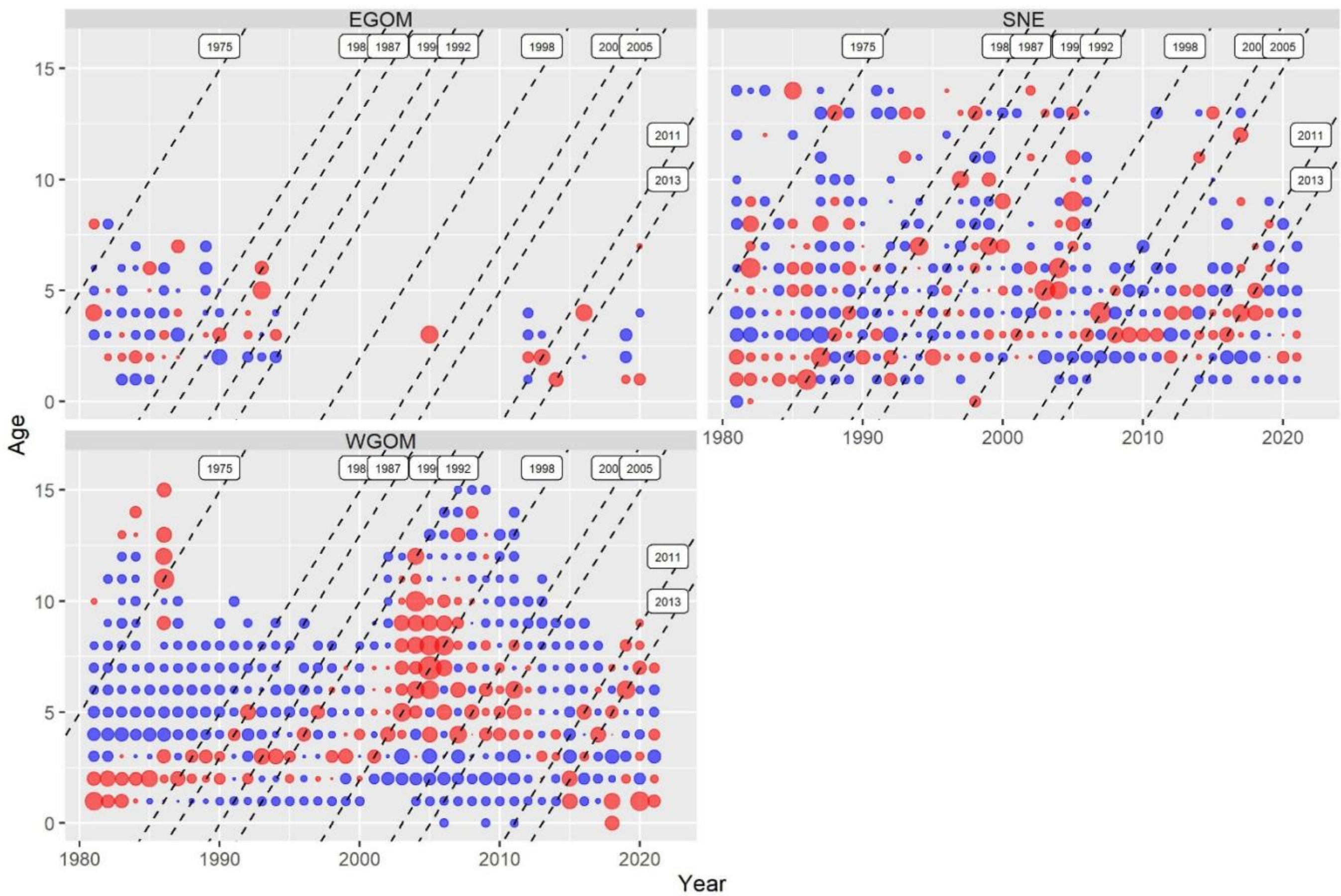

Figure 2: Standardized proportion-at-age (SPAY) for the estimates of recreational catches for each stock (panels). See Figure 1 for additional description.

Additional References

Bousquet, N., Cadigan, N., Duchesne, T. and Rivest, L.P., 2010. Detecting and correcting underreported catches in fish stock assessment: trial of a new method. Canadian Journal of Fisheries and Aquatic Sciences, 67(8), pp.1247-1261.

Cadigan, N. G. 2016. A state-space stock assessment model for northern cod, including under-reported catches and variable natural mortality rates. Canadian Journal of Fisheries and Aquatic Sciences, 73(2): 296-308. http://dx.doi.org/10.1139/cjfas-2015-0047.

DFO. 2020. Stock Assessment of NAFO Subdivision 3Ps Cod. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/2018. https://waves-vagues.dfo-mpo.gc.ca/Library/40877413.pdf.

DFO. 2017. Assessment of the Atlantic Mackerel Stock for the Northwest Atlantic (Subareas 3 and 4) in 2016. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2017/034. https://waves-vagues.dfo-mpo.gc.ca/Library/40619576.pdf.

Hammond, T.R. and Trenkel, V.M., 2005. Censored catch data in fisheries stock assessment. ICES Journal of Marine Science, 62(6), pp.1118-1130.

Thorson, J.T., Monnahan, C.C. and Hulson, P.J.F., 2023. Data weighting: An iterative process linking surveys, data synthesis, and population models to evaluate mis-specification. Fisheries Research, 266, p.106762.

Van Beveren, E., Duplisea, D., Castonguay, M., Doniol-Valcroze, T., Plourde, S. and Cadigan, N., 2017. How catch underreporting can bias stock assessment of and advice for northwest Atlantic mackerel and a possible resolution using censored catch. Fisheries Research, 194, pp.146-154.

ToR 3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.

In addition to the extensive information provided in the WG report, four WP were provided: WP8 NEFSC Trawl Survey Expanded Figs., WP9 Survey Time Series Correlations, WP10 Integrated Survey Indices (VAST), WP11 EGOM Sentinel Index Modification. I agree with the Panels conclusion that this ToR has been met in full for all four cod stocks reviewed (WGOM, GB, SNE and EGOM).

I agree with the Panel recommendations, particularly the one for integrated survey analyses. When there are spatial differences in the sampling domains of surveys, and changes in the spatial distribution of a stock over time, then this can create spurious trends in survey indices that space-aggregated stock assessment models have no ability to sort out. A spatial integrated analysis of the survey information is a good solution. A computational challenge will be integrating survey catch-at-length with an appropriate count data model. Fine-scale space-time-length models are often computationally prohibitive. However, this could include the unconverted Bigelow and Albatross survey catches and provide an opportunity to integrate comparative fishing data in the model, maybe as coming from a “dummy” survey and year similar to Benoît and Cadigan (2014). This could also provide an opportunity to examine if there were important spatial differences in the relative efficiency of the Bigelow survey compared to the Albatross.

Similar to the fishery age compositions, the Panel concluded that the WG provided insufficient information about how well cohorts tracked through the various survey age compositions. This is a basic first step when implementing age-based cohort models. I computed SPAY plots for the 4 stocks, which are provided in Figures 3-6 at the end of this ToR. Between-age pairwise correlation plots of survey indices of cohort size are another common way to examine how well cohorts track through survey age compositions (see Figures 7-17 in Appendix 4 at the end of this report).

For the WGOM region, I find overall that the NEFSC Fall survey tracks cohorts the best, and the Spring survey is second-best. The NEFSC survey indices at age zero have reasonably good correlations with age one indices a year later (Figures 9 and 10) but the SPAY plot (Figure 3) suggests less consistent tracking of cohorts at age zero. Hence, it seems reasonable that age zero was not used as a recruitment index in the WGOM assessment model, but it also seems reasonable to re-visit this in the future. The NEFSC Spring indices seem to track cohorts better at ages 6-8, while the fall NEFSC survey indices do better at younger ages.

The survey indices in the GB region do not track cohorts as well. The NEFSC Spring indices (Fig. 11) and the Canadian Spring survey indices (Fig. 13) at ages 1-7 have fairly good correlations. The NEFSC Fall survey (Fig. 12) does not seem to track cohorts as well, although relatively weak and strong years classes are evident in the age compositions (Fig. 4).

The survey cohort information for cod in the EGOM region seems less reliable than in the WGOM or GB regions, although the lag 1 cohort correlations are all usually significantly positive (Figs. 14-15) at ages 2-5 for the NEFSC surveys. The ME_NH inshore trawl surveys have not yet provided evidence of cohort tracking. The survey age composition information in the SNE region is limited and has not tracked cohorts well. This could also occur if there have not been relatively weak or strong cohorts in this region in recent years.

The fundamental stock assessment information provided by age-based survey indices is the total mortality rate, Z. Estimates of F largely depend on assumptions about M, but the Z total should be more reliable, especially for surveys that have flat-topped selectivity. A good way to estimate survey Zs is by using SURBA, and in a research track process I recommend that this be done as an initial step to assess how similar are the recruitment and Z signals from various surveys in a region. The full assessment model will integrate and average this information, and also include catch information which I often find only scales the assessment and otherwise provides little information about population dynamics unless strong selectivity assumptions are made. To get a rough sense of the survey Z signals, I compute the cohort average Z's for ages that seemed fully selected (Figs. 18-32).

In the WGOM region, overall I do not see evidence from the NEFSC surveys (Figs. 18-19) that average Z had changed much for cohorts since 1980. This is similar to the GB region (Figs. 25-26). However, the Canadian Spring index in this region does indicate higher Zs for cohorts since about 2005 (Fig. 27). Survey Zs in the EGOM and SNE are noisy and do not indicate trends in Z (Figs. 28-32).

Using the same age data to produce age-based indices from differences surveys is using data twice, and can create complicated correlations. This can be avoided in an age-length structured model.

I recommend that analyses of survey weight-at-age be provided in future research tracts. There are some simple models that can be applied to reduce sampling errors and fill-in sampling gaps (e.g., Cadigan, 2020; Cadigan and Rideout, 2022). In fact, similar models of biological processes are now provided as an option with the ICES SAM stock assessment software. There can be substantial differences in survey and fishery weights-at-age (i.e., Cadigan, 2020) and it is important for assessments to understand this. The change from deriving stock weights (used to calculate SSB) from commercial sampling to weights derived from surveys was a major contributor to assessment retrospective differences in biomass and SSB for 3Ps cod (DFO, 2020). For 3Ps cod, weights from surveys are thought to better represent population weights-at-age.

Additional References

Benoît, H.P., and Cadigan, N. 2014. Model-based estimation of commercial-sized snow crab (*Chionoecetes opilio*) abundance in the southern Gulf of St. Lawrence, 1980-2013, using data from two bottom trawl surveys. DFO Can. Sci. Advis. Sec. Res. Doc. 2014/082. v + 24 p.

Cadigan, N. 2020. A Simple Mixed-Effects Model to Smooth and Extrapolate Weights-at-Age for 3Ps Cod. DFO Can. Sci. Advis. Sec. Res. Doc. 2023/024. iv + 49 p. https://www.dfo-mpo.gc.ca/csas-sccs/Publications/ResDocs-DocRech/2023/2023_024-eng.pdf

Cadigan, N and R. Rideout. 2022. A Mixed-Effects Model to Smooth and Interpolate Survey Weights-at-Age for 3NO Cod. NAFO SCR Doc. 22/040. 16 p. https://www.nafo.int/Portals/0/PDFs/sc/2022/scr22-040.pdf

DFO. 2020. Stock Assessment of NAFO Subdivision 3Ps Cod. DFO Can. Sci. Advis. Sec. Sci. Advis. Rep. 2020/2018. https://waves-vagues.dfo-mpo.gc.ca/library-bibliotheque/40877413.pdf

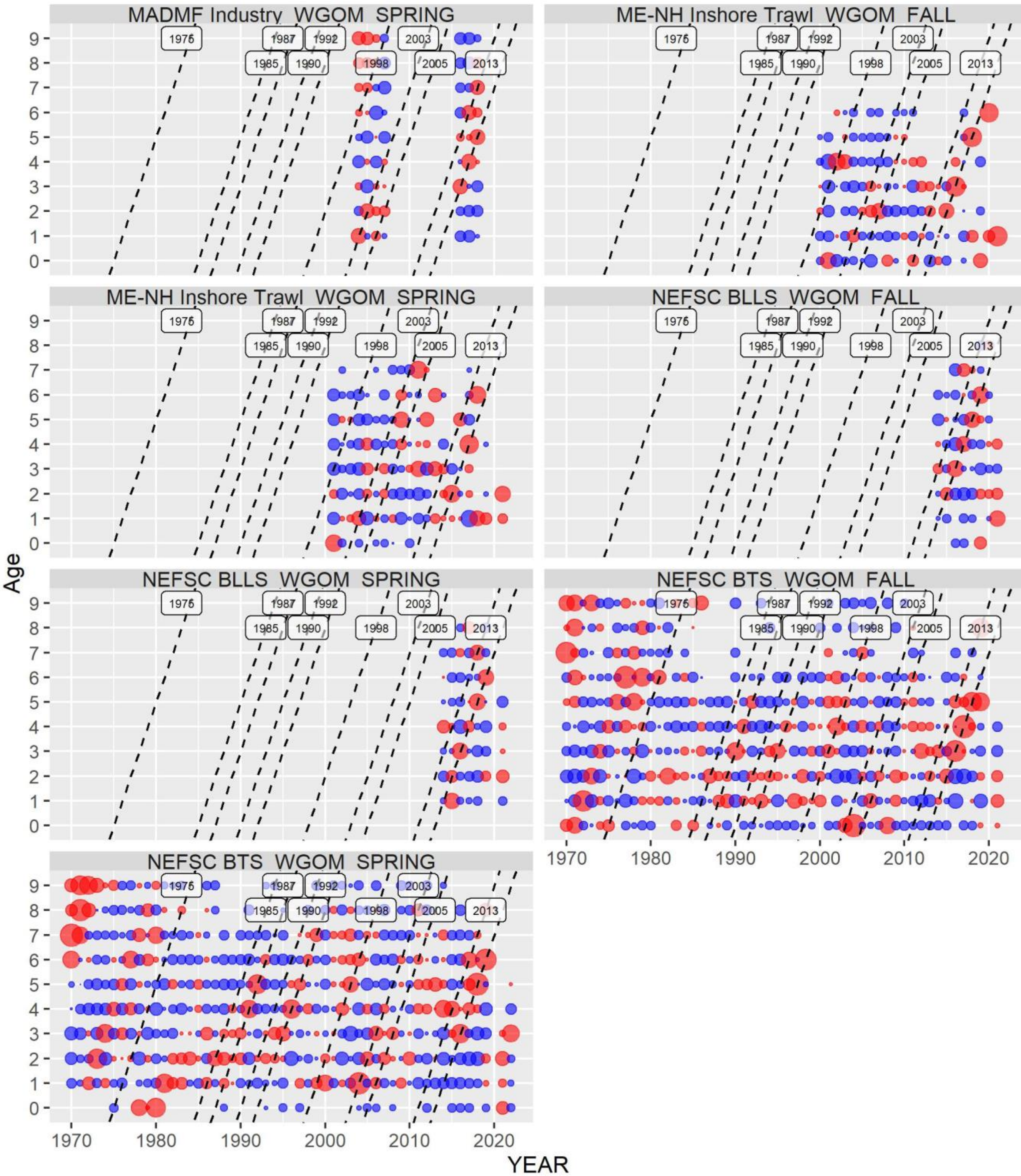

Figure 3: Spay plots for WGOM cod survey (panels) age compositions. See Figure 1 for additional description.

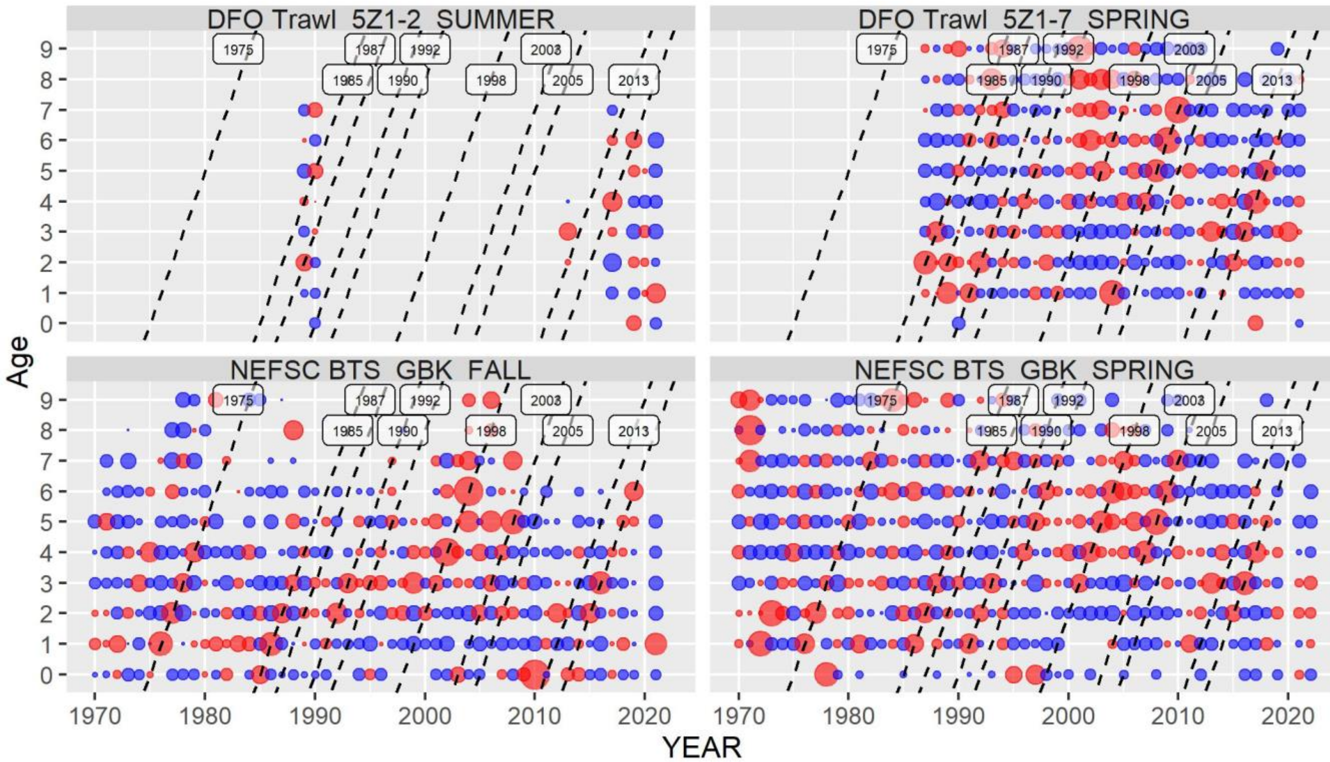

Figure 4: Spay plots for GB cod survey (panels) age compositions. See Figure 1 for additional description.

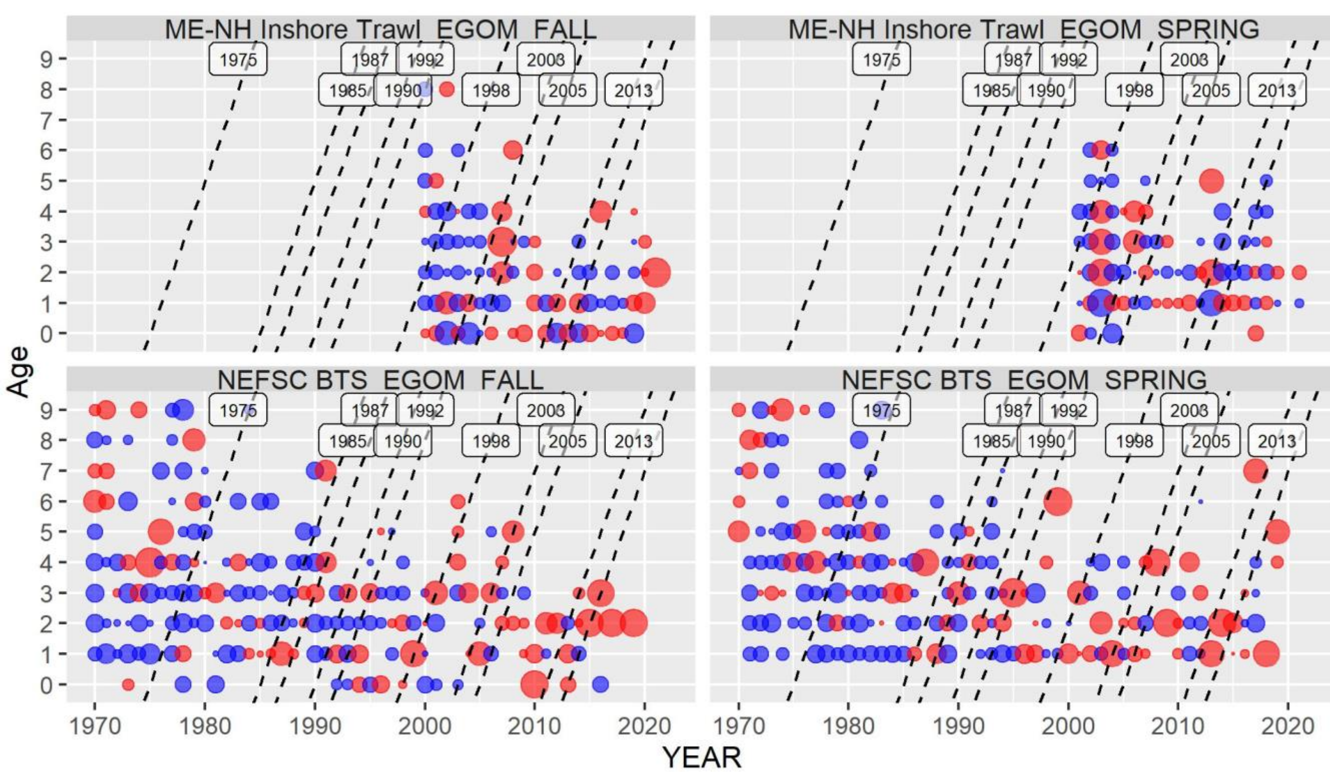

Figure 5: Spay plots for EGOM cod survey (panels) age compositions. See Figure 1 for additional description.

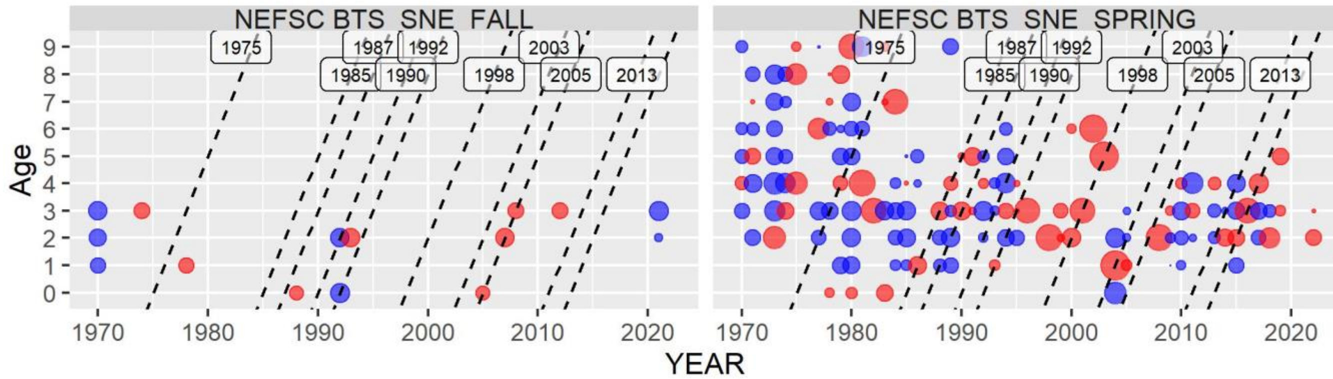

Figure 6: Spay plots for SNE cod survey (panels) age compositions. See Figure 1 for additional description.

ToR 4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.

The Panel concluded that ToR 4 had been mostly met for each of the cod stocks, and I agree with this. The resulting assessments are accepted for use in subsequent management track processes, although there are some research recommendations that must be addressed before using the models in management track processes (see ToR 7). This was for EGOM, GB, and SNE cod. I don't think that addressing these recommendations will require substantial new investigations. I also don't think advice will change substantially, given the status of these stocks; however, assessment teams should strive to provide the best assessments possible, regardless of possible stock status conclusions. For the rest of this ToR, I give my additional perspective and recommendations on some assessment issues beyond the panel summary.

I conclude that the WHAM model is a very useful state-space stock assessment package for age-structured assessment. I was impressed with the flexibility of WHAM. However, there is a need to extend the approach to include length sample data. At present, WHAM only accommodates age data and this resulted in cases where length samples were not used, or the same age information was used to convert length-based survey indices into age-based ones. These are not "best practice" approaches. I am aware that there is a "growth" branch of WHAM under development and I recommend that this continue so that, in future research tracks, the survey and catch length sampling can be better utilized. This is longer-term research because WHAM may need to be extended to be fully age and length structured and include time-varying growth rates. This will be computationally more complex than the age-based WHAM, and more complex models may not produce better assessments. Much research will be required to understand what more useful model configurations in this extended setting are.

An example of this involves how to use the comparative fishing information collected for the Albatross to Bigelow vessel change. I prefer assessment models that can use survey length compositions and conditional age compositions, and models with length-based survey catchability. With these models the original indices can be used for model fitting, and the comparative fishing data can be included as a likelihood component to constrain the ratio of catchabilities (i.e., the length-based relative efficiency) for the Albatross and Bigelow survey series. In this way, the uncertainty in the survey calibrations is accounted for. A complication is that the calibration curves in Miller (2013) may only be appropriate for the stock regions the comparative fishing data were obtained from, and these calibration curves may not be appropriate for each of the four cod stocks. This issue requires further investigation.

I am very skeptical about life-history methods used to infer values of M. Nonetheless, I find that the base values used in the assessments seemed reasonable. Estimating the overall level of M within age-based stock assessments similar to those for GB and WGOM cod is a challenge (e.g., Aldrin et al., 2021), and more so for SNE and EGOM cod with the sparser data for these stocks. Estimating change in M is more feasible (e.g., Aldrin et al., 2021) but this did not produce converged models for GB and WGOM cod. However, in most Canadian cod stock assessments, time-varying M is included in assessment models for management advice. I am unsure why this did not work in the GB and WGOM cod assessments, but this was not a focus of our review. I recommend that there should be a more North Atlantic system-wide approach to understanding how to model M in age-based and state-space stock assessments with available data similar to GB and WGOM cod, NAFO Divisions 2J3KL, 3NO, 3Ps, 3Pn4RS, and 4T cod in Canada, and perhaps some ICES cod stocks, and possibly other groundfish stocks in these regions. Much of the literature (e.g., Maunder et al., 2023) about estimating M in stock assessments is not in the state-space context or directly relevant to the types of assessments for the above Atlantic cod stocks. I think extending investigations like Miller and Hyun (2018) to include the above stocks is useful, and finding criteria that can indicate situations where M or time-change in M can be identified (e.g., Clark, 2022) is useful. This is a research project that may be more suitable for a group like the ICES Methods Working Group (although their members probably have full research programs already), and I am not proposing this for the Atlantic cod research track process.

Tables of parameter estimates for final models should be provided. These could be split into catchability parameter tables, and tables for correlation and variance parameters, etc. Comparing variance parameters between runs can give an overall description of how well a model fits various data sources, and how much process error was required to achieve these fits.

Some of the simulation self-test results were quite poor (e.g., SNE cod). Zheng and Cadigan (2023) showed that some bias may be expected, depending on the details of the simulation procedures and assessment data. The bias is like the smoothing bias that occurs with GAMs and other smoothers. This issue will not exist in assessment models without random effects. The efficacy of simulation self-tests requires additional research in the state-space stock assessment context, but this is a broader recommendation than appropriate for the Atlantic cod research track process.

There was some evidence presented in the main WP suggesting that the weight-at-length relationship for cod has changed over time. This has been the case for cod stocks off the coast of Newfoundland and Labrador. Changes in weight-at-length, and in particular the proportion of fish in very poor condition, have been linked to time-varying M (e.g., Björnsson et al., 2022; Regular et al., 2022; Varkey et al., 2022; Weerasekeraa et al., 2023). I recommend more detailed investigations of the weight-length relationships for the US Atlantic cod stocks should continue. However, there may be within-stock spatial variation in the weight-length relationship that can confound temporal variation if there have been changes in the sampling locations for weights over time (e.g., Cadigan et al., 2022).

I find it very helpful to compare observed and model predicted survey indices, plotted by cohort and age like the top panel in Fig. 18. Systematic differences in the observed and model predicted catches may indicate change in M or mis-reported catch. I recommend that these should be produced for future research track reviews.

Including age zero survey indices can be problematic, but this depends on the size and distribution of the age zero cod. In a major nursery area for Northern cod (e.g., Gregory et al., 2016), weak catch rates for cohorts at age 0 were usually followed by weak catch rates for the same cohort at age 1. However, strong age zero catch rates did not necessarily produce strong age 1 catch rates. This seemed to indicate high time-variation in M at age zero. Hence, in a stock assessment context, a low age zero index may indicate a weak cohort recruiting to future fisheries more reliably than a high age zero index would indicate a stronger cohort. Hence, including an age zero index in an assessment model may require time-varying M at age 0 to fit the index well. However, an age zero index may have value if it reliably indicates weak cohorts, even if it does not reliably indicate strong cohorts.

My experience is that the size (i.e., SD) of cohort process errors is better determined for stocks with multiple age-based surveys. The cohort errors represent deviations in population dynamics that should be reflected by all or most survey indices. If there are multiple surveys then cohort process errors, when not included in the model fitting, should create common residual patterns in the survey indices. However, when there is only one age-based survey index then it is more difficult to distinguish between cohort process errors and survey measurement errors. Hence, it was not surprising to me that SNE model runs with process errors in selectivity, natural mortality, catchability of recreational LPUE (iid), and NAA (i.e., full state space model) did not converge.

Additional References

Aldrin, M., Aanes, F.L., Tvete, I.F., Aanes, S. and Subbey, S., 2021. Caveats with estimating natural mortality rates in stock assessment models using age aggregated catch data and abundance indices. Fisheries Research, 243, p.106071.

Björnsson, B., Sólmundsson, J. and Woods, P.J., 2022. Natural mortality in exploited fish stocks: annual variation estimated with data from trawl surveys. ICES Journal of Marine Science, 79(5), pp.1569-1582.

Cadigan, N., Robertson, M.D., Nirmalkanna, K. and Zheng, N., 2022. The complex relationship between weight and length of Atlantic cod off the south coast of Newfoundland. Canadian Journal of Fisheries and Aquatic Sciences, 79(11), 1798-1819. https://doi.org/10.1139/cjfas-2021-0325

Clark, W.G., 2022. Why natural mortality is estimable, in theory if not in practice, in a data-rich stock assessment. Fisheries Research, 248, p.106203.

Gregory, R.S., Morris, C, Newton, B., and Sargent, P. 2016. Relative strength of the 2010, 2011 and 2012 year-classes, from nearshore surveys of demersal age 0 and 1 Atlantic cod in Newman Sound, Bonavista Bay. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/081. iv + 10 p

Maunder, M.N., Hamel, O.S., Lee, H.H., Piner, K.R., Cope, J.M., Punt, A.E., Ianelli, J.N., Castillo-Jordán, C., Kapur, M.S. and Methot, R.D., 2023. A review of estimation methods for natural mortality and their performance in the context of fishery stock assessment. Fisheries Research, 257, p.106489.

Miller, T.J. and Hyun, S.Y., 2018. Evaluating evidence for alternative natural mortality and process error assumptions using a state-space, age-structured assessment model. Canadian Journal of Fisheries and Aquatic Sciences, 75(5), pp.691-703.

Regular, P.M., Buren, A.D., Dwyer, K.S., Cadigan, N.G., Gregory, R.S., Koen-Alonso, M., Rideout, R.M., Robertson, G.J., Robertson, M.D., Stenson, G.B. and Wheeland, L.J., 2022. Indexing starvation mortality to assess its role in the population regulation of Northern cod. Fisheries Research, 247, p.106180. https://doi.org/10.1016/j.fishres.2021.106180.

Varkey, D., Babyn, J., Regular, P., Ings, D., Kumar, R., Rogers, B., Champagnat, J., Morgan, M., 2022. A state-space model for stock assessment of cod *Gadus morhua* stock in NAFO Subdivision 3Ps. DFO Canada Science Advisory Secretariat Research Document 2022/022. iv + 78 p.

Weerasekeraa, S.J.W.W.M.M.P., Cadigan, N.G., Nirmalkanna, K., Regular P.M., and R.M. Rideout. A spatiotemporal model to derive length-based starvation mortality for Atlantic cod (*Gadus morhua*) on the Southern Grand Bank of Newfoundland. Submitted.

Zheng, N. and Cadigan, N., 2023. Frequentist Conditional Variance for Nonlinear Mixed-Effects Models. Journal of Statistical Theory and Practice, 17(1), p.3.

ToR 5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.

I agree with the Panel's conclusion that this ToR has been met in full for all four cod stocks reviewed (WGOM, GB, SNE and EGOM).

I recommend that detecting change in the stock-recruitment relationship should be based on stock-recruitment residuals. We expect the recruitment rate (R/SSB) to decline as SSB increases, which is why it is not sufficient to examine only time-trends in recruitment rates, and especially time-trends in recruitment. I favor the approach of Tang et al. (2021) for detecting change in the stock-recruitment relationship.

The long-term stochastic projections used to derive the SSB_{MSY} proxy seemed simplistic and did not account for future variability in SSB weights, catch weights, selectivity, and maturity. I am also uncomfortable that these projections assumed that future recruitment is independent of SSB, even though all four cod stock assessments indicated strong relationships with the amount of recruitment and parental SSB. I agree that the common choice of F40% Fmsy proxy also needs rethinking, considering prevailing stock productivity. However, I appreciate that the assessment authors choose procedures to define F and B_{MSY} reference points that are typical of their region and in the US. I agree with the WG that "a broader discussion (i.e., across species) of appropriate methods for defining time windows to characterize prevailing conditions for stocks is needed in the region." I suggest that this broader discussion should be extended to include how to conduct long-term stochastic projections to define SSB_{MSY} and even F_{MSY}.

There were some large differences in catch and SSB weights-at-age assumed for reference point calculations across the four stocks. The age 9+ SSB weight for GB was 9.2 kg, whereas it was 14.21 kg in the EGOM. These are large differences and may reflect some measurement error. I recommend that improved procedures be investigated to derive weights-at-age.

Additional References

Tang, X., Zheng, N., Rideout, R.M., Wang, S. and Zhang, F., 2021. Identification of recruitment regime shifts with a hidden Markov stock-recruitment model. ICES Journal of Marine Science, 78(7), pp.2591-2602.

ToR 6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.

This ToR has been fully met for all four cod stocks reviewed (WGOM, GB, SNE and EGOM). The WG did not consider the sensitivity of projections to assumptions, as required by this ToR, and this should be considered before a management track. I agree with the WG recommendation of using WHAM to conduct short-term projections internally. I also agree with the WG recommendation about "monitoring projection accuracy in future assessments and developing methods that continue to improve projections". The ICES analog of WHAM (i.e., SAM) has some capacity to internally model weights-at-age and maturity, via the "biopar" feature, and include uncertainty in these processes in projections. The biopar models are simple. This could be investigated as a future feature for WHAM.

ToR 7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

I agree with the Panels conclusion for this ToR, and the prioritization of WG research recommendations. The Panel also provided some additional recommendations as part of other ToRs above. I think it would have been useful for the Panel to also summarize those additional new recommendations as part of this ToR; however, there was not enough time during the Panel meeting to do this. I also provided some additional research recommendations above that I summarize below in the Conclusions and Recommendations section.

ToR 8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.

I agree with the Panel that this ToR was met for all four stocks. However, specific backup alternatives were not reviewed by the Panel so I suggest that if the proposed assessment approaches are rejected in a future management track assessment, then more investigations will be required to develop a credible backup approach.

ToR 9. Apply the findings of the Atlantic Cod Stock Structure Working Group and identify what assessment approaches the available data can support in defining the appropriate scale of Atlantic cod stock assessment. Consider implications for management processes and other practical limitations in the final units and boundaries used for stock assessments.
I agree with the Panel that this ToR was met in full.

Conclusions and Recommendations

I mostly agree with the Panel conclusions for the WGOM, GB, SNE and EGOM cod stocks. My conclusions for ToR 6 and 8 are a little different than the Panel conclusion. I conclude that:

1. ToR 1 was met.
2. ToR 2 was met. I was not sure what "*Characterize the uncertainty*" meant. Uncertainties in landings and discards were discussed during the review but there is a need to quantify uncertainty in a way that can be included in age/length based stock assessment models.
3. ToR 3 was met in full.
4. ToR 4 was met in full for WGOM cod, and was mostly met for EGOM, GB, and SNE cod. The resulting assessments are accepted for use in subsequent management track processes, although there are some research recommendations that must be addressed for EGOM, GB, and SNE cod before using the models in management track processes.
5. ToR 5 was met in full.
6. ToR 6 was mostly met; however, the WG did not consider the sensitivity of projections to assumptions, as required by this ToR. This should be considered before a management track.
7. ToR 7 was met in full.
8. ToR 8 was mostly met; however, specific backup alternatives were not reviewed by the Panel so I suggest that if the proposed assessment approaches are rejected in a future management track assessment, then more investigations will be required to develop a credible backup approach.
9. ToR 9 was met in full.

The documentation provided was extensive. This included descriptions of commercial and recreational landings and discards, methods used to estimate the age composition of catches and discards, and surveys. We did not have time to evaluate the impacts of the numerous different choices for deriving assessment model inputs, and if there were better options. Reviewing four stock assessments, including revisions to all inputs, was quite ambitious for the time available.

During the panel meeting we examined how well cohorts tracked in commercial and survey age compositions. I conclude that the GB and WGOM fishery age compositions are sufficient to support a cohort-based stock assessment model. There is less information for EGOM and SNE cod, but enough to attempt fitting cohort models. For the WGOM region, I find overall that the NEFSC Fall survey tracks cohorts the best, and the Spring survey is second-best. The GB NEFSC Spring indices and the Canadian Spring survey indices at ages 1-7 have good correlations. The GB NEFSC Fall survey does not seem to track cohorts as well, although relatively weak and strong years classes are evident in the age compositions. The survey cohort information for cod in the EGOM region seems less reliable than in the WGOM or GB regions, although the lag 1 cohort correlations were usually significantly positive at ages 2-5 for the NEFSC surveys. The survey age composition information in the SNE region is limited and has not tracked cohorts well. This could also occur if there have not been relatively weak or strong cohorts in this region in recent years.

Although ToR 4 was not met fully for EGOM, GB, and SNE cod, and there are some research recommendations that must be addressed before using the models in management track processes, I don't think that addressing these recommendations will require substantial new investigations. I also don't think advice will change substantially, given the status of these stocks; however, assessment teams should strive to provide the best assessments possible, regardless of possible stock status conclusions.

I conclude that the WHAM model is a very useful state-space stock assessment package for age-structured assessments. I was impressed with the flexibility of WHAM. I agree with the WG recommendation of using WHAM to conduct short-term projections internally. However, there is a need to extend the approach to include length sample data. At present, WHAM only accommodates age data and this resulted in cases where length samples were not used, or the same age information was used to convert length-based survey indices into age-based ones. These are not "best practice" approaches.

I am very skeptical about life-history methods used to infer values of M. Nonetheless, I find that the base values used in the assessments seemed reasonable. Estimating the overall level of M within age-based stock assessments similar to those for GB and WGOM cod is a challenge, and more so for SNE and EGOM cod with their sparser data. Estimating change in M is more feasible but this did not produce converged models for GB and WGOM cod. However, in most Canadian cod stock assessments, time-varying M is included in assessment models for management advice. I am unsure why this did not work in the GB and WGOM cod assessments, but this was not a focus of our review.

Panel Research Recommendations

1. The Panel recommends targeted sampling in time and areas of cod and seal overlap to better inform seal predation effects in any future work on indicators. (Hammill et al., 2014).
2. The Panel considers that the development of integrated modeling for surveys (e.g., VAST) should be continued, especially considering the postulated distributional shift in cod on GB in spring and concerns about conflicting signals from the NEFSC and DFO spring surveys for this stock.
3. The Panel recommends research continue to standardize and quantify the uncertainty and sampling distribution of age compositions (e.g., Thorson, 2014; Thorson and Haltuch, 2019; Thorson et al., 2023)
4. The Panel recommended that the lack of convergence of WHAM configurations be explored more fully in future research track processes (for example, by resetting parameter bounds if lack of convergence is due to bounds being reached).
5. The Panel recommends that it would be useful if the WHAM model provided leave-one-survey-out diagnostics.
6. For WGOM, the Panel recommends a four block logistic-age based version of the model be tested in future.
7. The Panel recommends further investigations of robust approaches to modeling recruitment in the assessment model. This recommendation was made for WGOM but would apply to all four stocks.
8. The Panel recommends that the effect of removing the correlation between age 1 process errors and those for ages 2-9+ be considered in a future assessment for WGOM.
9. The Panel recommends additional evaluation of the age for a plus group should be conducted, since survey indices and catch information at older model ages did not track cohorts well for most stocks.
10. For GB cod:

a. Further explorations of differences in NEFSC and DFO survey catch rates during 2000-present are required.
b. Investigate a model using unconverted Bigelow indices, similar to WGOM cod. If the differences in NEFSC spring and fall survey indices and the Canadian indices cannot be resolved, then two WHAM models should be formulated (one for US indices, one for Canadian indices) to investigate differences in assessment results and management advice.
c. Jan1 and SSB weights-at-age were calculated using the Rivard approach applied to commercial weights. However, these weights are known to be over-estimates of the stock weights at young ages because of the size-selectivity of commercial fisheries. Survey weights-at-age may be better for stock weights because survey gears may be less size-selective. This should be investigated.

11. The Panel recommends that the recreational LPUE index for SNE cod be re-estimated with zero hauls included.
12. The Panel recommends that criteria be agreed to determine how much bias in self-tests is too large and when this should lead to rejection.
13. The Panel recommends that the poor fit of several indices and the absence of estimates of dead discards in the lobster trap fishery be addressed and that the assessment of EGOM cod must be updated with estimates of dead discards in the lobster trap fishery before the management track assessment.
14. The panel recommends that the issue of what historical period to use in estimating future recruitment in projections be re-examined in the future, possibly as part of a "broader discussion (i.e., across species) of appropriate methods for defining time windows to characterize prevailing conditions for stocks" as recommended by the WG.
15. The panel recommends that in the future, projection assumptions are spelled out clearly for all stocks as was done for the GB stock, i.e., including the type of random effect specified.
16. The Panel suggests that age-length-based models that can use length data directly as sampled should be investigated in the medium-long term.

CIE Reviewer Additional Recommendations for Research Track

1. Tables of parameter estimates for final models should be provided. Comparing variance parameters between runs can give an overall description of how well a model fits various data sources, and how much process error was required to achieve these fits.
2. Comparison plots of observed and model-predicted cohort catch curves are a useful diagnostic of model fit.

Longer-term and generic research recommendations (not part of the cod research track process).

1. When there are spatial differences in the sampling domains of surveys, and changes in the spatial distribution of a stock over time, then this can create spurious trends in survey indices that space-aggregated stock assessment models have no ability to sort out. A spatial integrated analysis of the survey information is a good solution for this problem. A computational challenge will be integrating survey catch-at-length with an appropriate count data model. Fine-scale space-time-length models are often computationally prohibitive. However, this could include the unconverted Bigelow and Albatross survey catches and provide an opportunity to integrate comparative fishing data in the model, maybe as coming from a "dummy" survey and year, similar to Benoît and Cadigan (2014).

This could also provide an opportunity to examine if there were important spatial differences in the relative efficiency of the Bigelow survey compared to the Albatross.

2. Development of the "growth" branch of WHAM should continue so that, in future cod research tracks, the survey and catch length sampling can be better utilized. This is longer-term research because WHAM may need to be extended to be fully age and length structured and include time-varying growth rates. Much research will be required to understand what the more useful model configurations in this extended setting are. Such a model will remove the need to use the same age data to produce age-based indices from different surveys, which is using data twice and creates complicated correlations in survey indices. With a length-age structured model, it will be possible to include unconverted Bigelow and Albatross survey catches for model fitting, and the comparative fishing data can be included as a likelihood component to constrain the ratio of catchabilities (i.e., the length-based relative efficiency) for these survey series. A complication is that the calibration curves in Miller (2013) may only be appropriate for the stock regions the comparative fishing data were obtained from, and these calibration curves may not be appropriate for each of the four cod stocks. This issue requires further investigation (see Res. Recomm. #2).
3. There is a need for a North Atlantic system-wide study of how to model M in age-based and state-space stock assessments with data similar to GB and WGOM cod, and also cod in NAFO Divisions 2J3KL, 3NO, 3Ps, 3Pn4RS, and 4T in Canada, and perhaps some ICES cod stocks, and possibly other groundfish stocks in these regions. Extending investigations like Miller and Hyun (2018) to include the above stocks is useful, and finding criteria that can indicate situations where M or time-change in M can be identified (e.g., Clark, 2022) is useful.
4. Uncertainties in landings and discards need to be quantified in a way that can be included in age/length based stock assessment models.
5. I agree with the WG that "a broader discussion (i.e., across species) of appropriate methods for defining time windows to characterize prevailing conditions for stocks is needed in the region." I suggest that this broader discussion should be extended to include how to conduct long-term stochastic projections to define SSB_{MSY} and even F_{MSY}. This is related to the WG recommendation about "monitoring projection accuracy in future assessments and developing methods that continue to improve projections". The ICES analog of WHAM (i.e., SAM) has some capacity to internally model weights-at-age and maturity, via the "biopar" feature, and include uncertainty in these processes in projections. This could be investigated as a future feature for WHAM.
6. Comparing recruitment trends and Zs from surveys before including them in an assessment model is useful. This could be achieved by applying a SURBA model to each survey, and then comparing results.
7. Further analyses of survey and commercial weights-at-age are required, to get better estimates of stock weights-at-age. More detailed investigations of the weight-length relationships for the US Atlantic cod stocks should continue and include a focus on how many fish are in very poor condition.
8. The efficacy of simulation self-tests requires additional research in the state-space stock assessment context.
9. Detecting change in the stock-recruitment relationship should be based on stock-recruitment residuals. We expect the recruitment rate (R/SSB) to decline as SSB increases, which is why it is not sufficient to examine only time-trends in recruitment rates, and especially time-trends in recruitment.

10. It seems reasonable that age zero was not used as a recruitment index in the WGOM assessment model, but it also seems reasonable to re-visit this in the future. This may require a model with time-varying M, at least at age zero.

Appendix 1: Bibliography of materials provided for review

Assessment Report

Research Track Assessment of Atlantic Cod. Atlantic Cod Research Track Working Group. July 14, 2023. 441 pp.

Working Papers:

WP1 Stakeholder Meeting Summary
WP2 Development of Ecosystem Indicators
WP3 Environmental Influences on Cod
WP4 Stakeholder Meeting 2 Summary
WP5 Rec Discard Mortality
WP6 FDD Exploration
WP7 Cod LPUE
WP8 NEFSC Trawl Survey Expanded Figs.
WP9 Survey Time Series Correlations
WP10 Integrated Survey Indices (VAST)
WP11 EGOM Sentinel Index Modification
WP12 Time Varying Cod Maturity
WP13 Time Varying Cod LW
WP14 Estimating Cod M by Stock
WP15 Atlantic Cod Model Selection Procedure
WP16 EGOM Assessment Model ToR 4
WP17 WGOM Assessment Model ToR 4
WP18 SNE Assessment Model ToR 4
WP19 GB Assessment Model ToR 4
WP20 Reference Points

Other Papers

Stock, B.C., Xu, H., Miller, T.J., Thorson, J.T. and Nye, J.A., 2021. Implementing two-dimensional autocorrelation in either survival or natural mortality improves a state-space assessment model for Southern New England-Mid Atlantic yellowtail flounder. Fisheries Research, 237, p.105873.

Stock, B.C. and Miller, T.J., 2021. The Woods Hole Assessment Model (WHAM): a general state-space assessment framework that incorporates time-and age-varying processes via random effects and links to environmental covariates. Fisheries Research, 240, p.105967.

Technical Documentation for ASAP Version 3.0. NOAA Fisheries Toolbox. September 2012. 71 pp.

User Manual for ASAP 3. September 2012

Appendix 2: CIE Statement of Work

Performance Work Statement (PWS)
National Oceanic and Atmospheric Administration (NOAA)
National Marine Fisheries Service (NMFS)
Center for Independent Experts (CIE) Program
External Independent Peer Review
Under Contract #1305M219DNFFK0025

Atlantic Cod Research Track Peer Review
July 31 – August 3, 2023

Background

The National Marine Fisheries Service (NMFS) is mandated by the Magnuson-Stevens Fishery Conservation and Management Act, Endangered Species Act, and Marine Mammal Protection Act to conserve, protect, and manage our nation's marine living resources based upon the best scientific information available (BSIA). NMFS science products, including scientific advice, are often controversial and may require timely scientific peer reviews that are strictly independent of all outside influences. A formal external process for independent expert reviews of the agency's scientific products and programs ensures their credibility. Therefore, external scientific peer reviews have been and continue to be essential to strengthening scientific quality assurance for fishery conservation and management actions.

Scientific peer review is defined as the organized review process where one or more qualified experts review scientific information to ensure quality and credibility. These expert(s) must conduct their peer review impartially, objectively, and without conflicts of interest. Each reviewer must also be independent from the development of the science, without influence from any position that the agency or constituent groups may have. Furthermore, the Office of Management and Budget (OMB), authorized by the Information Quality Act, requires all federal agencies to conduct peer reviews of highly influential and controversial science before dissemination, and that peer reviewers must be deemed qualified based on the OMB Peer Review Bulletin standards[1].

Scope

The Research Track Peer Review meeting is a formal, multiple-day meeting of stock assessment experts who serve as a panel to peer-review tabled stock assessments and models. The research track peer review is the cornerstone of the Northeast Region Coordinating Council stock assessment process, which includes assessment development, and report preparation (which is done by Working Groups or Atlantic States Marine Fisheries Commission (ASMFC) technical committees), assessment peer review (by the peer review panel), public presentations, and document publication. The results of this peer review will be incorporated into future management track assessments, which serve as the basis for developing fishery management recommendations.

The purpose of this meeting will be to provide an external peer review of the Atlantic cod stocks. The requirements for the peer review follow. This Performance Work Statement (PWS) also includes: **Annex 1**: TORs for the research track, which are the responsibility of the analysts; **Annex 2:** a draft meeting agenda; **Annex 3:** Individual Independent Review Report Requirements; and **Annex 4:** Peer Reviewer Summary Report Requirements.

---

[1] https://www.whitehouse.gov/wp-content/uploads/legacy_drupal_files/omb/memoranda/2005/m05-03.pdf

Requirements
NMFS requires three reviewers under this contract (i.e. subject to CIE standards for reviewers) to participate in the panel review. The chair, who is in addition to the three reviewers, will be provided by either the New England or Mid-Atlantic Fishery Management Council's Science and Statistical Committee; although the chair will be participating in this review, the chair's participation (i.e. labor and travel) is not covered by this contract.

Each reviewer will write an individual review report in accordance with the PWS, OMB Guidelines, and the TORs below. Modifications to the PWS and ToRs cannot be made during the peer review, and any PWS or ToRs modifications prior to the peer review shall be approved by the Contracting Officer's Representative (COR) and the CIE contractor. All TORs must be addressed in each reviewer's report. The reviewers shall have working knowledge and recent experience in the use and application of index-based, age-based, and state-space stock assessment models, including familiarity with retrospective patterns, model diagnostics from various population models, and how catch advice is provided from stock assessment models. In addition, knowledge and experience with simulation analyses is helpful.

Tasks for Reviewers

- Review the background materials and reports prior to the review meeting
    - Two weeks before the peer review, the Assessment Process Lead will electronically disseminate all necessary background information and reports to the CIE reviewers for the peer review.
- Attend and participate in the panel review meeting
    - The meeting will consist of presentations by NMFS and other scientists, stock assessment authors and others to facilitate the review, to provide any additional information required by the reviewers, and to answer any questions from reviewers
- Conduct an independent peer review in accordance with the requirements specified in this PWS and TORs, in adherence with the required formatting and content guidelines.
- Reviewers are not required to reach a consensus. Individual reviewer perspectives should be provided in their individual reports, and any lack of consensus should be clearly described in the panel's summary report.
- Each reviewer shall assist the Peer Review Panel Chair with contributions to the Peer Review Panel's Summary Report.
- Deliver individual Independent Reviewer Reports to NMFS according to the specified milestone dates.
- This report should explain whether each research track Term of Reference was or was not completed successfully during the peer review meeting, using the criteria specified below in the "Tasks for Peer Review Panel."
- If any existing Biological Reference Points (BRP) or their proxies are considered inappropriate, the Independent Report should include recommendations and justification for suitable alternatives. If such alternatives cannot be identified, then the report should indicate that the existing BRPs are the best available at this time.
- During the meeting, additional questions that were not in the Terms of Reference but that are directly related to the assessments and research topics may be raised. Comments on these questions should be included in a separate section at the end of the Independent Report produced by each reviewer.

- The Independent Report can also be used to provide greater detail than the Peer Reviewer Summary Report on specific stock assessment Terms of Reference or on additional questions raised during the meeting.

Tasks for Review panel

- During the peer review meeting, the panel is to determine whether each research track Term of Reference (TOR) was or was not completed successfully. To make this determination, panelists should consider whether the work provides a scientifically credible basis for developing fishery management advice. Criteria to consider include: whether the data were adequate and used properly, the analyses and models were carried out correctly, and the conclusions are correct/reasonable. If alternative assessment models and model assumptions are presented, evaluate their strengths and weaknesses and then recommend which, if any, scientific approach should be adopted. Where possible, the Peer Review Panel chair shall identify or facilitate agreement among the reviewers for each research track TOR.
- If the panel rejects any of the current BRP or BRP proxies (for B_{MSY} and F_{MSY} and MSY), the panel should explain why those particular BRPs or proxies are not suitable, and the panel should recommend suitable alternatives. If such alternatives cannot be identified, then the panel should indicate that the existing BRPs or BRP proxies are the best available at this time.
- Each reviewer shall complete the tasks in accordance with the PWS and Schedule of Milestones and Deliverables below.

Tasks for Peer Review Panel chair and reviewers combined:
Review the Report of Atlantic Cod Research Track Working Group.

The Peer Review Panel Chair, with the assistance from the reviewers, will write the Peer Reviewer Summary Report. Each reviewer and the chair will discuss whether they hold similar views on each research track Term of Reference and whether their opinions can be summarized into a single conclusion for all or only for some of the Terms of Reference of the peer review meeting. For terms where a similar view can be reached, the Peer Reviewer Summary Report will contain a summary of such opinions.

The chair's objective during this Peer Reviewer Summary Report development process will be to identify or facilitate the finding of an agreement rather than forcing the panel to reach an agreement. Again, the CIE reviewers are not required to reach a consensus. The chair will take the lead in editing and completing this report. The chair may express their opinion on each research track Term of Reference, either as part of the group opinion, or as a separate minority opinion. The Peer Reviewer Summary Report will not be submitted, reviewed, or approved by the Contractor.

Place of Performance
The place of performance shall be remote, via WebEx video conferencing.

Period of Performance
The period of performance shall be from the time of award through October 2023. Each reviewer's duties shall not exceed **14** days to complete all required tasks.

Schedule of Milestones and Deliverables: The contractor shall complete the tasks and deliverables in accordance with the following schedule.

Schedule	Milestones and Deliverables
Within 2 weeks of award	Contractor selects and confirms reviewers
Approximately 2 weeks later	Contractor provides the pre-review documents to the reviewers
July 31 – August 3, 2023	Panel review meeting
Approximately 2 weeks later	Reviewers submit draft peer-review reports to the contractor for quality assurance and review
Within 2 weeks of receiving draft reports	Contractor submits final reports to the Government

* The Peer Reviewer Summary Report will not be submitted to, reviewed, or approved by the Contractor.

Applicable Performance Standards
The acceptance of the contract deliverables shall be based on three performance standards:
(1) The reports shall be completed in accordance with the required formatting and content (2) The reports shall address each TOR as specified (3) The reports shall be delivered as specified in the schedule of milestones and deliverables.

Travel
No travel is necessary, as this meeting is being held remotely.

Restricted or Limited Use of Data
The contractors may be required to sign and adhere to a non-disclosure agreement.

NMFS Project Contact
Michele Traver, NEFSC Assessment Process Lead
Northeast Fisheries Science Center
166 Water Street, Woods Hole, MA 02543
Michele.Traver@noaa.gov

Annex 1. Generic Research Track Terms of Reference

1. Identify relevant ecosystem and climate influences on the stock. Characterize the uncertainty in the relevant sources of data and their link to stock dynamics. Consider findings, as appropriate, in addressing other TORs. Report how the findings were considered under impacted TORs.
2. Estimate catch from all sources including landings and discards. Describe the spatial and temporal distribution of landings, discards, and fishing effort. Characterize the uncertainty in these sources of data.
3. Present the survey data used in the assessment (e.g., indices of relative or absolute abundance, recruitment, state surveys, age-length data, application of catchability and calibration studies, etc.) and provide a rationale for which data are used. Describe the spatial and temporal distribution of the data. Characterize the uncertainty in these sources of data.
4. Use appropriate assessment approach to estimate annual fishing mortality, recruitment and stock biomass (both total and spawning stock) for the time series, and estimate their uncertainty. Compare the time series of these estimates with those from the previously accepted assessment(s). Evaluate a suite of model fit diagnostics (e.g., residual patterns, sensitivity analyses, retrospective patterns), and (a) comment on likely causes of problematic issues, and (b), if possible and appropriate, account for those issues when providing scientific advice and evaluate the consequences of any correction(s) applied.
5. Update or redefine status determination criteria (SDC; point estimates or proxies for BMSY, BTHRESHOLD, FMSY and MSY reference points) and provide estimates of those criteria and their uncertainty, along with a description of the sources of uncertainty. If analytic model-based estimates are unavailable, consider recommending alternative measurable proxies for reference points. Compare estimates of current stock size and fishing mortality to existing, and any redefined, SDCs.
6. Define appropriate methods for producing projections; provide justification for assumptions of fishery selectivity, weights at age, maturity, and recruitment; and comment on the reliability of resulting projections considering the effects of uncertainty and sensitivity to projection assumptions.
7. Review, evaluate, and report on the status of research recommendations from the last assessment peer review, including recommendations provided by the prior assessment working group, peer review panel, and SSC. Identify new recommendations for future research, data collection, and assessment methodology. If any ecosystem influences from TOR 2 could not be considered quantitatively under that or other TORs, describe next steps for development, testing, and review of quantitative relationships and how they could best inform assessments. Prioritize research recommendations.

8. Develop a backup assessment approach to providing scientific advice to managers if the proposed assessment approach does not pass peer review or the approved approach is rejected in a future management track assessment.
9. Apply the findings of the Atlantic Cod Stock Structure Working Group and identify what assessment approaches the available data can support in defining the appropriate scale of Atlantic cod stock assessment. Consider implications for management processes and other practical limitations in the final units and boundaries used for stock assessments.

Research Track TORs:

General Clarification of Terms that may be Used in the Research Track Terms of Reference

Guidance to Peer Review Panels about "Number of Models to include in the Peer Reviewer Report":

In general, for any TOR in which one or more models are explored by the Working Group, give a detailed presentation of the "best" model, including inputs, outputs, diagnostics of model adequacy, and sensitivity analyses that evaluate robustness of model results to the assumptions. In less detail, describe other models that were evaluated by the Working Group and explain their strengths, weaknesses and results in relation to the "best" model. If selection of a "best" model is not possible, present alternative models in detail, and summarize the relative utility each model, including a comparison of results. It should be highlighted whether any models represent a minority opinion.

On "Acceptable Biological Catch" (DOC Nat. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

Acceptable biological catch (ABC) is a level of a stock or stock complex's annual catch that accounts for the scientific uncertainty in the estimate of Overfishing Limit (OFL) and any other scientific uncertainty…" *(p. 3208) [In other words, OFL ≥ ABC.]*

ABC for overfished stocks. For overfished stocks and stock complexes, a rebuilding ABC must be set to reflect the annual catch that is consistent with the schedule of fishing mortality rates in the rebuilding plan. *(p. 3209)*

NMFS expects that in most cases ABC will be reduced from OFL to reduce the probability that overfishing might occur in a year. (p. 3180)

ABC refers to a level of ''catch'' that is ''acceptable'' given the ''biological'' characteristics of the stock or stock complex. As such, Optimal Yield (OY) does not equate with ABC. The specification of OY is required to consider a variety of factors, including social and economic factors, and the protection of marine ecosystems, which are not part of the ABC concept. (p. 3189)

On "Vulnerability" (DOC Natl. Stand. Guidelines. Fed. Reg., v. 74, no. 11, 1-16-2009):

“Vulnerability. A stock’s vulnerability is a combination of its productivity, which depends upon its life history characteristics, and its susceptibility to the fishery. Productivity refers to the capacity of the stock to produce Maximum Sustainable Yield (MSY) and to recover if the population is depleted, and susceptibility is the potential for the stock to be impacted by the fishery, which includes direct captures, as well as indirect impacts to the fishery (e.g., loss of habitat quality).” (p. 3205)

Participation among members of a Research Track Working Group:

Anyone participating in peer review meetings that will be running or presenting results from an assessment model is expected to supply the source code, a compiled executable, an input file with the proposed configuration, and a detailed model description in advance of the model meeting. Source code for NOAA Toolbox programs is available on request. These measures allow transparency and a fair evaluation of differences that emerge between models.

Annex 2. Draft Review Meeting Agenda
{Final Meeting agenda to be provided at time of award}

Atlantic Cod Track Assessment Peer Review Meeting

July 31 –August 3 , 2023

WebEx link: TBD

DRAFT AGENDA* (v. 4/27/2023)
**All times are approximate, and may be changed at the discretion of the Peer Review Panel chair. The meeting is open to the public; however, during the Report Writing sessions we ask that the public refrain from engaging in discussion with the Peer Review Panel.*
Monday, July 31, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:30 a.m.	Welcome/Logistics Introductions/Agenda /Conduct of Meeting	Michele Traver, Assessment Process Lead Russ Brown, PopDy Branch Chief Panel Chair	
9:30 a.m. - 10:30 a.m.	TOR #1		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #1 cont.		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #2		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #2 cont.		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	

5 p.m.	Adjourn		

Tuesday, August 1, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead Panel Chair	
9:15 a.m. - 10:30 a.m.	TOR #3		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #3 cont.		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #4		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #4 cont.		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Wednesday, August 2, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 9:15 a.m.	Welcome/Logistics	Michele Traver, Assessment Process Lead	

		Panel Chair	
9:15 a.m. - 10:30 a.m.	TOR #5		
10:30 a.m. - 10:45 a.m.	Break		
10:45 a.m. - 11:45 a.m.	TOR #6		
11:45 a.m. - 12:15 p.m.	Discussion/Summary	Review Panel	
12:15 p.m. - 12:30 p.m.	Public Comment	Public	
12:30 p.m. - 1:30 p.m.	Lunch		
1:30 p.m. - 3 p.m.	TOR #7		
3 p.m. - 3:15 p.m.	Break		
3:15 p.m. - 4:15 p.m.	TOR #8-9		
4:15 p.m. - 4:45 p.m.	Discussion/Summary	Review Panel	
4:45 p.m. - 5 p.m.	Public Comment	Public	
5 p.m.	Adjourn		

Thursday, August 3, 2023

Time	Topic	Presenter(s)	Notes
9 a.m. - 5 p.m.	Report Writing	Review Panel	

Annex 3. Individual Independent Peer Reviewer Report Requirements

1. The independent Peer Reviewer report shall be prefaced with an Executive Summary providing a concise summary of whether they accept or reject the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.).

2. The report must contain a background section, description of the individual reviewers' roles in the review activities, summary of findings for each TOR in which the weaknesses and strengths are described, and conclusions and recommendations in accordance with the TORs. The independent report shall be an independent peer review, and shall not simply repeat the contents of the Peer Reviewer Summary Report.

   a. Reviewers should describe in their own words the review activities completed during the panel review meeting, including a concise summary of whether they accept or reject the work that they reviewed, and explain their decisions (strengths, weaknesses of the analyses, etc.), conclusions, and recommendations.

   b. Reviewers should discuss their independent views on each TOR even if these were consistent with those of other panelists, but especially where there were divergent views.

   c. Reviewers should elaborate on any points raised in the Peer Reviewer Summary Report that they believe might require further clarification.

   d. The report may include recommendations on how to improve future assessments.

3. The report shall include the following appendices:

Appendix 1: Bibliography of materials provided for review
Appendix 2: A copy of this Performance Work Statement
Appendix 3: Panel membership or other pertinent information from the panel review meeting.

Annex 4. Peer Reviewer Summary Report Requirements

1. The main body of the report shall consist of an introduction prepared by the Research Track Peer Review Panel chair that will include the background and a review of activities and comments on the appropriateness of the process in reaching the goals of the peer review meeting. Following the introduction, for each assessment /research topic reviewed, the report should address whether or not each Term of Reference of the Research Track Working Group was completed successfully. For each Term of Reference, the Peer Reviewer Summary Report should state why that Term of Reference was or was not completed successfully. It should also include whether they **accept or reject** the work that they reviewed, with an explanation of their decision (strengths, weaknesses of the analyses, etc.)

   To make this determination, the peer review panel chair and reviewers should consider whether or not the work provides a scientifically credible basis for developing fishery management advice. If the reviewers and peer review panel chair do not reach an agreement on a Term of Reference, the report should explain why. It is permissible to express majority as well as minority opinions.

   The report may include recommendations on how to improve future assessments.

2. If any existing Biological Reference Points (BRPs) or BRP proxies are considered inappropriate, include recommendations and justification for alternatives. If such alternatives cannot be identified, then indicate that the existing BRPs or BRP proxies are the best available at this time.

3. The report shall also include the bibliography of all materials provided during the peer review meeting, and relevant papers cited in the Peer Reviewer Summary Report, along with a copy of the CIE Performance Work Statement.

   The report shall also include as a separate appendix the assessment Terms of Reference used for the peer review meeting, including any changes to the Terms of Reference or specific topics/issues directly related to the assessments and requiring Panel advice.

Appendix 3: Panel Membership or other pertinent information from the panel review meeting.

DFO - Department of Fisheries and Oceans (Canada)
GARFO - Greater Atlantic Regional Fisheries Office
GMRI - Gulf of Maine Research Institute
MADMF - Massachusetts Division of Marine Fisheries
MEDMR - Maine Department of Marine Resources
NEFMC - New England Fisheries Management Council
NEFSC - Northeast Fisheries Science Center
SMAST - University of Massachusetts School of Marine Science and Technology
~~~~~~

Panel Members

JJ Maguire – Chair
Steven Holmes - CIE Panel
Noel Cadigan - CIE Panel
Coby Needle - CIE Panel

Panel Advisors

Russ Brown - NEFSC, Population Dynamics Branch Chief
Michele Traver - NEFSC, Assessment Process Lead

Stock Assessment Team Members

Lisa Kerr - University of Maine
Irene Andruschchenko - DFO
Steve Cadrin - SMAST
Micah Dean - MADMF
Alex Hansell - NEFSC
Scott Large – NEFSC
Rich McBride - NEFSC
Charles Perretti - NEFSC
Kathy Sosebee – NEFSC

Other Participants

Alex Dunn - NEFSC
Alicia Miller - NEFSC
Alison Frey - NEFSC
Amanda Hart - GMRI
Angela Forristall - NEFMC Staff
Andy Jones - NEFSC
Anna Mercer - NEFSC
Brian Linton - NEFSC
Burton Shank - NEFSC
~~~~~~

Caira Clark - Nature Conservancy of Canada
Carla Guenther - Maine Center for Coastal Fisheries
Cate O'Keefe - NEFMC Executive Director
Charles Adams - NEFSC
Chris Kellogg - NEFMC Deputy Director
Chris Legault - NEFSC
Cole Carrano - SMAST
Dave McElroy - NEFSC
Doug Butterworth - University of Cape Town (South Africa)
Frank Blount - Frances Fleet
Gareth Lawson - Conservation Law Foundation
Gary Nelson - MADMF
Jackie Odell - Northeast Seafood Coalition
Jamie Behan - GMRI
Jamie Cournane - NEFMC Staff
Jessica Blaylock - NEFSC
John Pappalardo - Cape Cod Hook Fishermen's Association
Jon Deroba - NEFSC
Julie Nieland - NEFSC
Katie Lankowicz - GMRI
Kelly Whitmore - MADMF
Kiersten Curti - NEFSC
Kristan Blackhart - NEFSC
Libby Etrie - NEFMC Member
Lisa Hendrickson - NEFSC
Liz Brooks - NEFSC
Liz Sullivan - GARFO
Mark Terceiro - NEFSC
Max Grezlik - SMAST
Melanie Barrett - DFO
Melanie Griffin - MADMF
Nicholas Calabrese - SMAST
Paul Nitschke - NEFSC
Rebecca Peters - MEDMR
Rebecca Rademeyer - Independent Consultant (South Africa)
Rick Bellavance - NEFMC Member
Robin Frede - NEFMC Staff
Robyn Linner - Stony Brook University
Roger Brothers - GMRI
Spencer Talmage - GARFO
Susan Wigley - NEFSC
Tara Dolan - MADMF
Tim Barrett - DFO
Tim Miller - NEFSC
Tim O'Donnell - Gloucester Marine Genomics Institute
Toni Chute - NEFSC

Yanjun Wang – DFO

Appendix 4. Additional figures

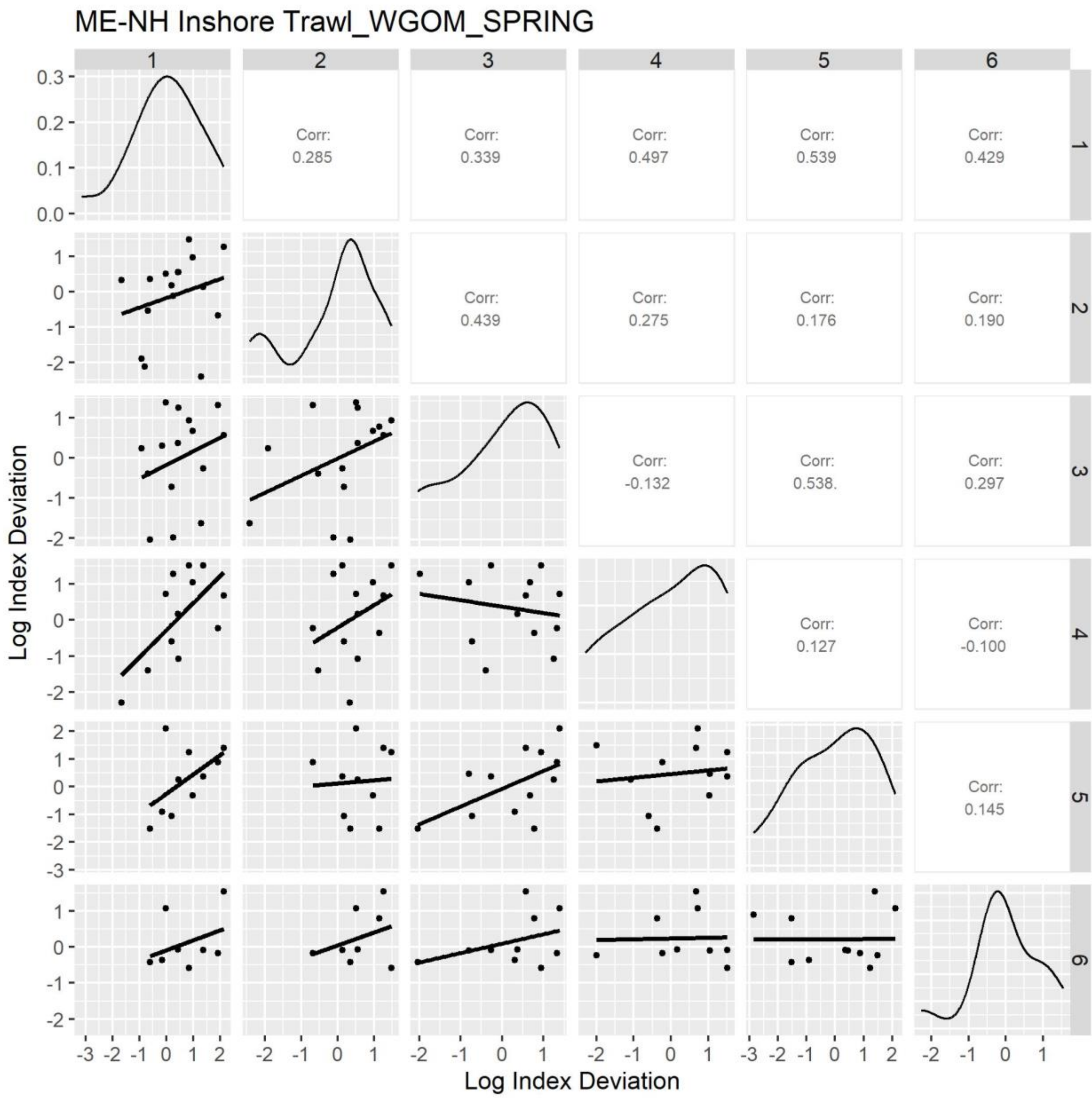

Figure 7: Between-age pairwise correlation plots of survey indices of cohort size. Survey ages are the rows and columns. Spearman's rank correlations are shown in the upper panels. The survey is listed at the top.

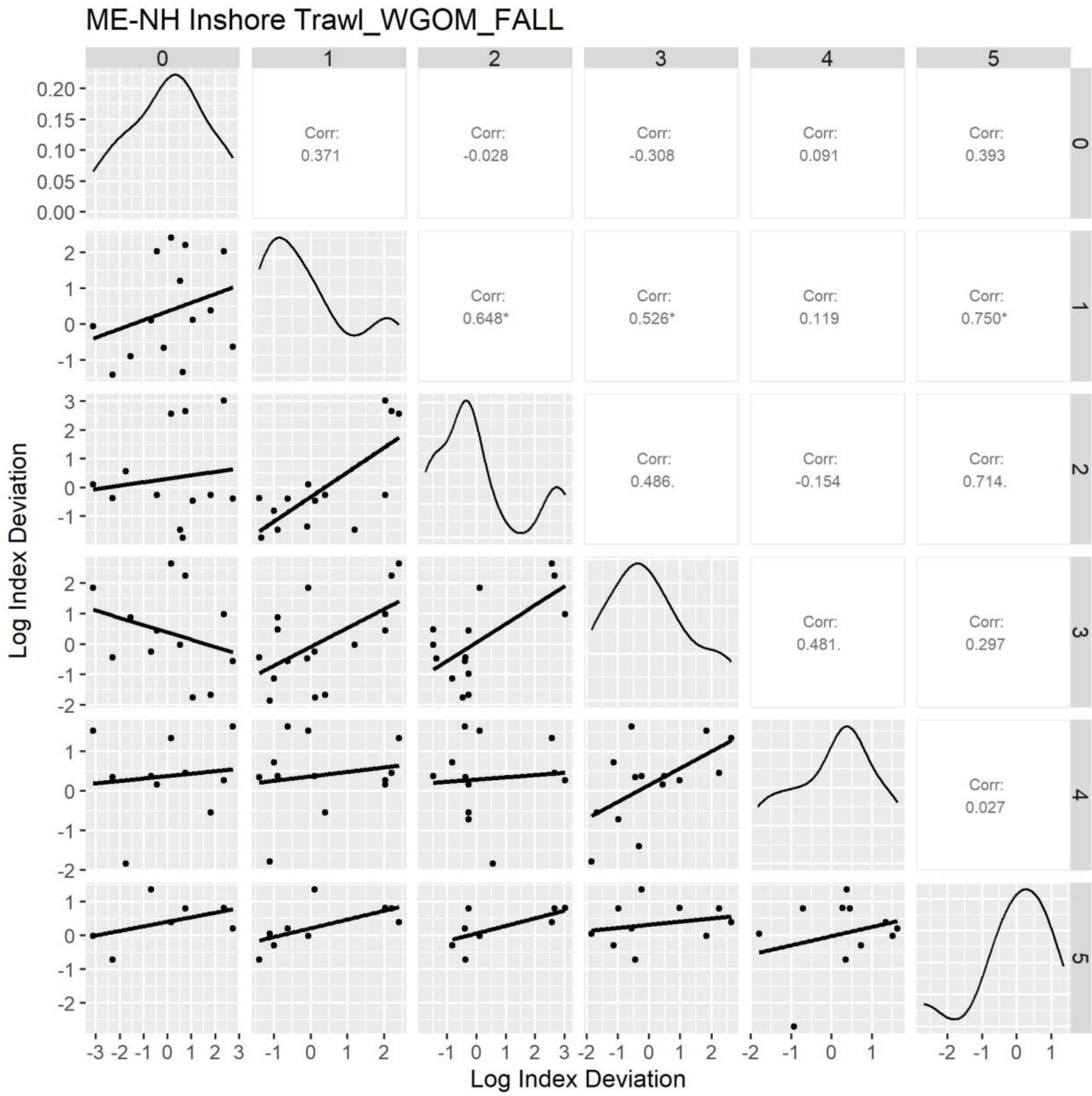

Figure 8: See Figure 7 for details.

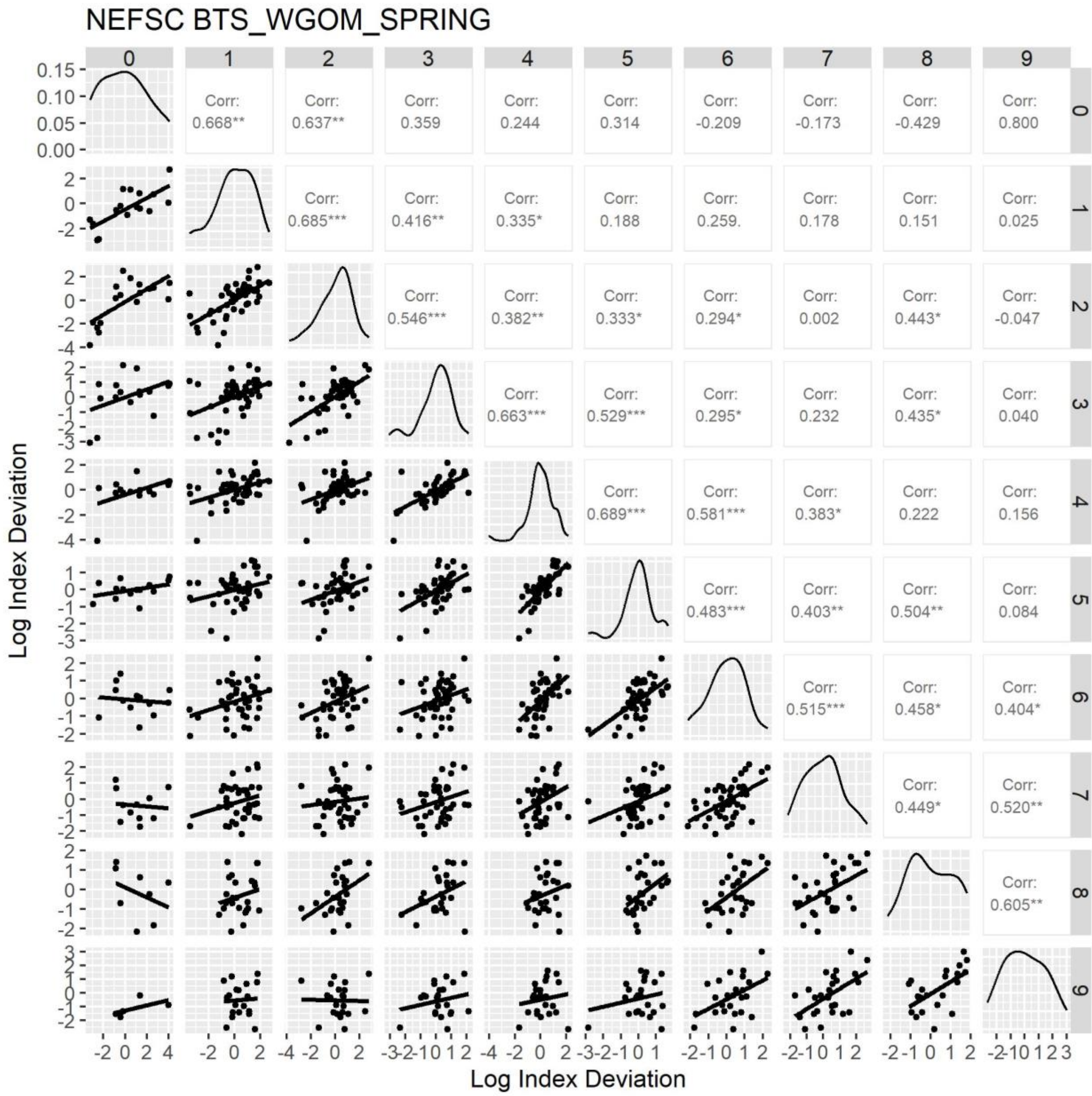

Figure 9: See Figure 7 for details.

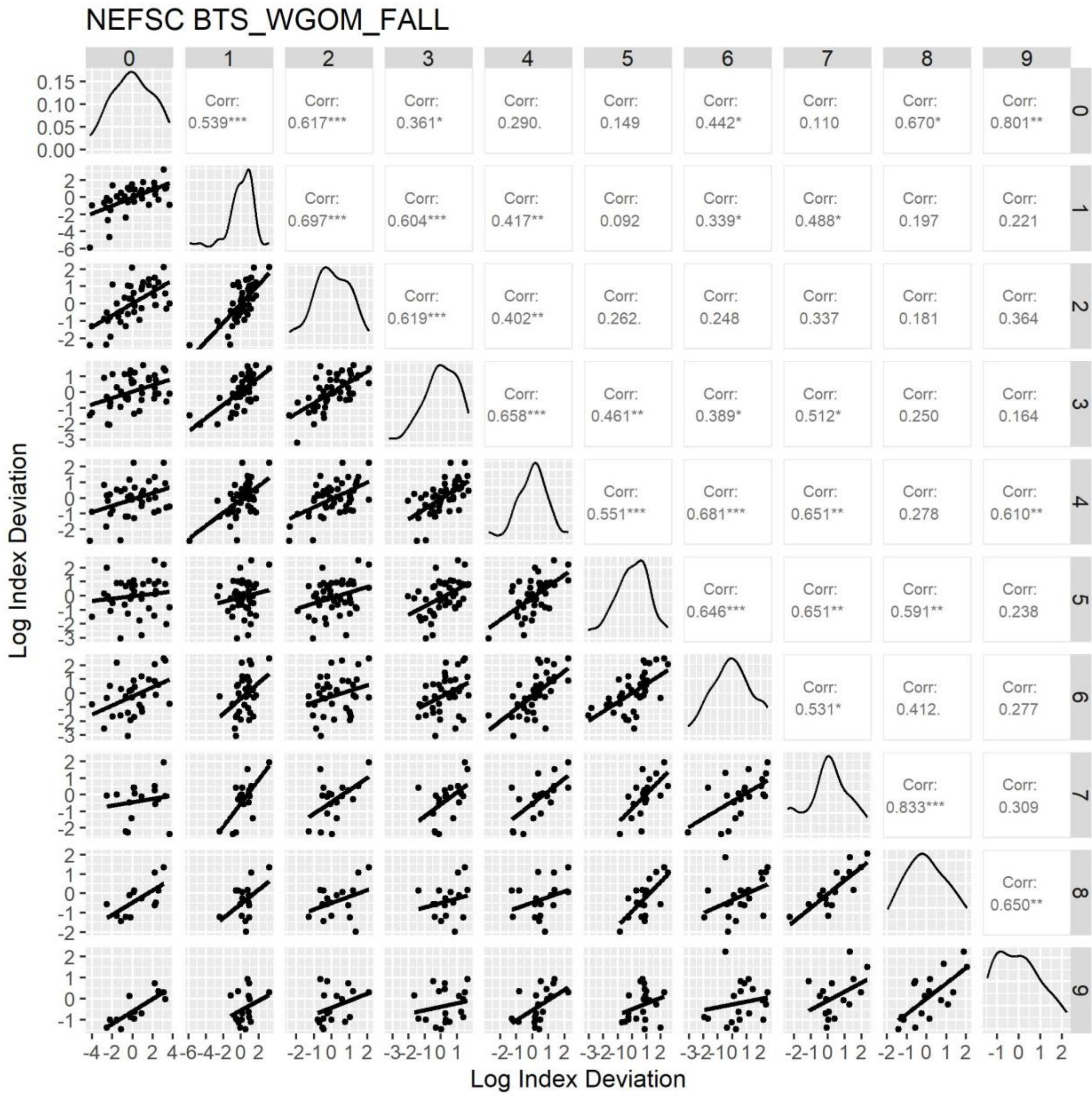

Figure 10: See Figure 7 for details.

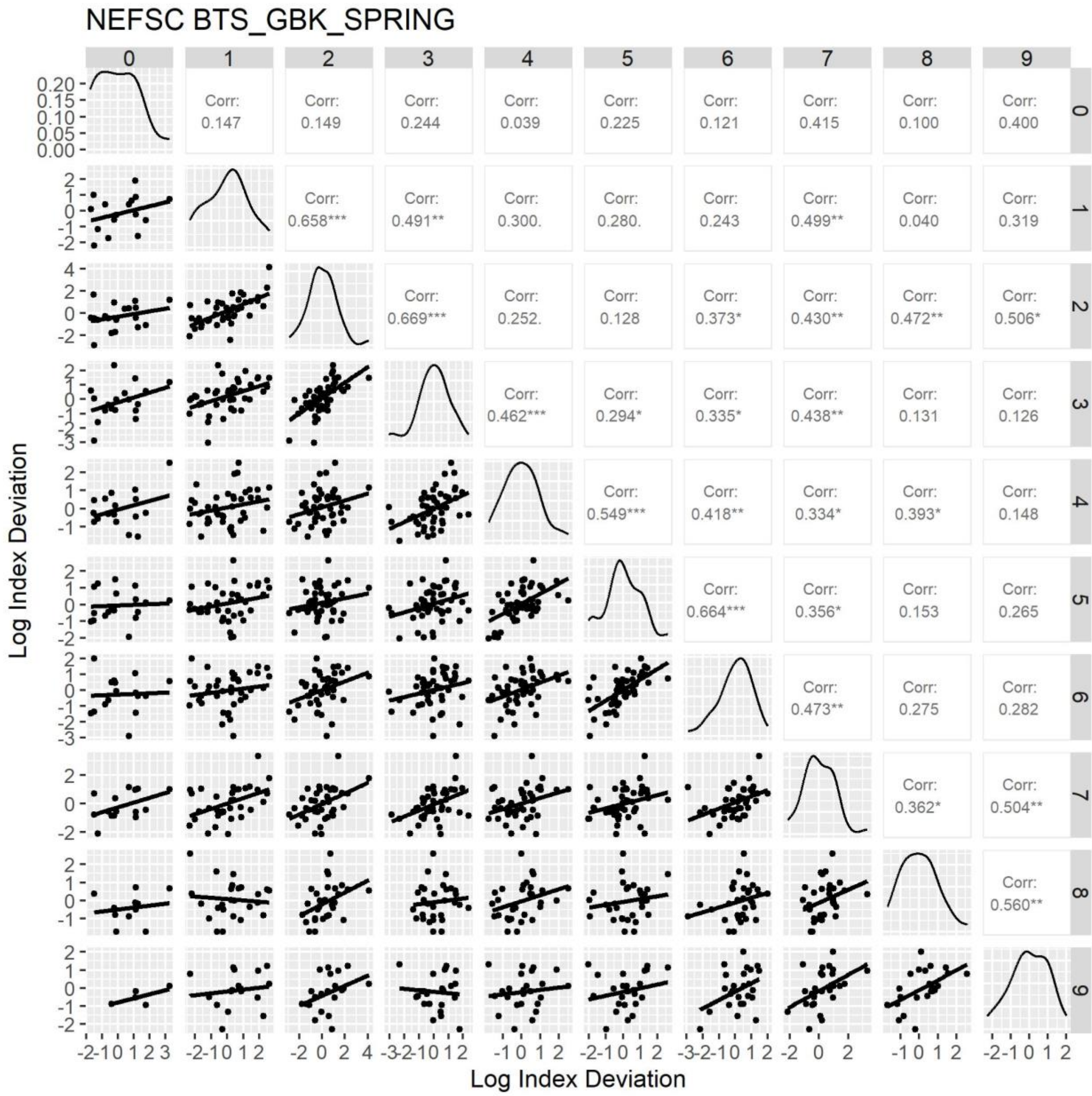

Figure 11: See Figure 7 for details.

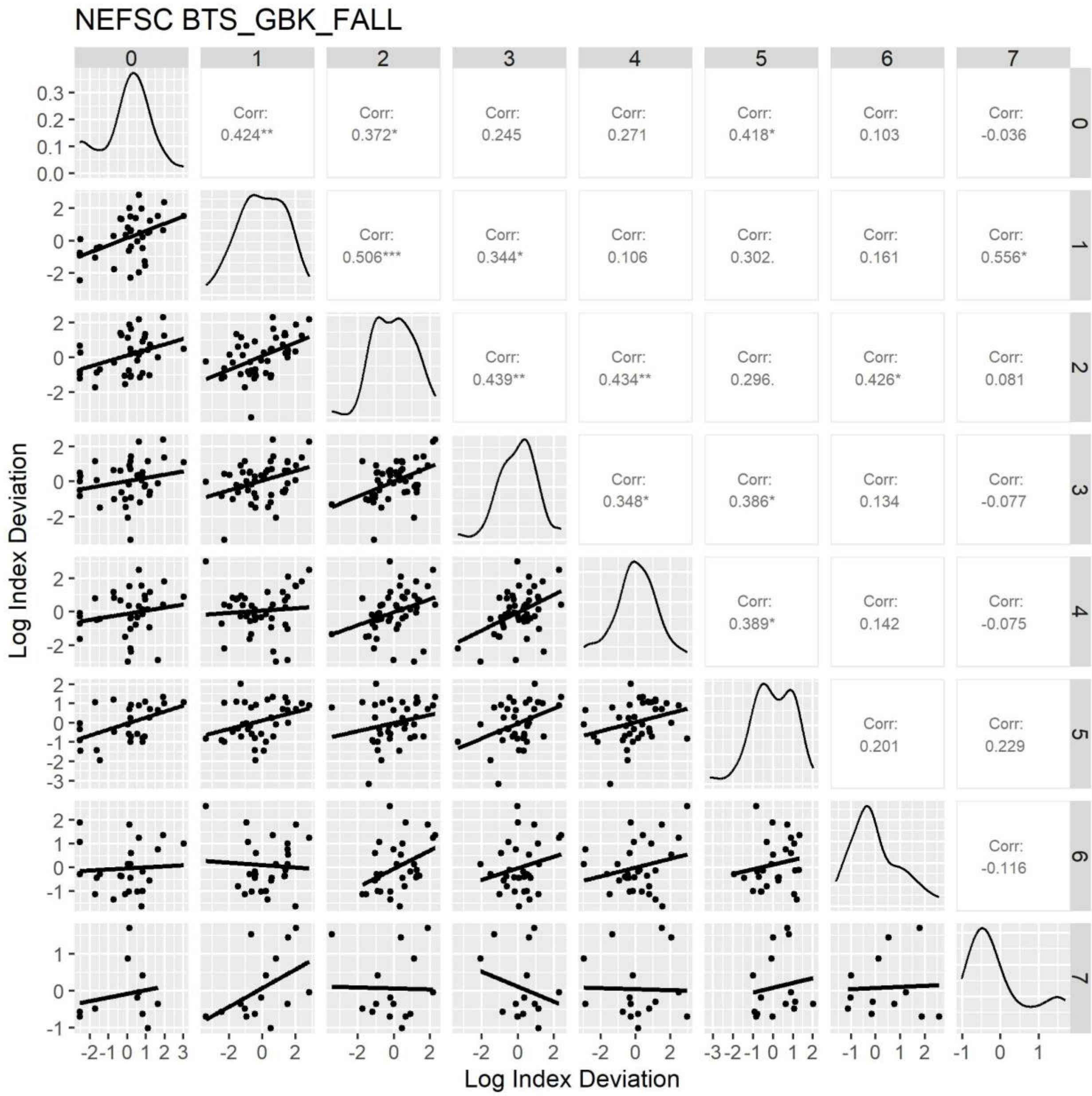

Figure 12: See Figure 7 for details.

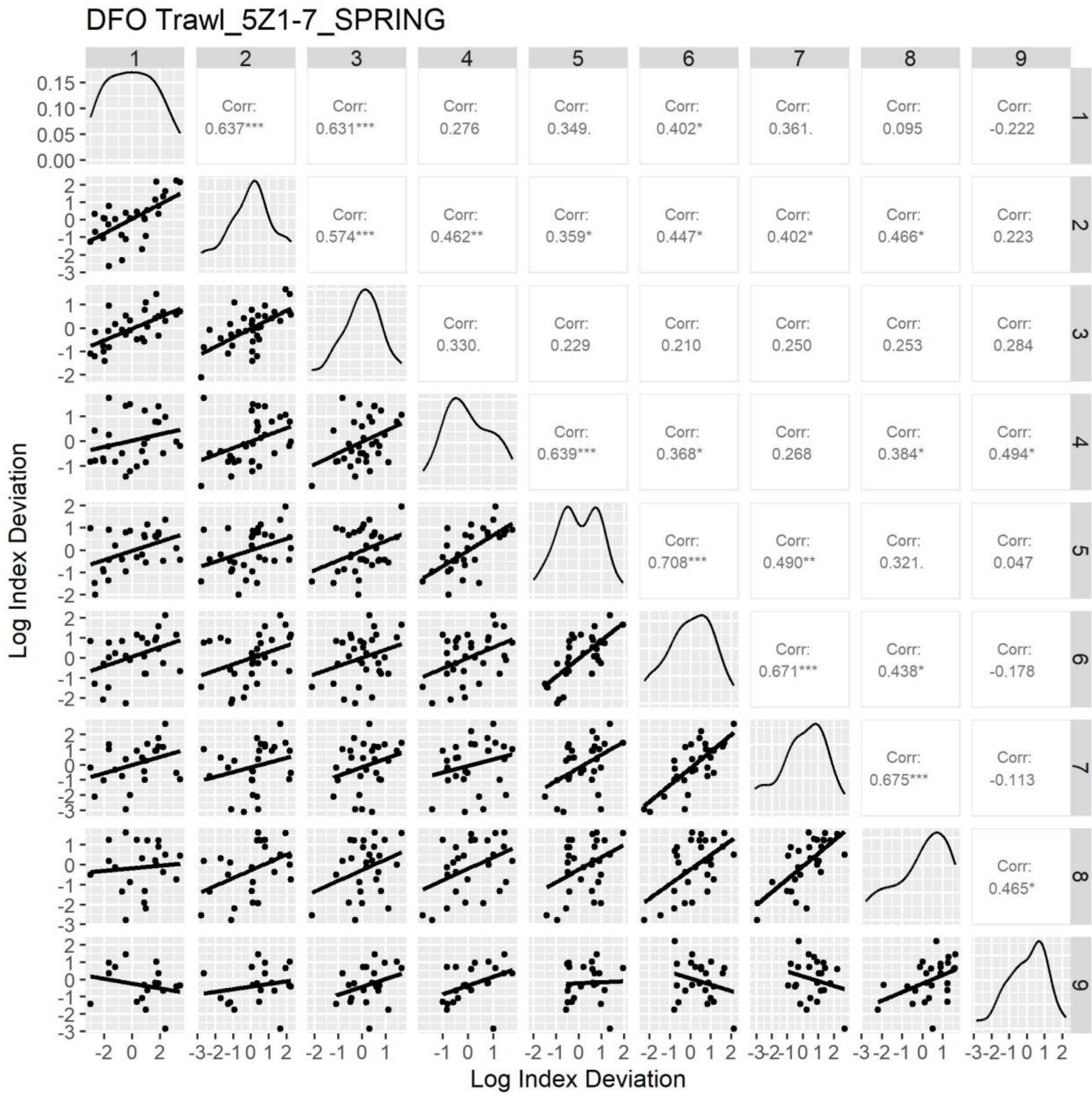

Figure 13: See Figure 7 for details.

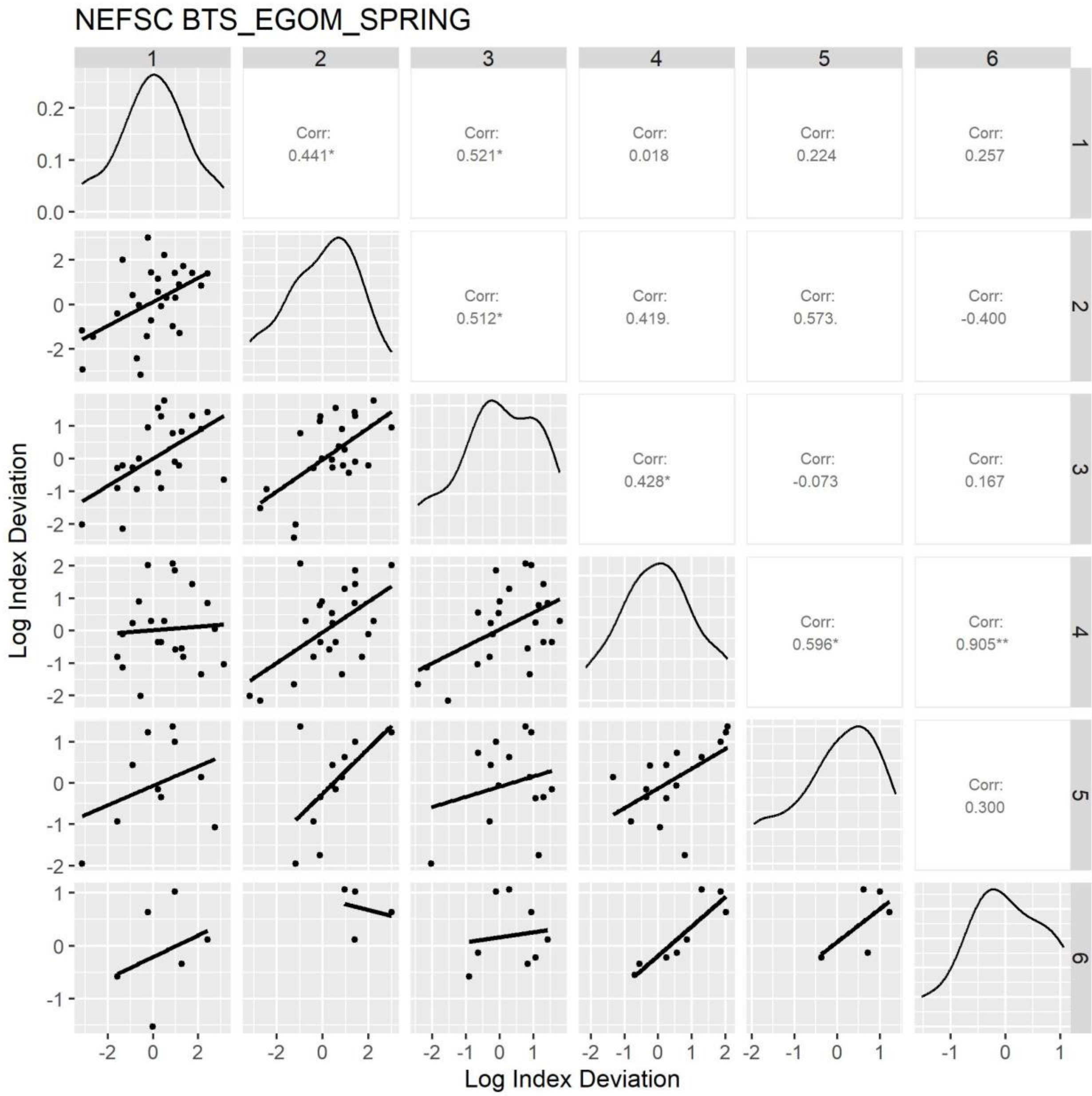

Figure 14: See Figure 7 for details.

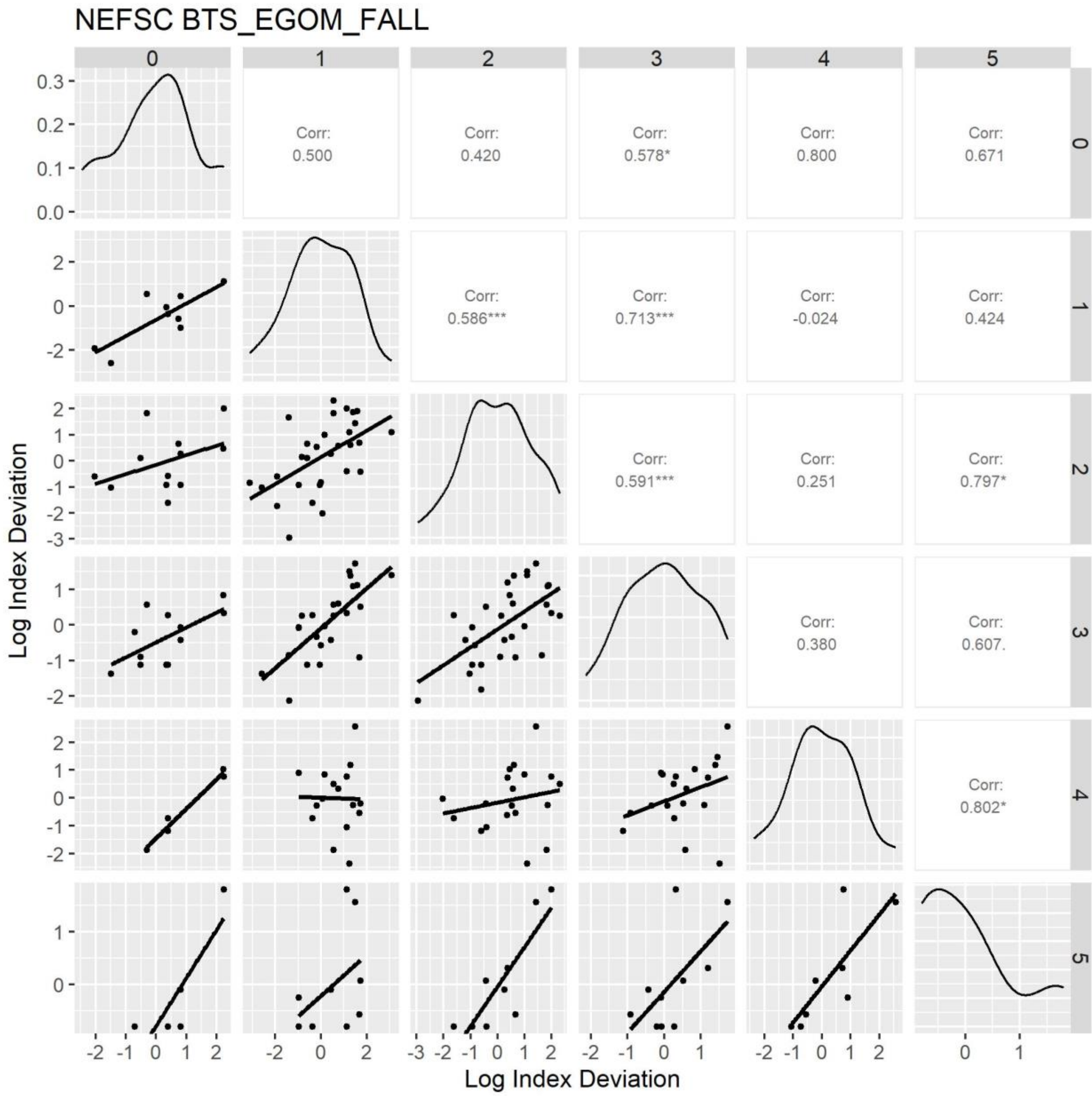

Figure 15: See Figure 7 for details.

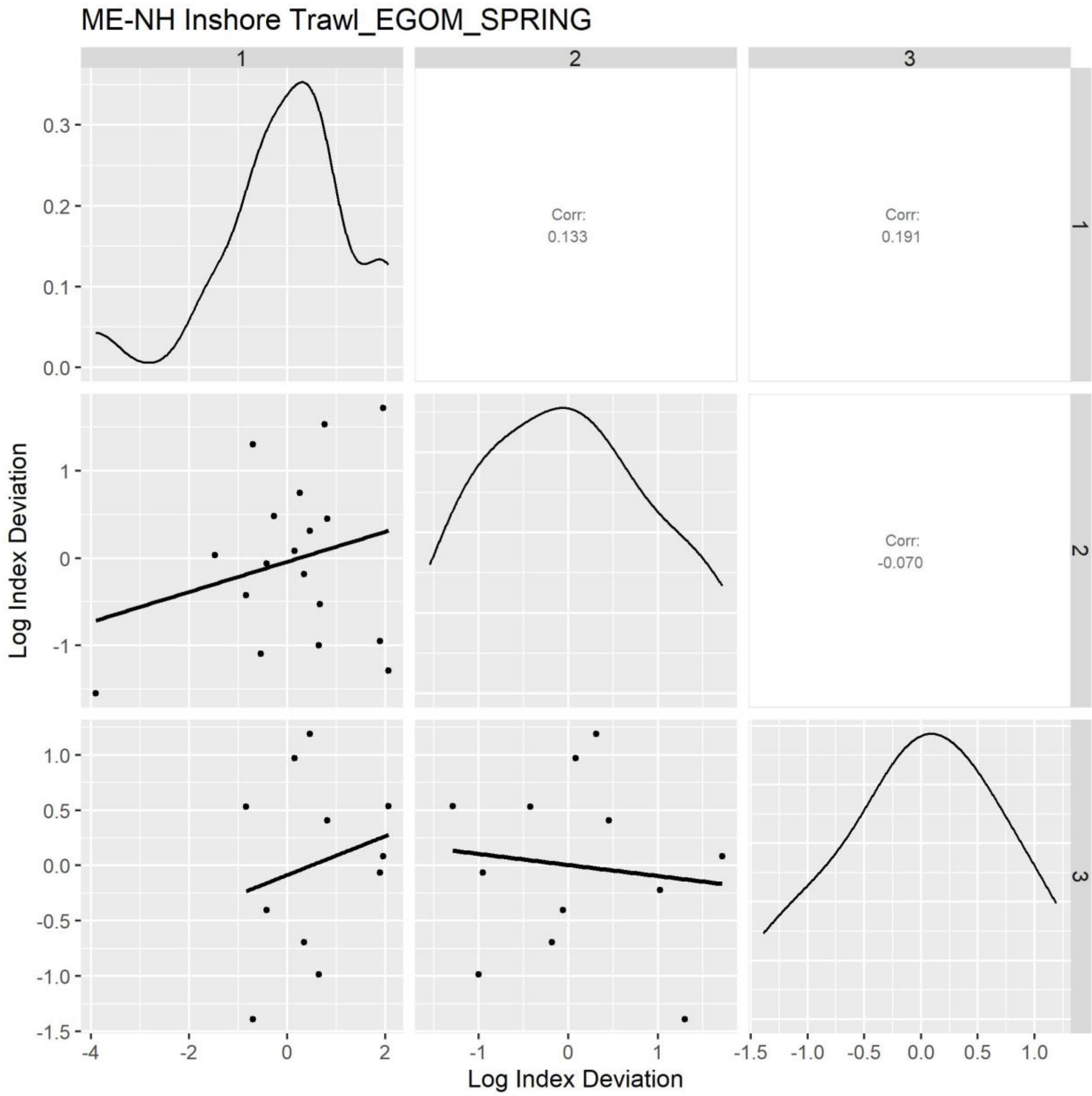

Figure 16: See Figure 7 for details.

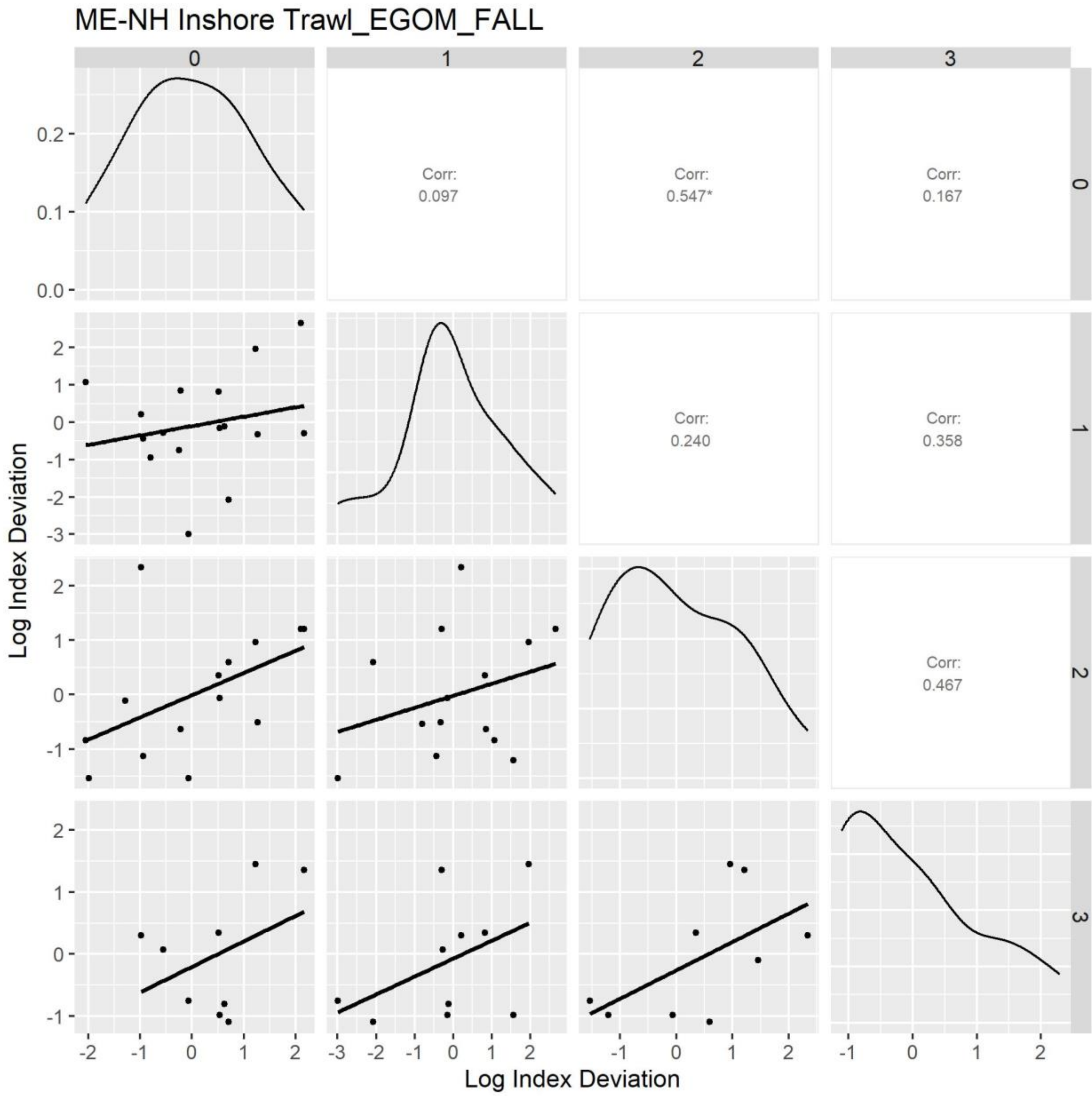

Figure 17: See Figure 7 for details.

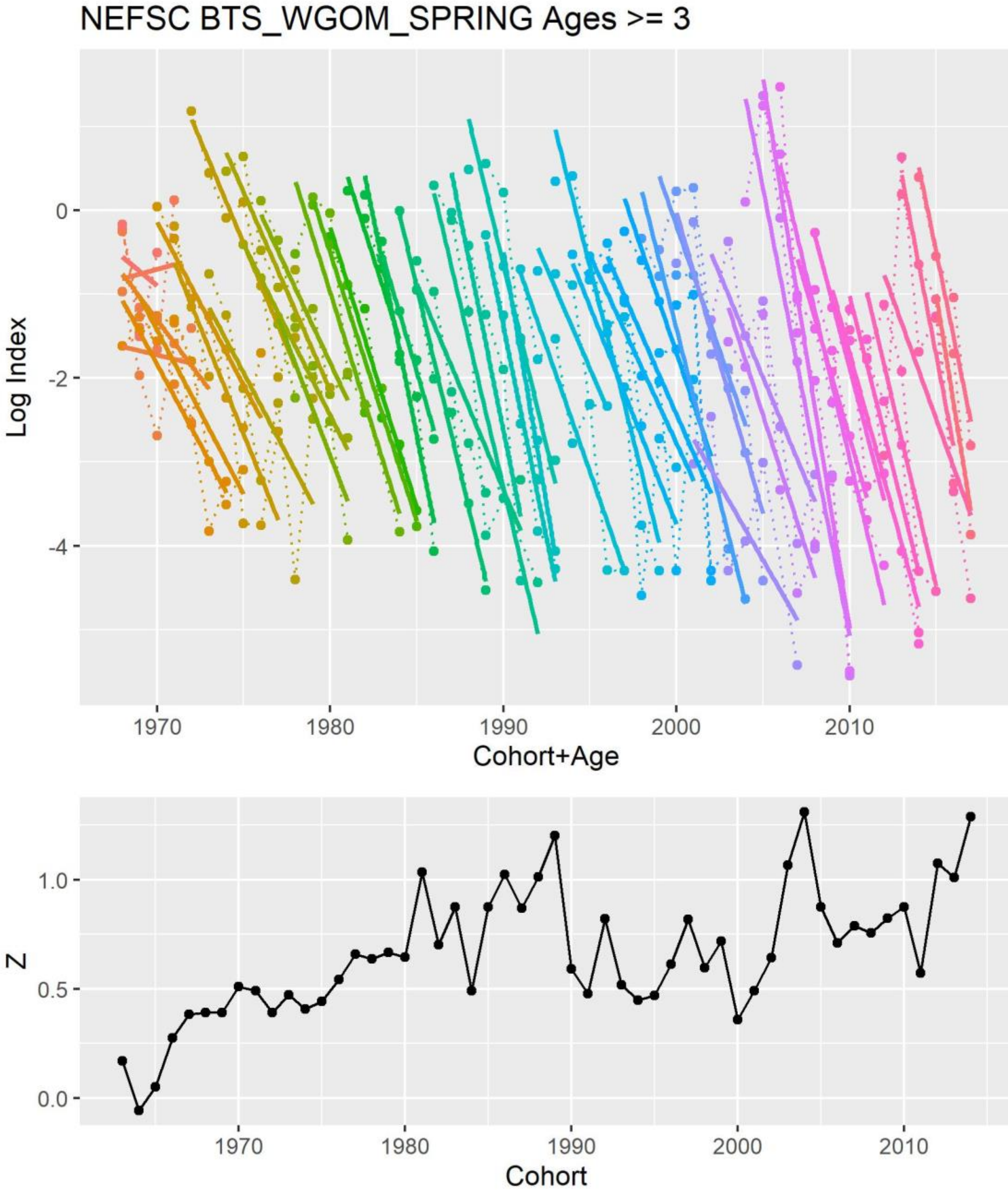

Figure 18: Top panel: The points and dotted lines indicate log survey index versus age for each cohort (colors). The solid lines are linear regression fits for each cohort. The regression slopes (i.e., Zs) are shown in the bottom panel. The survey and ages are indicated at the top.

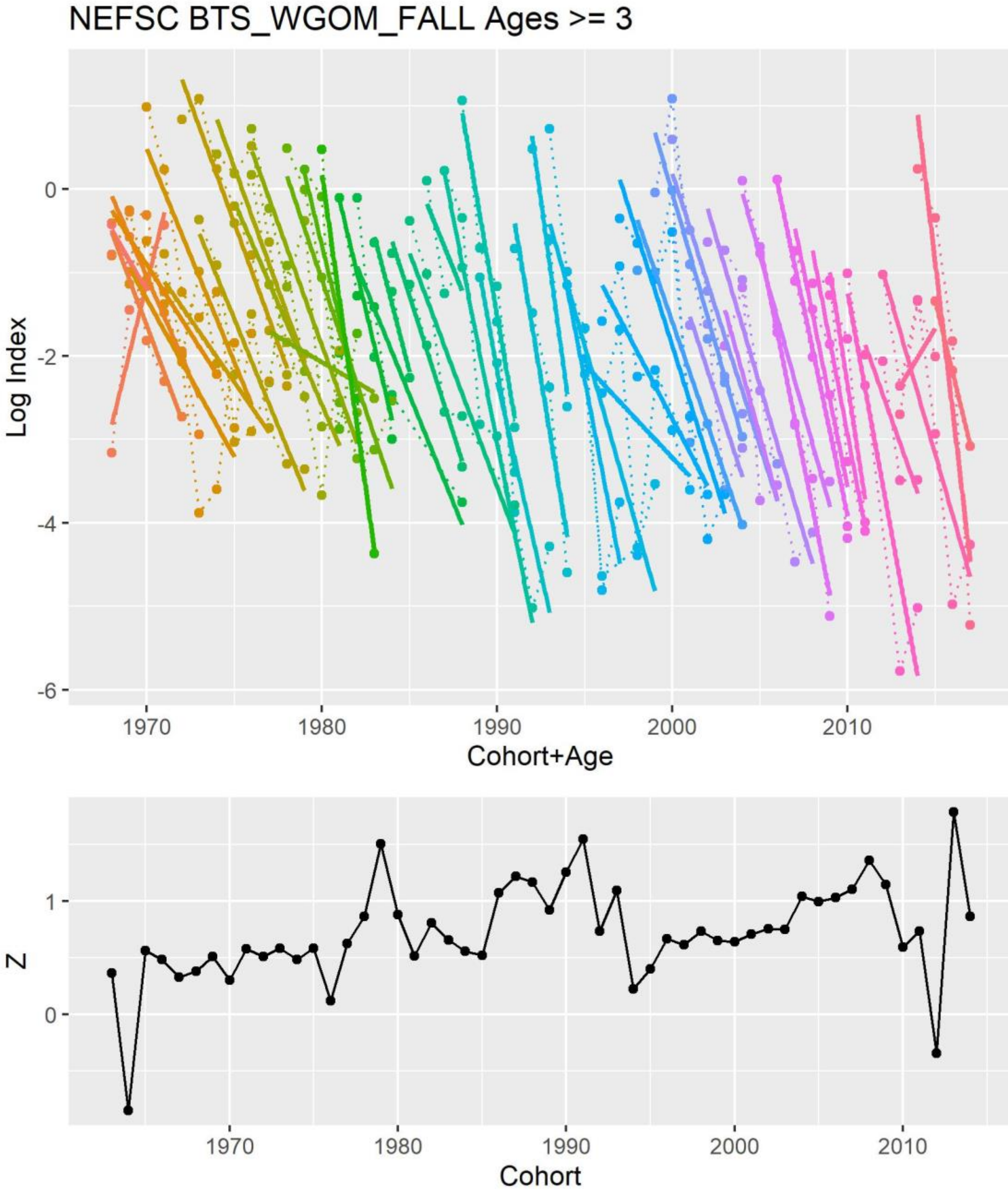

Figure 19: See Figure 18 for details.

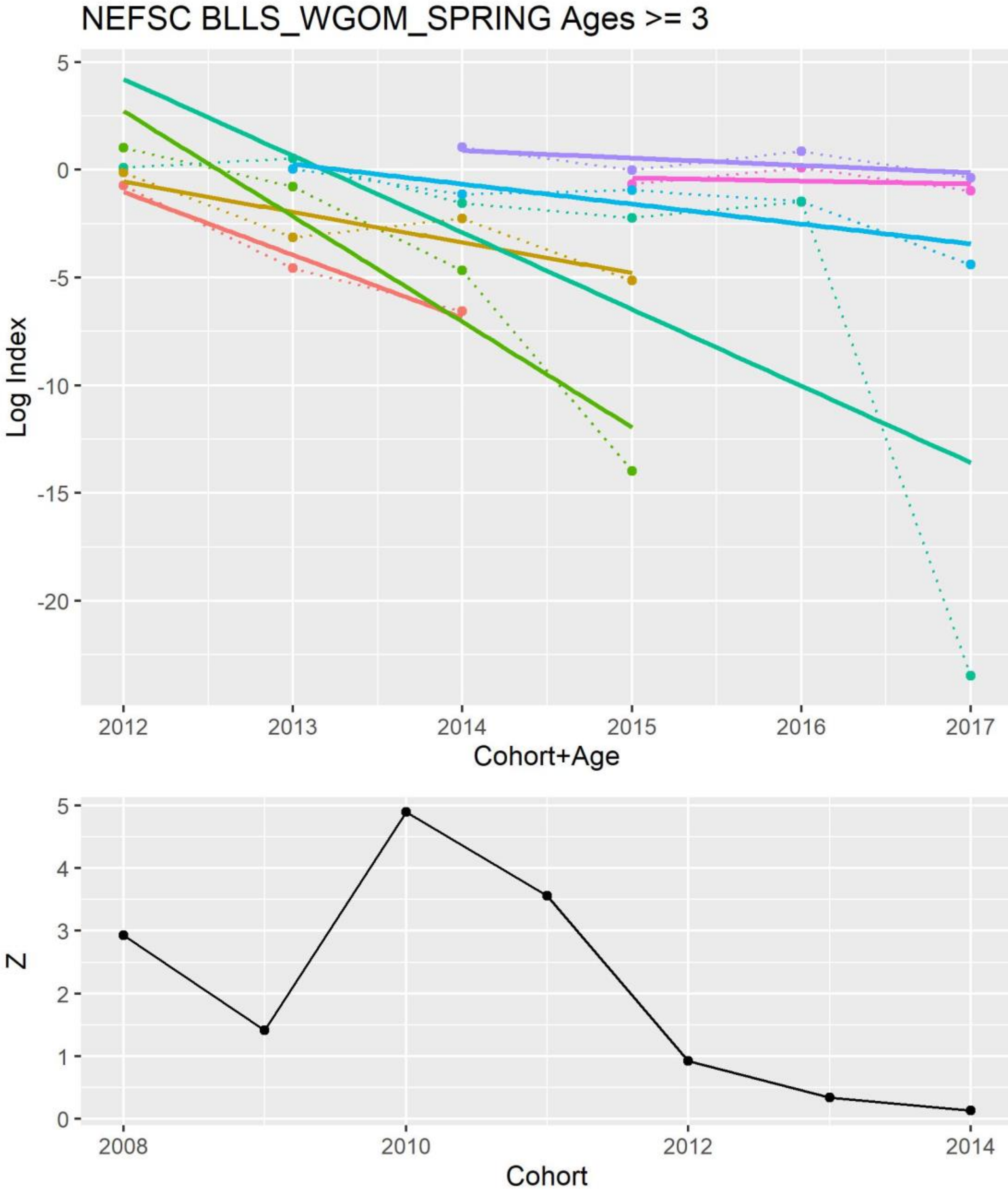

Figure 20: See Figure 18 for details.

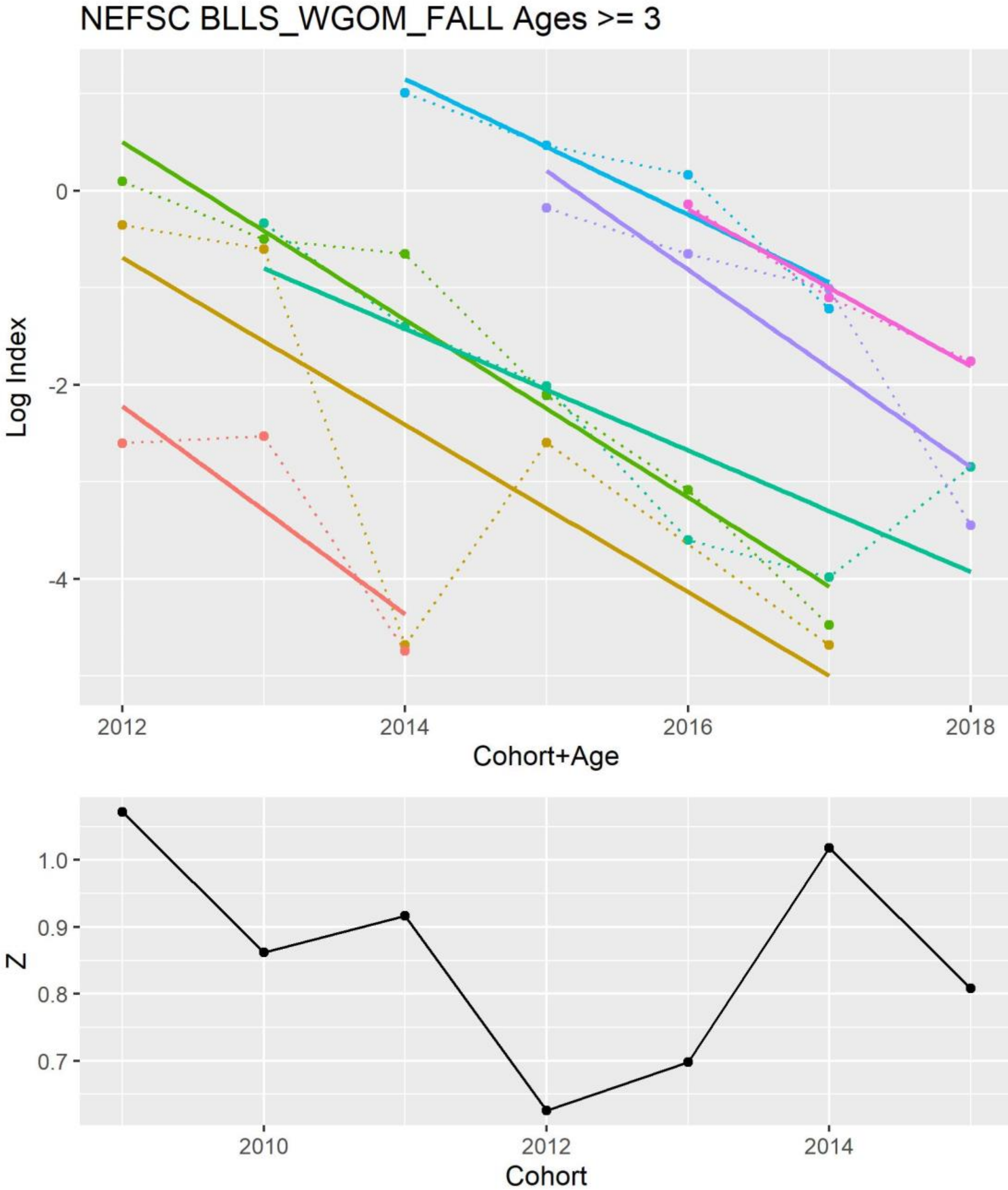

Figure 21: See Figure 18 for details.

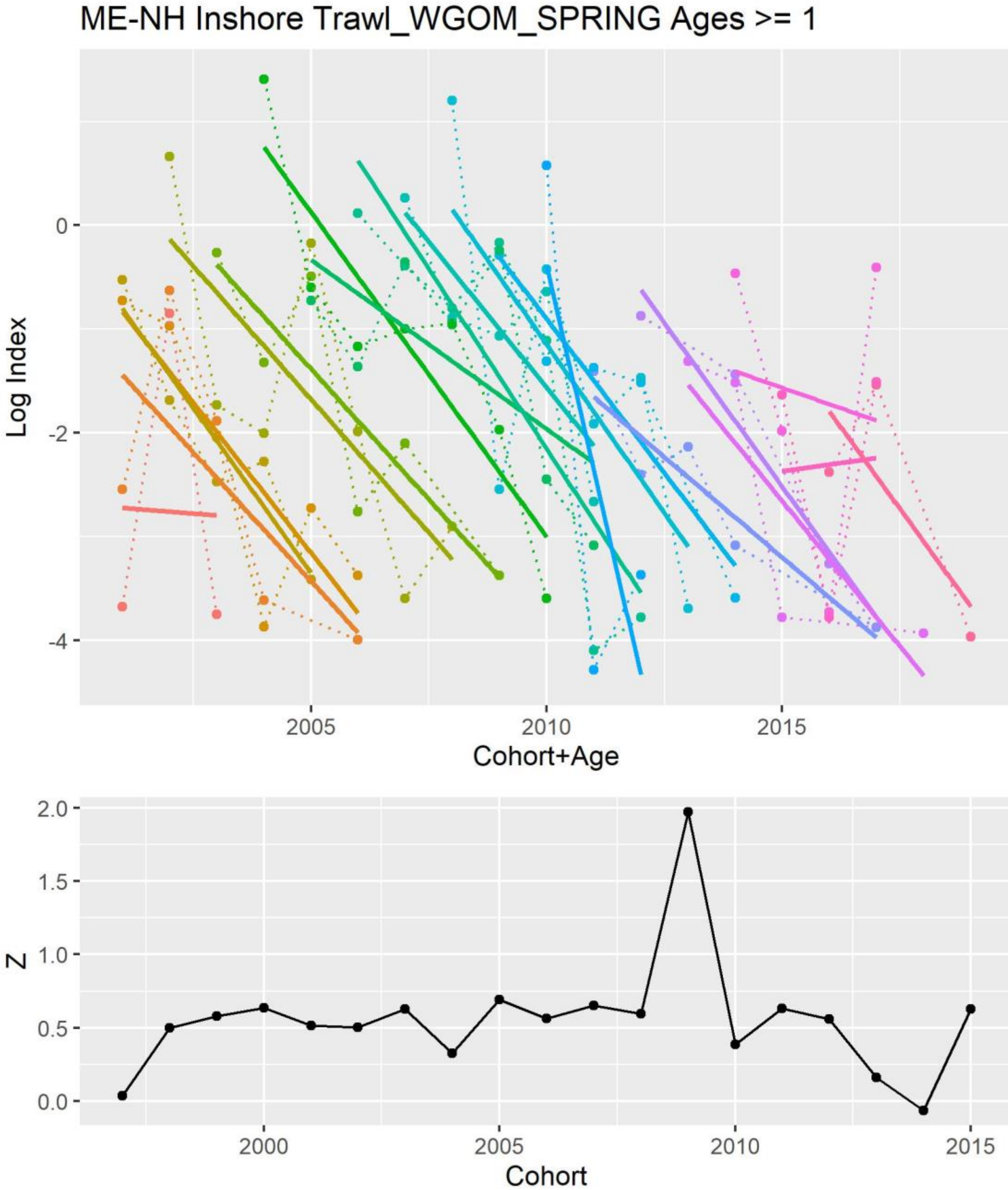

Figure 22. See Figure 18 for details.

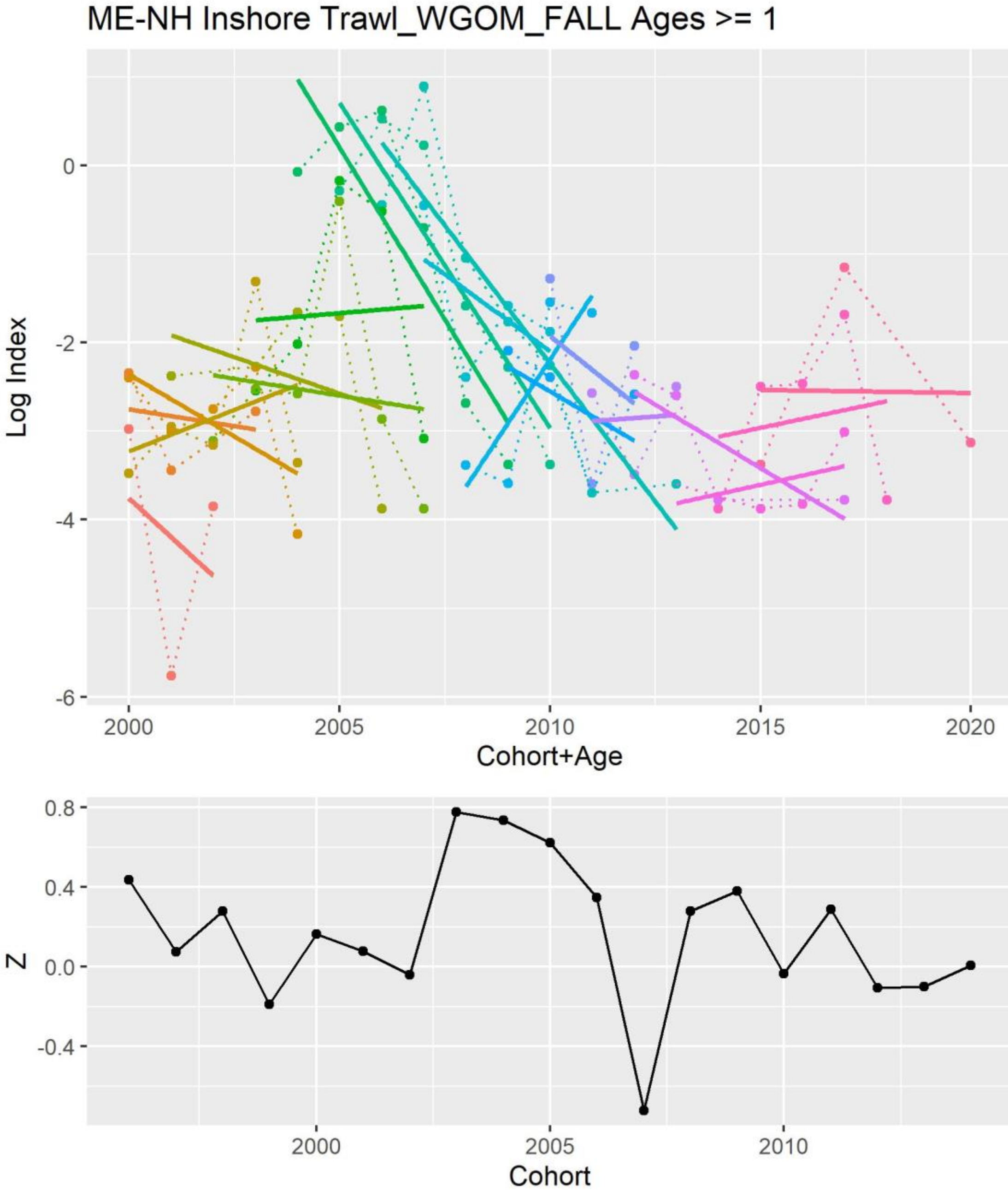

Figure 23: See Figure 18 for details.

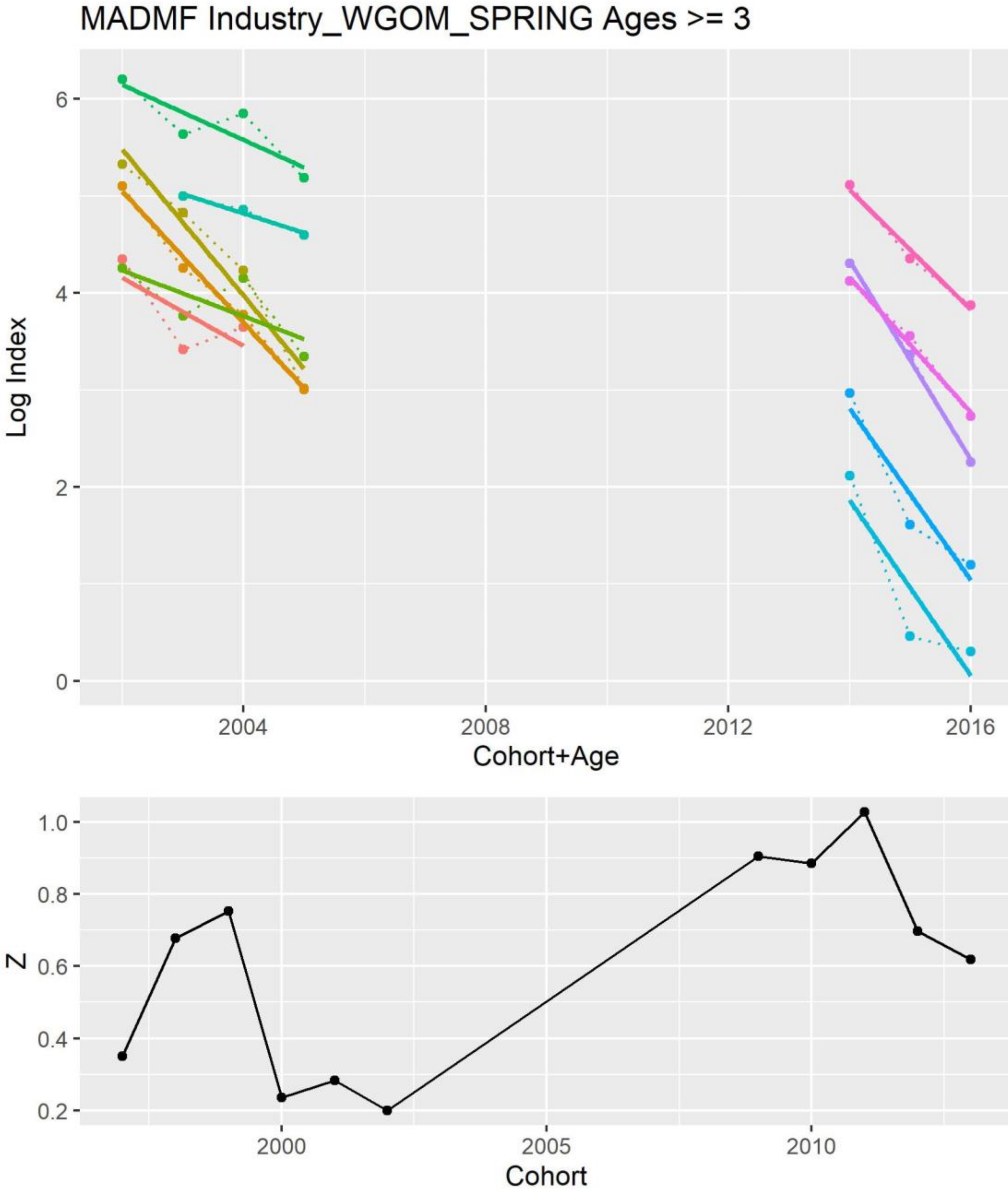

Figure 24: See Figure 18 for details.

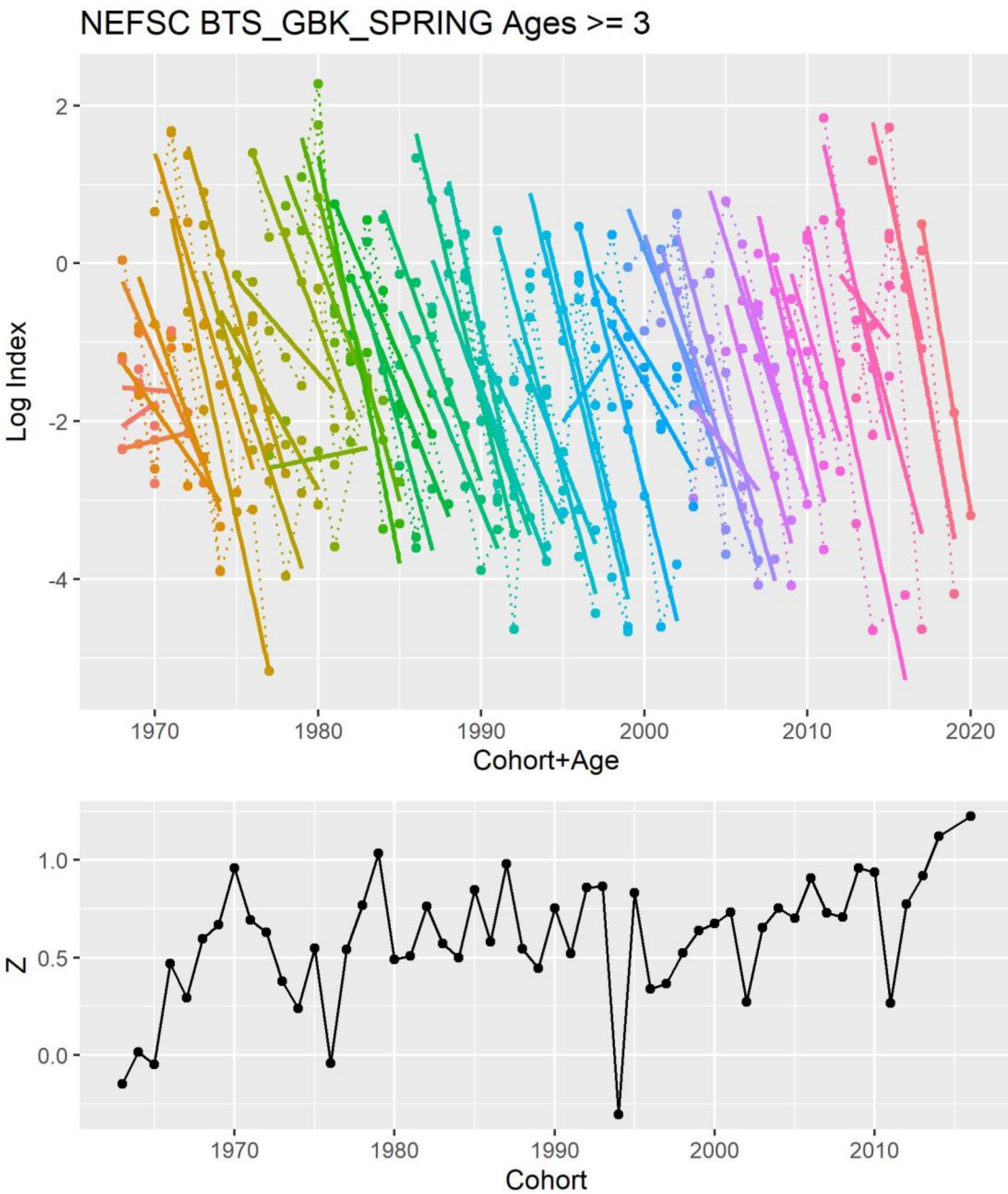

Figure 25: See Figure 18 for details.

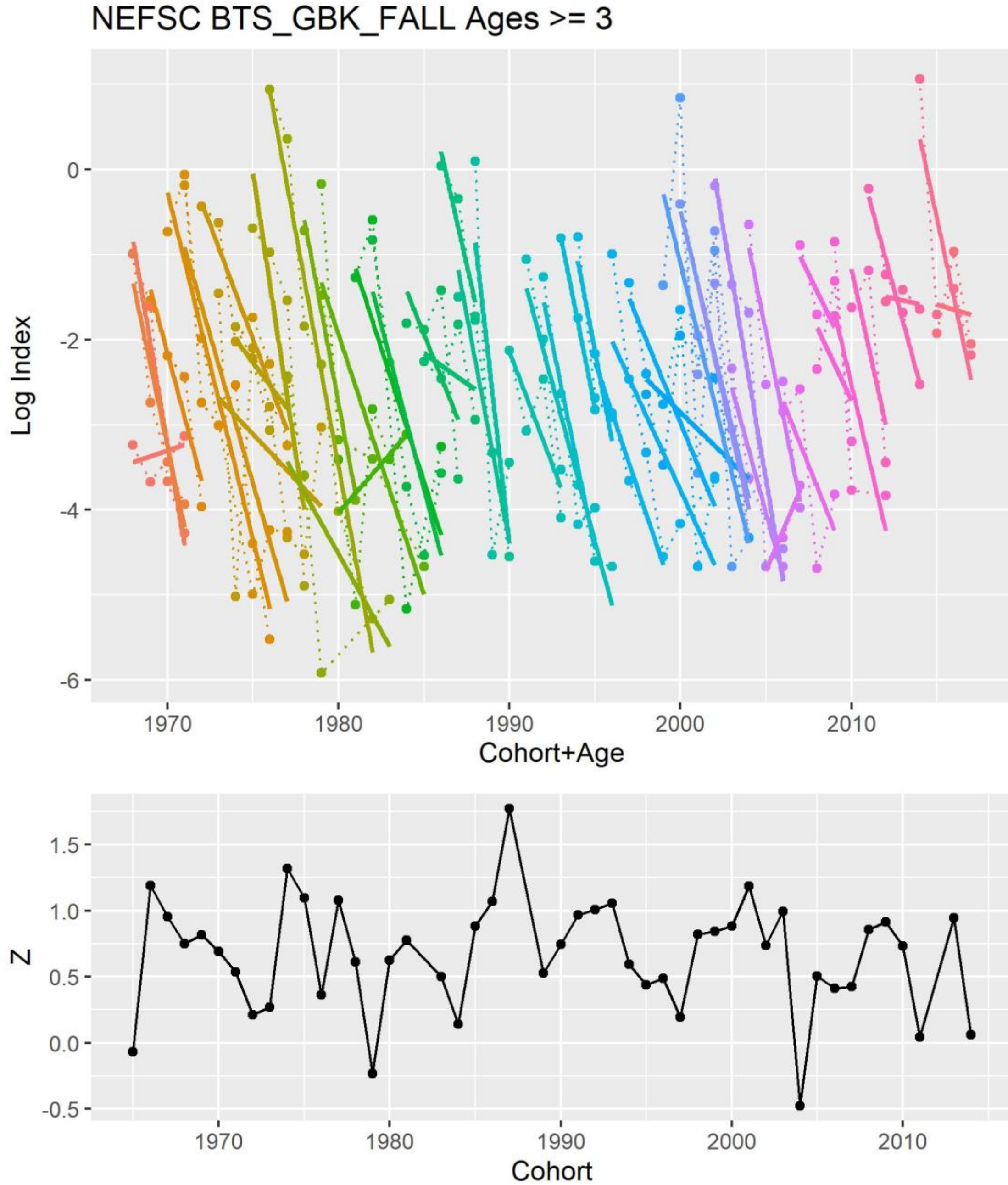

Figure 26: See Figure 18 for details.

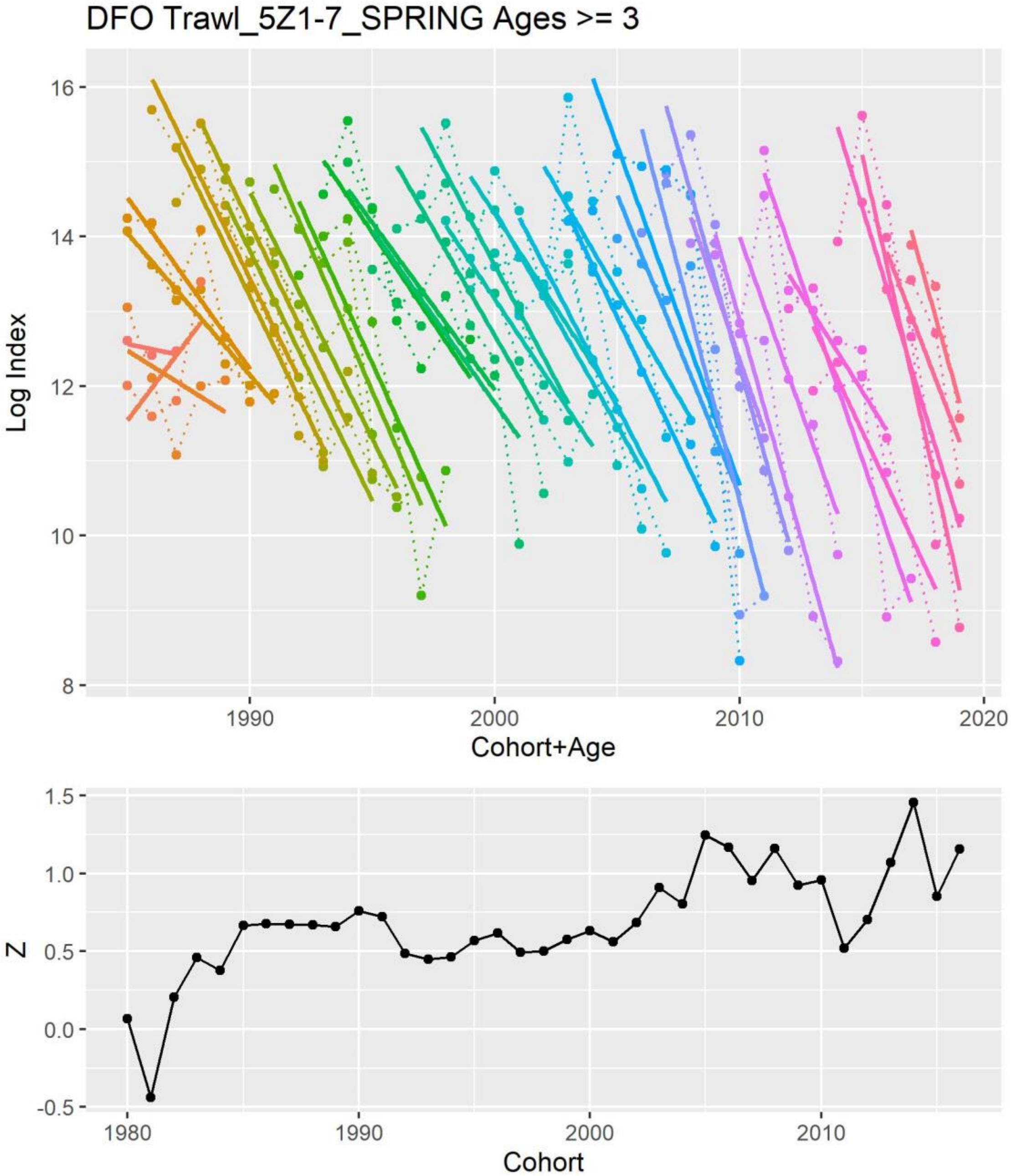

Figure 27: See Figure 18 for details.

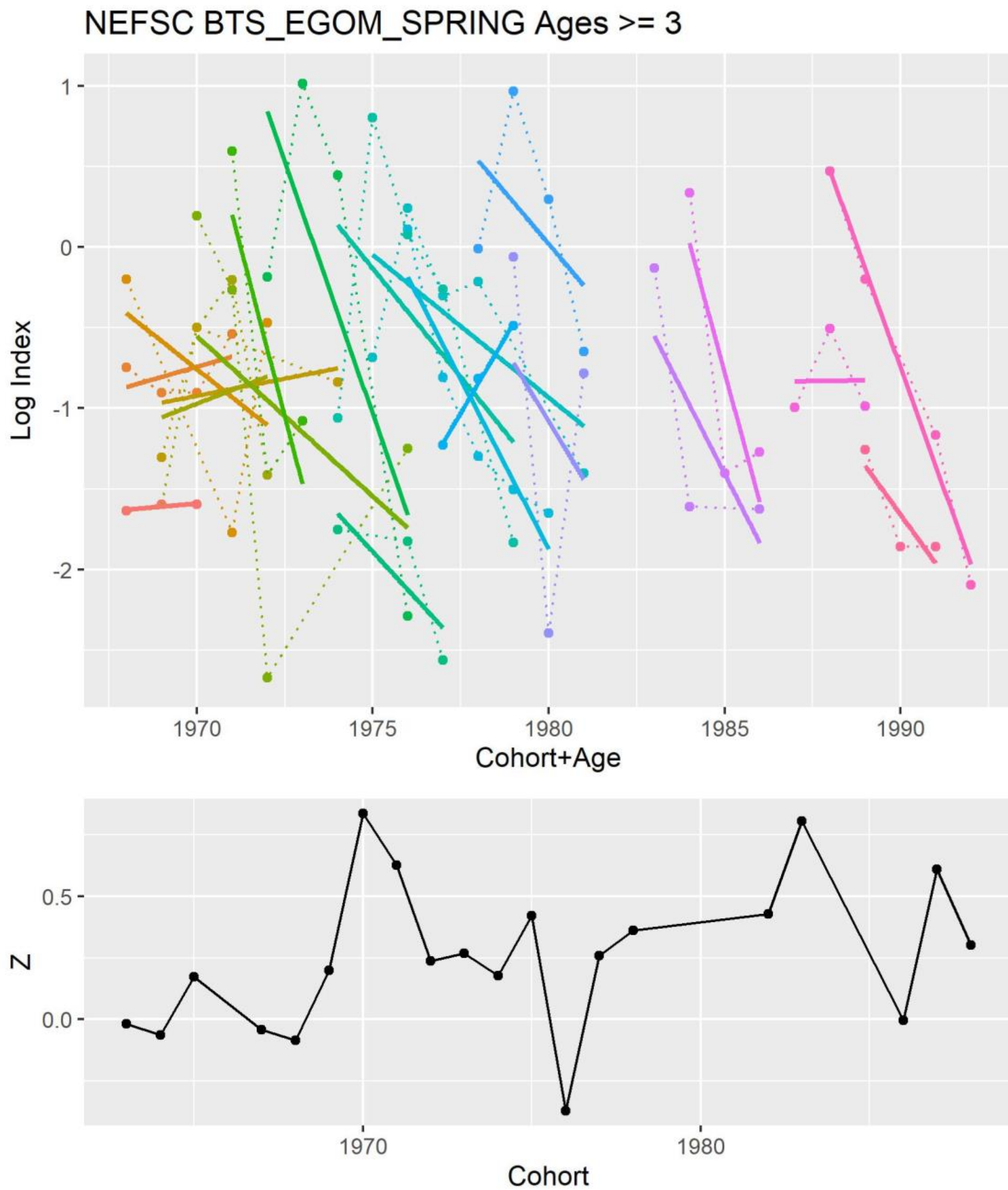

Figure 28: See Figure 18 for details.

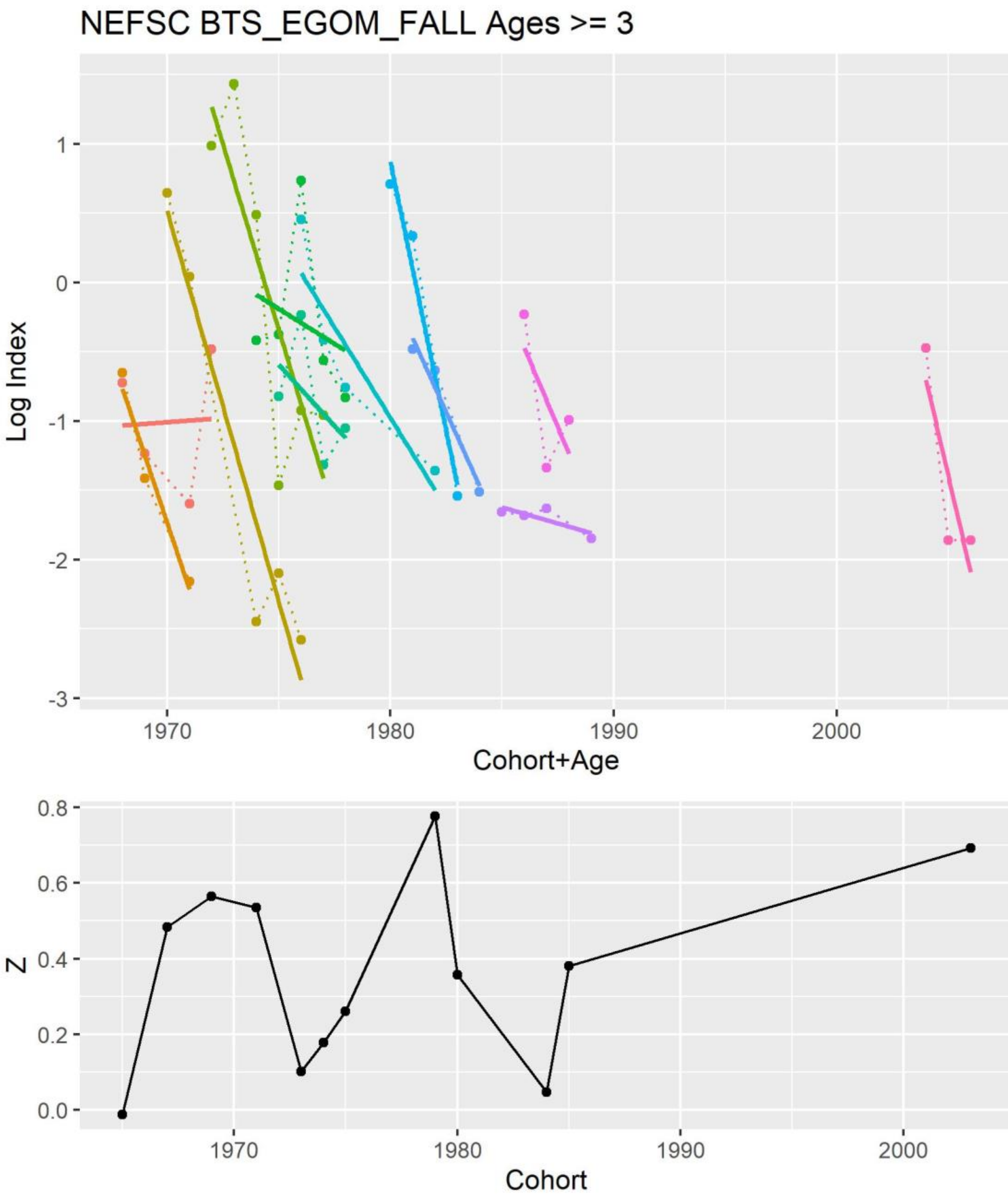

Figure 29: See Figure 18 for details.

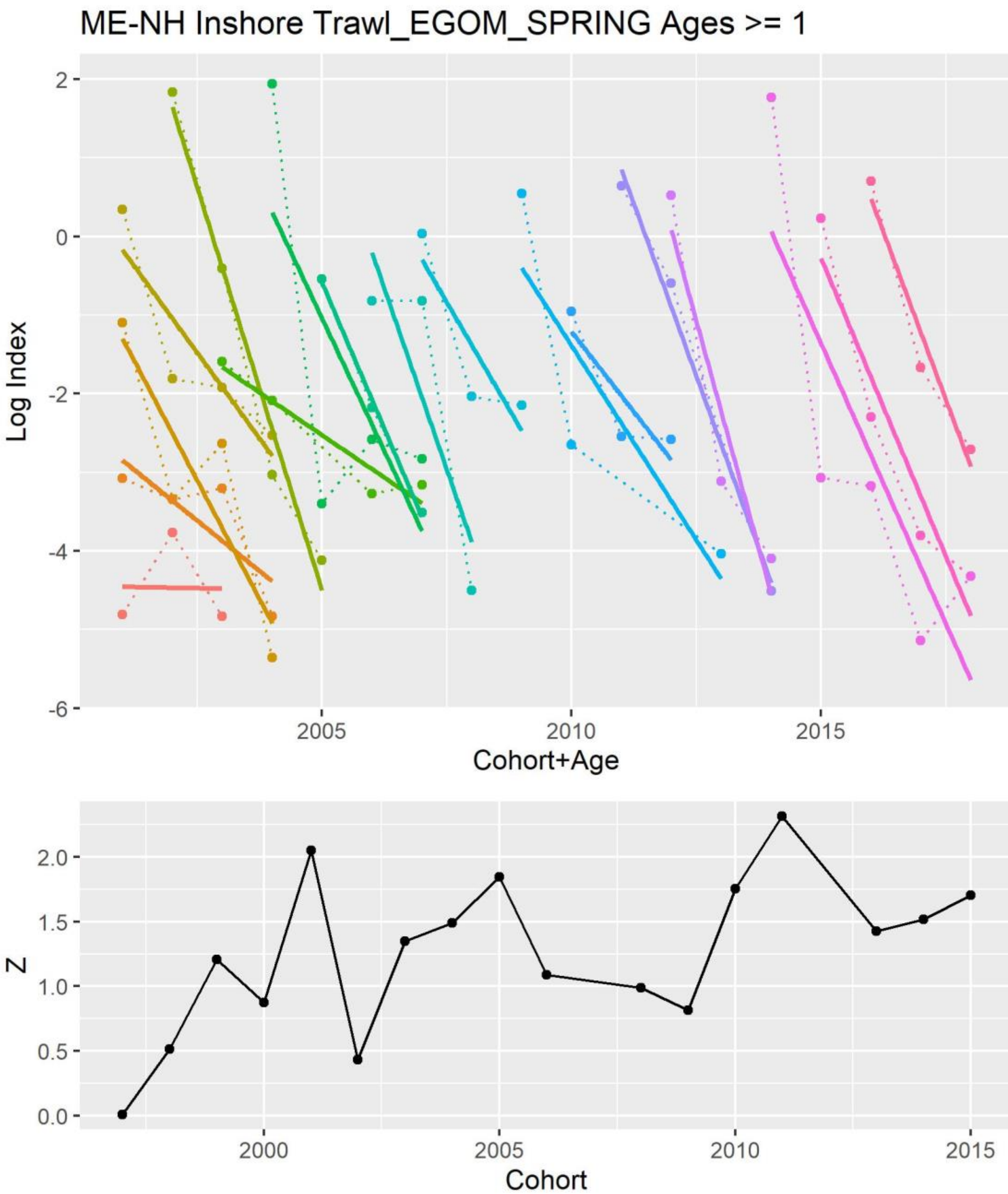

Figure 30: See Figure 18 for details.

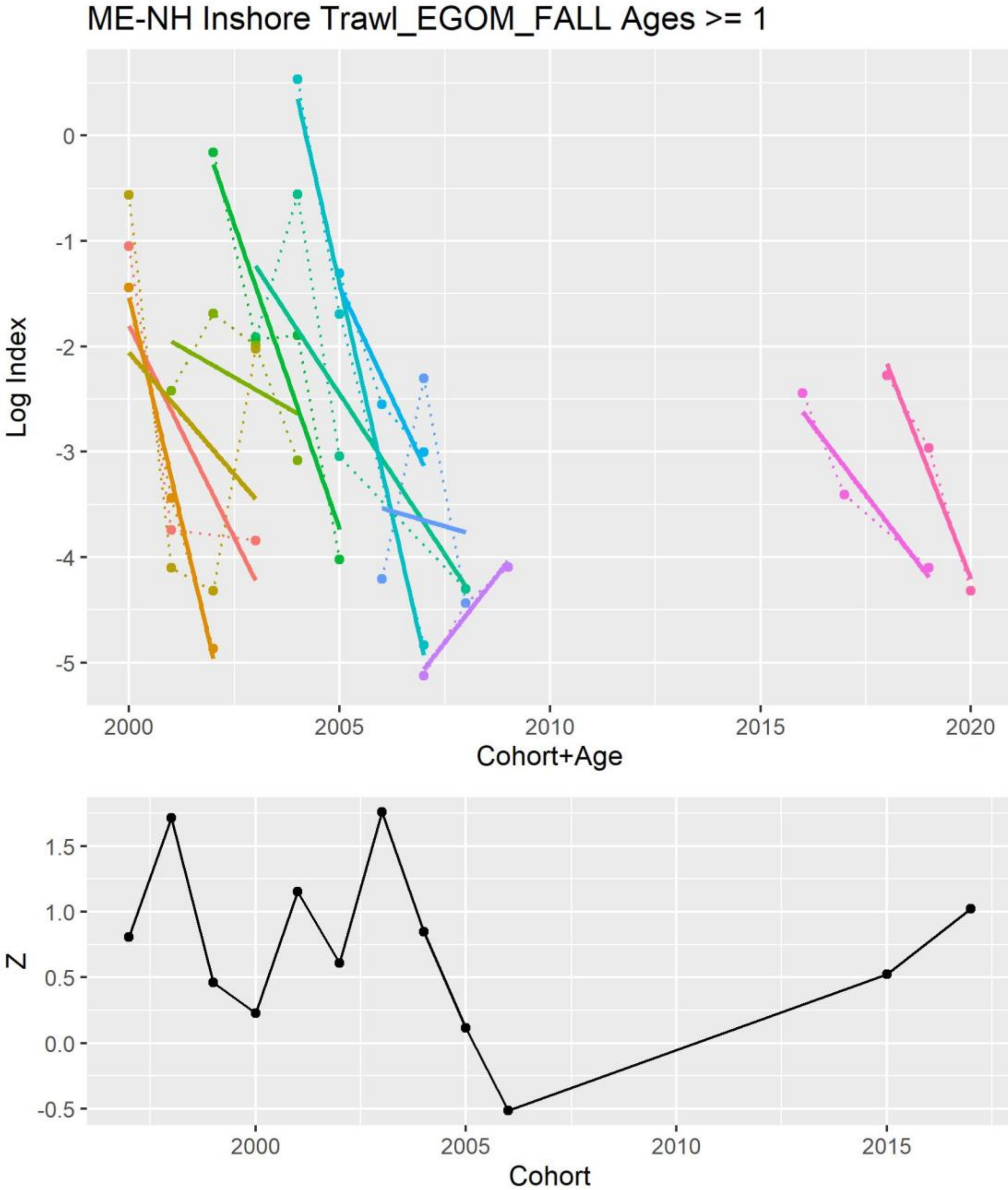

Figure 31: See Figure 18 for details.

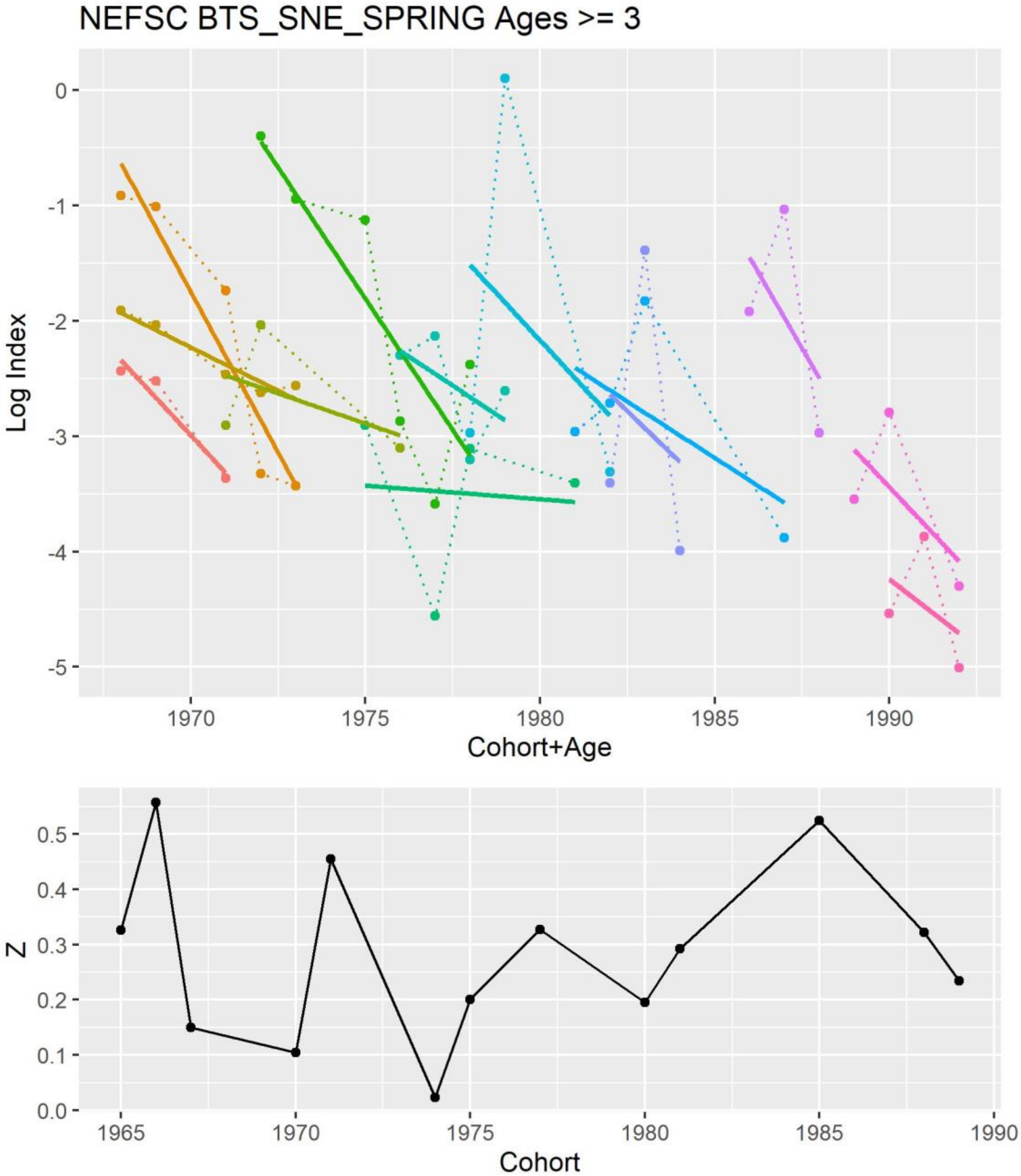

Figure 32: See Figure 18 for details.