EASTERN ATLANTIC BLUEFIN TUNA ASSESSMENT REVIEW REPORT FROM THE DATA PREPARATION MEETING HELD IN APRIL

James Ianelli[1]

SUMMARY

This report is based on the data-workshop meeting held in April and addresses the requested (from the Terms of Reference) broad evaluations on: a) indices and index development methods and evaluate the appropriateness of statistical methods used to develop indices considering how the raw data was/is collected, b) the adequacy of the biological assumptions (especially natural mortality, growth, fecundity), c) key data inputs and their assumptions such as input landings and size composition of fish destined to farms. Additionally, I was requested to provide a list of key issues, if any exist, that could result in the models being rejected, if not adequately addressed. Broadly evaluate the adequacy, appropriateness, and application of assessment methods and provide broad evaluation on CPUE treatment in models (variance scaling, selectivity). I conclude with some comments and recommendations for assessment models, model structure, suggestions for parameterizations and sensitivities.

RÉSUMÉ

Ce rapport se base sur l'atelier sur les données tenu en avril et traite des évaluations générales demandées (par les Termes de référence) en ce qui concerne : les indices et les méthodes de développement des indices et évaluer la pertinence des méthodes statistiques utilisées pour développer des indices compte tenu de la façon dont les données brutes sont/ont été collectées, l'adéquation des postulats biologiques (notamment la mortalité naturelle, la croissance et la fécondité), et les données d'entrée clés et leurs postulats, tels que les données d'entrée de débarquements et de composition par taille des poissons destinés aux fermes. En outre, on m'a demandé de fournir une liste des principaux problèmes, le cas échéant, qui pourraient conduire au rejet des modèles s'ils ne sont pas correctement traités. Évaluer globalement l'adéquation, la pertinence et l'application des méthodes d'évaluation et fournir une évaluation générale du traitement des CPUE dans les modèles (mise à l'échelle de la variance, sélectivité). Je conclus par des commentaires et des recommandations pour les modèles d'évaluation, la structure des modèles, et des suggestions pour le paramétrage et les analyses de sensibilité.

RESUMEN

Este informe se basa en la reunión del taller de datos celebrada en abril y aborda las amplias evaluaciones solicitadas (en los Términos de referencia) sobre índices y métodos de desarrollo de índices, y evalúa la idoneidad de los métodos estadísticos utilizados para desarrollar índices teniendo en cuenta cómo se recopilaron/recopilan los datos brutos, la idoneidad de los supuestos biológicos (especialmente la mortalidad natural, el crecimiento, la fecundidad), y las entradas de datos clave y sus supuestos, como la introducción de datos de desembarque y la composición por peces de los peces destinados a las granjas. Además, se pidió que se proporcionara una lista de problemas clave, si es que los hay, que podrían dar lugar a que se rechacen los modelos, si no se abordan adecuadamente. Evaluar ampliamente la idoneidad, la conveniencia y la aplicación de los métodos de evaluación y proporcionar una amplia evaluación sobre el tratamiento de la CPUE en los modelos (escala de varianza, selectividad). Se concluyó con algunos comentarios y recomendaciones para los modelos de evaluación, la estructura del modelo, sugerencias de parametrizaciones y sensibilidades.

KEYWORDS

Stock assessment

[1] Jim.Ianelli@gmail.com, 10628 Marine View Drive SW, Seattle WA 98146

1 MSE process

The process for the E-BFT MSE where data used in the assessment matches as close as possible to that used in the MSE OM seemed appropriate (and vice-versa, apparently). I was unable to completely review how the OM conditioning had been achieved and noted that the terminal year was 2020 which would differ in coverage to the assessment data. The document providing extensive details on the OM conditioning was dated 2015 and it was unclear if some advances (e.g., in MCMC integration) had been implemented. Some other general comments are:

- the system of getting feedback for performance measures (presumably through "Ambassador meetings" and in Panel 2) seemed difficult or at least hard to assess the effectiveness of communications (also the extent non-technical people contribute to questions and discussions at technical meetings may slow the process)
- the "quilt table" as presented seemed biased towards catch metrics (though it seemed this topic is being addressed at different levels within Commission/MSE process);
- the candidate management procedures (CMPs) levels of tuning seemed incomplete (but this will likely be updated)
- adhering to some members "legal" definition of overfishing in some member jurisdictions seems to potentially constrain, complicate and confuse the goal of an MSE
- the developments and facility presented in the ABFTMSE package seemed exceptionally useful as a tool for all participants.

The group heard a detailed description on issues related to deriving F_{MSY} levels or proxies thereof. A new treatment (U_{MSY}) was presented and included in specifications document. This resolved some complications related to how the myriad fisheries appear to have highly variable selectivity patterns over time. Nonetheless, it seems that even this formulation requires some assumptions about stationarity or lack of trend in the relative vulnerability of different age groups over time.

Relative to tuning and CMP evaluations, I found the term "satisficing" to be a reasonable, formalized way to reduce the number of CMPs and includes a broad evaluation of performance statistics. The reference set of OM conditions appears to provide sufficient contrast across CMPs and may help selection and reduce the need for many robustness tests.

The workload remaining for completing the MSE seems very heavy. My main comment relative to this is a general worry that having a full "assessment" process in the midst is slowing development. Nonetheless, document PLE_123 /2021 quite strongly encourages that the assessment activity does not negatively impact on MSE and other SCRS prioritized work. The outline for next steps and timings seems doable, and of critical importance is the agreement needed on how "exceptional circumstances" will be determined and agreed approaches to reacting to them should they occur.

2 E-BFT Data preparation

The Commission had requested that the assessment data for an EABFT stock assessment be compatible with those being used for the developing MSE. They note that these are meant to be as strict-updates only. In one report, it was noted that "…Exchanges with various national scientists had not indicated any need to re-evaluate indices." I can only trust that this was a reasonable conclusion. However, there seemed to be some contradiction with respect to indices used in the previous assessment compared to what is used in the MSE (which for the present assessment cycle is meant to be a "strict update" from the MSE). Some clarity in communicating how indices in the last assessment differed from that used in the MSE is warranted.

2.1 Development of fishery-independent indices

An index using acoustic methods in the Bay of Biscayne was presented. I support the group's general recommendation that the survey should be extended in time but was unable to assess if the level of financial support needed was appropriate relative to other research programs used for the assessment analysis.

Related to the updated French aerial survey index, the group recognized that some environmental factors could be affecting the "availability" of BFT to this survey. Work is underway to evaluate this and should be available in 2023. They adopted the use of this index in 2017 and 2020 assessments and agreed to retain it for this year.

Regarding the larval index, this shows a major increase in recent years and is responsible for some, rather incredible, increased stock sizes in projections. It is rare that such age-zero or age 1 indices work well in predicting

recruitment to fisheries at much older age classes. I agree with the group's recommendation to inflate the recent years' CVs as they are relatively small (and among the smallest in the series).

2.2 Biological assumptions

The paper describing some alternative back-calculated growth by Stewart et al. (SCRS/P/2022/011) was of particular interest to me. These fish were from northwestern Atlantic and were apparently likely growing to bigger sizes than those from the eastern stock. Nonetheless, I think that further investigation of these data relative to the E-BFT assessment is warranted, at least in the form of comparisons with data used in the assessment. Perhaps it could help inform growth during juvenile to early adulthood periods.

Another paper examining maturity-at-age was presented (SCRS/P/2022/012). This study found that samples both stocks had similar maturation. but no comparison with actual maturity that's used in the assessment (despite the assertion that such a "comparison" is presented in Table 4 of the April meeting report, both the growth and the maturity work of Stewart et al. are missing). *I recommend that the different data sets about growth be plotted and compared.* Another suggestion would be to evaluate how residuals from a global "mean" growth curve compare spatially, if such data are available. It seems such information could be drawn from the database work presented by Rodriguez-Marin (SCRS/2022/075), even if a relatively coarse spatial grid.

Length at age data to estimate growth seem to be limited for this stock in general, particularly within the Mediterranean. This poses a problem for several assessment approaches including what type of values (and/or priors) to use for natural mortality, mean somatic mass-at-age, and effective fecundity. The application of conditional-age-at-length (CAAL) data provided by Rodriguez et al. has been adopted within SS3; and in principle, should help with growth estimation (noting that the SS3 model presented was configured to estimate a single growth curve within the model). However, they use two data sources (spines and otoliths) and the extent they are consistent is unclear. For example, apparently the spines perform poorly for age-determinations beyond age of 13 yrs. There may also be issues with how the different CAAL data are weighted within the assessment. I note that for the SS3 runs, they will account for age-determination error so some of the differences can perhaps be addressed. Additionally, it was noted that by fixing the growth in the final assessment model (after comparing/conditioning with the CAAL and the work presented in Stewart et al. (2022) might add robustness in future model evaluations (including adding flexibility in time-varying selectivity).

2.3 Key data inputs and assumptions

2.3.1 General

During the meetings I was able to obtain important data files used in different assessment software packages. This provided the ability to understand the extent and issues related to how scattered the data appear through time by fishery. I was able to run the draft versions of stock synthesis (SS3) to understand issues related to that approach. I also followed closely the development of an ASAP application and comment on those along with the VPA approach in the "Key issues" section below.

In the assessment model specification report it was detailed that the assessment data should only extend to 2020. This seems unnecessary for several approaches (except VPA perhaps?). More recent data is available and was presented. For example, a key index (the French aerial survey) showed a slight decrease in 2021 compared to 2020.

Such information seems important to resolve near-term trends.

2.3.2 Catch

The landings data for 1996-2006 period (historically dubbed as inflated catch) seemed problematic but this was well recognized by the experts and efforts to account for additional mortality appear to have been addressed. Tables 6 and 7 of the meeting report highlight data availability showing the holes by fleet and major area.

2.3.3 Size composition

In an effort to evaluate this, I obtained the file "E-BFT_szfrq5_Apr2022.csv" from the working drive and simply plotted out the length compositions by gear type (which covered 5 types of gear and totaled some 16 'fisheries'; Attachment 2). These figures quite simply illustrate the complexity of the fisheries. Specific to the stereo camera work, the application seems important and is encouraged. However, from the materials presented at the data preparation workshop, it was difficult to ascertain the consistency of the actual sampling as it appears that some programs had ceased recordings. Also, Figure 3 of the meeting report (reproduced from a presentation) has an odd truncation of data at 100 cm SFL. It would be useful to understand that pattern.

2.3.4 Age data

The Stewart et al. paper presented some innovative studies on age determination. I understand that these samples were from the western north Atlantic, but it seems showing some comparisons with model assumptions (i.e., that used in the AgeIt software and results from SS3) would be useful. I examined the original document on the AgeIt software (Ortiz 2011) and wonder if it may be useful to run this to generate new age composition data (for ASAP and VPA in particular) at some alternative plausible growth curves to evaluate the sensitivities. See the section above for a discussion of issues related to applying the CAAL data within the SS3 model.

2.3.5 CPUE treatment in models

The paper presented on the Japanese longline CPUE was comprehensive and presented a new approach that applied a spatio-temporal delta-lognormal model to their data (VAST; Thorson 2017). They also provided results updating the past methods for contrast. I agree with their recommendation to adopt the VAST method. The diagnostics applied used state-of-the-art methods and appeared to give reasonable fits. On a minor note is that I also think they should abandon using SAS as that limits the transparency and transferability among members.

As far as implementation within assessment models, Table 9 of meeting report shows the index values and the CVs by year. I find these CVs to be quite high in general, even by most fisheries standards. This highlights the importance of seeing the sensitivity to different indices by dropping them (or several of them) in succession to evaluate their contribution to assessment models.

Regarding other fishery CPUE indices, they were taken as given and no formal presentations on their derivation was given.

2.4 Key issues

The data organization is necessarily complex given the number of fisheries and highly interested members of ICCAT. This poses serious challenges for any assessment approach. At this meeting, the three draft forms of the models presented were far from being appropriate as an "acceptable" assessment in the conventional sense. Nonetheless, each approach attempts to integrate the available data and on their own can provide insights on stock trends, issues with assumptions, and what type of data might best help in the future. The initial main issues as I see from the drafts presented in April were:

- SS3:
    - Size-selectivity may interact with growth in unclear ways
    - The need to handle variable selectivity over time (burdens computation time)
    - High fishing mortality issue and the "hybrid F" option (may need to estimate Fs directly)
    - Constant, fixed length-weight relationship
    - Having so many fleets adds to the complexity of assumptions and on the computation of reference points
    - The model may have difficulty anchoring the scale of the population (e.g., based on examinations of flat likelihoods when profiled over the R_0 parameter
- ASAP:
    - Single fleet (apparently to be revised in subsequent developments; but then same issues will arise a in the SS3 case)
    - "selectivity blocks" may require added evaluations/justifications (including adding variability); lacks flexibility in selectivity-at-age. Separable assumptions may be too severe
    - Unclear how body mass-at-age is derived/used and the extent it varies over time
    - Unclear how catch-at-age data can be shown as consistent by areas and times/seasons

- VPA:
    - Single fleet
    - Cohort sliced data
    - Ad-hoc methods
    - Lack of statistics, data are “known” variables
    - Known to have issues in terminal years

Each of these approaches as drafted can handle (to some degree) the uncertainty in the catch during the 1996-2006 period (historically dubbed as inflated catch) since this has been identified as an issue of importance. I think that the VPA approach, which lacks any formal statistical basis, is hard to defend for providing useful advice to the commission. The authors are striving to at least evaluate its application in traditional and rigorous ways, but at the end, the trade-off in the number of assumptions required may be less useful.

The absolute scale of the assessment is also a major uncertainty. There was a proposal to implement some priors on the population scale to constrain the assessments. Further work on profiling such scale parameters is warranted, particularly given other sources of uncertainty (e.g., age-specific natural mortality). In general, a sensitivity analysis where indices are dropped from the assessment (incorrectly termed “jack-knife” during the workshop; see link) should be continued to evaluate the contribution of different indices.

2.4.1 Other recommendations

Regarding the use of VPA, it seems unfortunate that such an approach was needed due to issues related to inflexibility in the separability in fishing mortality. The method seems always more ad-hoc than typical (arguably still subjective) assessment approaches. It seems that too much time was devoted to devising ad-hoc ways to make it “work”. I find that since VPA methods fundamentally ignore the uncertainty on the data used as input misses an important characteristic of the E-BFT assessment process—that the data are highly uncertain! An alternative modeling approach should be developed. One suggestion would be to allow some efficient non-parametric smoothing (see SPRFMO 2021 or Butterworth et al. 2003). These models are efficient (can be fitted to a broad variety of data without much computational overhead) and can provide full Bayesian integration based on recent experience with the adnuts R package (Monnahan and Kristensen 2018, Monnahan 2018).

3 References

Butterworth, D.S, J. N. Ianelli & R. Hilborn (2003) A Statistical Model for Stock Assessment of Southern Bluefin Tuna with Temporal Changes in Selectivity, African Journal of Marine Science, 25:1, 331-361, DOI: 10.2989/18142320309504021.

Ortiz, M., and Palma, C. 2011. Summary of comparison and verification of the AgeIT program for age-slicing of bluefin tuna catch-at-size (CAS) information. ICCAT Collective Volume of Scientific Papers, 66: 918–934.

Monnahan CC, Kristensen K (2018). “No-U-turn sampling for fast Bayesian inference in ADMB and TMB: Introducing the adnuts and tmbstan R packages.” PLoS ONE, 13(5), e0197954.

Monnahan CC (2018). adnuts: No-U-Turn MCMC Sampling for ADMB Models. . R package version 1.1.2.

SPRFMO (2021). 9th Scientific Committee meeting report. 79 p. Wellington, New Zealand 2021. See Technical annex.

4 Adopted Agenda

1. Opening, adoption of agenda and meeting arrangements and assignment of rapporteurs

MSE Process

2. Summary of developments on ABFT-MSE

    2.1 Report on 2022 March Panel 2 meeting on BFT MSE
    2.2 Report of the informal BFT MSE Technical Sub-group February 14-16th, 2022.
    2.3 Review of the scientific papers/presentations relevant to MSE
    2.4 Round-robin from CMPs and changes to CMPs based on Panel 2/Commission input
    2.5 Summary of CMP performance metrics based on Panel 2/Commission input
        2.5.1 Key figures and plots
    2.6 Update performance statistics based on initial operational management objectives
        2.6.1 Request from Panel 2 to provide a B_{LIM}
        2.6.2 Fishing mortality metric
        2.6.3 Other statistics
    2.7 Specification of final MSE robustness trials
    2.8 Decision process for CMP development and performance tuning and eventual selection
        2.8.1 Process for development tuning and performance tuning
        2.8.2 Satisficing
        2.8.3 Other considerations
            2.8.3.1 Description of stock recruitment relationships in operating models
    2.9 Initial cull of CMPs
    2.10 Communications material
        2.10.1 Key plots and outputs
        2.10.2 Develop presentation to Panel 2 on progress
        2.10.3 BFTMSE Ambassadors programme
    2.11 Path forward for the BFT MSE process
    2.12 Update of trial specification document (TSD)

E-ABFT Data Preparatory

3. Review of the scientific papers relevant to E-BFT stock assessment

4. Presentation of initial data inputs

    4.1 Biology and age data
    4.2 Size and age composition, update stereo-camera data through to 2020
    4.3 Catch Estimates
        4.3.1 Task I Nominal Catches
        4.3.2 Assumptions about catches in 2021and 2022 for projections
        4.3.3 Assumptions regarding past inflated catch and recent IUU
    4.4 Indices of abundance

5. Detailed ToRs for E-BFT stock assessment (VPA, Stock Synthesis, and ASAP)

    5.1 Specify runs

6. Workplan leading to the July assessment

7. General discussion of GBYP matters including Close-kin

8. Other matters

    8.1 BFT Technical Sub-group on Growth in farms and other Docs

9. Adoption of the report and closure

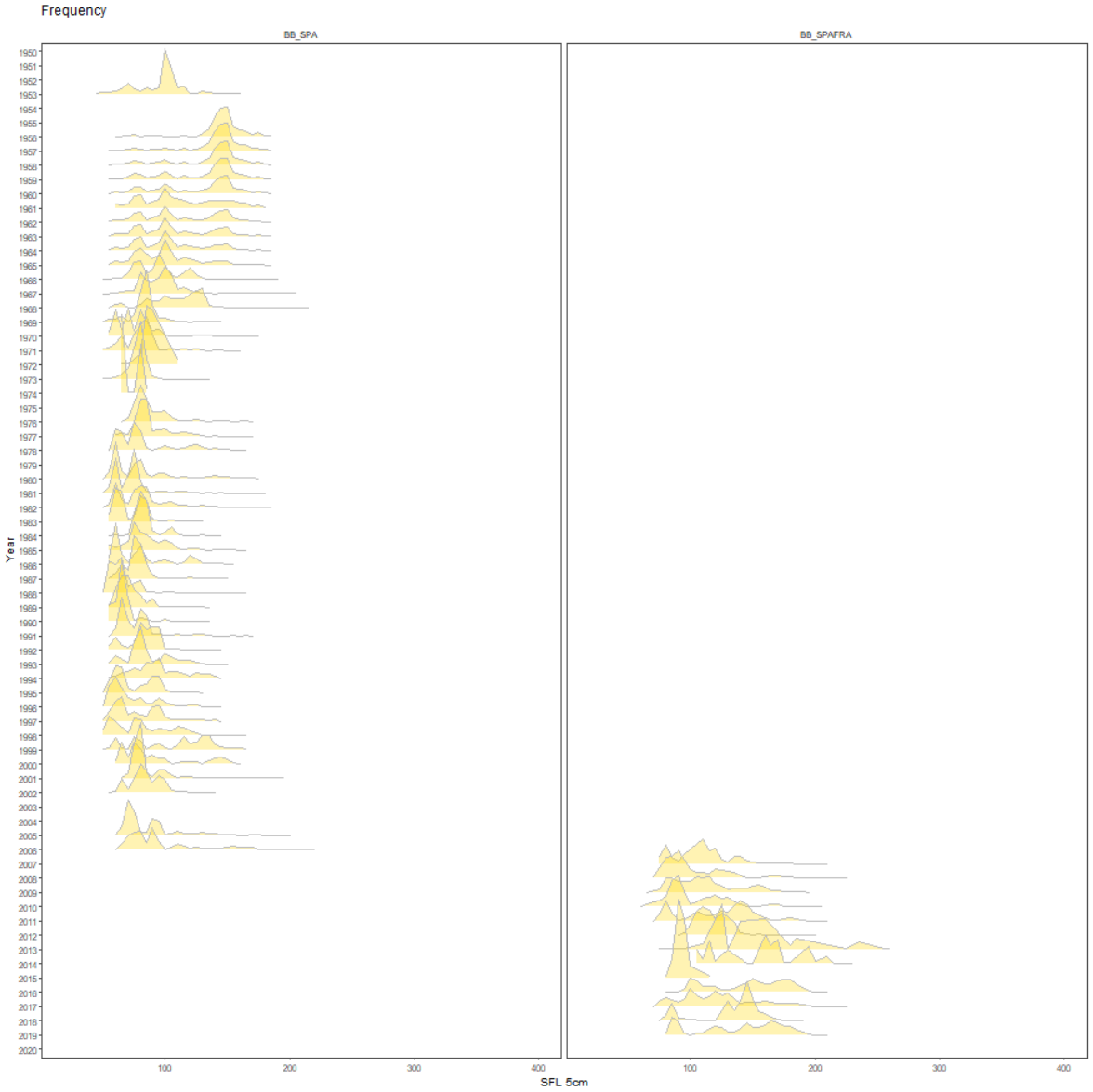
5 List of papers

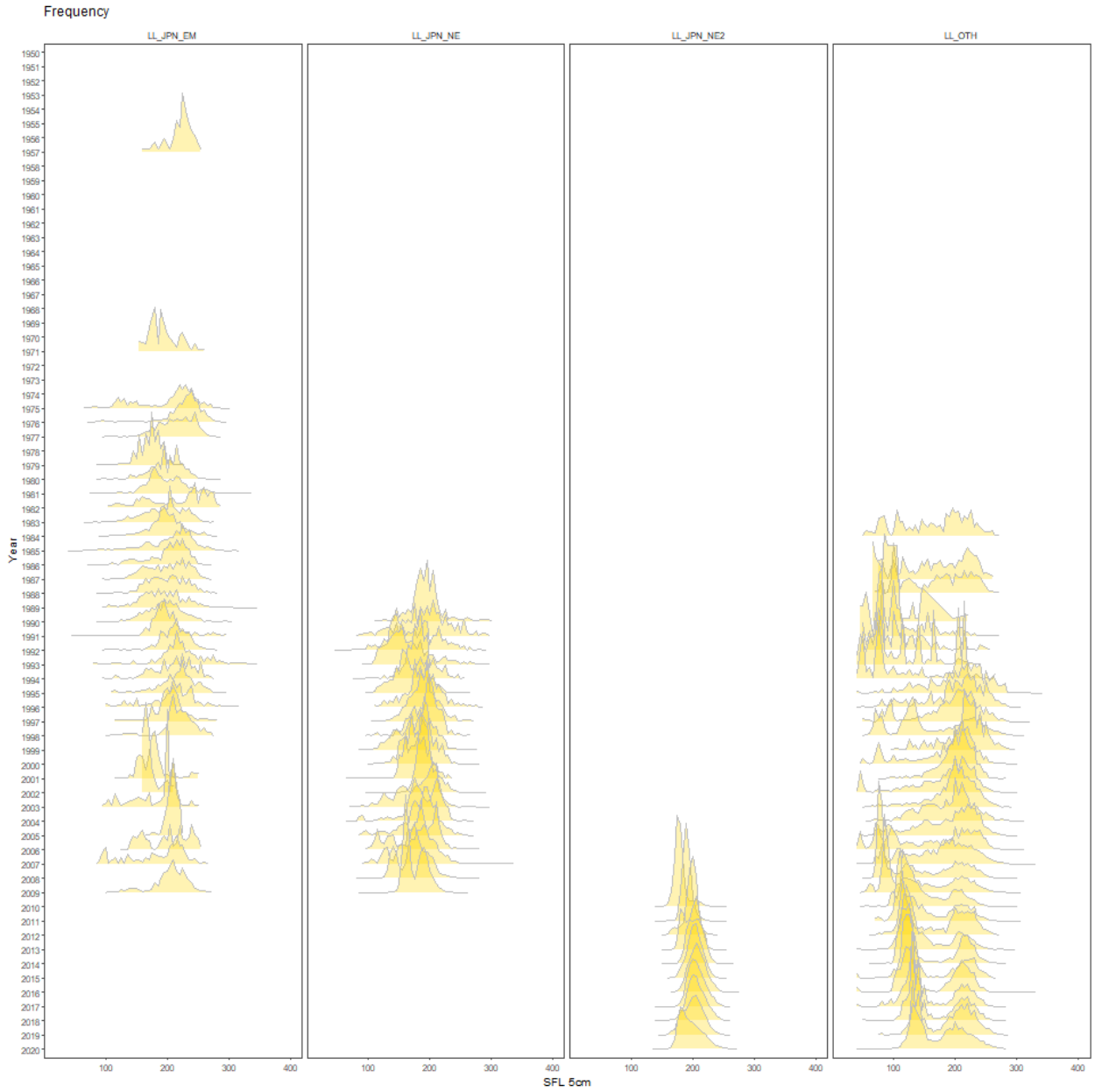
Number	Title	Authors
SCRS/2022/066	ACOUSTIC-BASED FISHERY-INDEPENDENT ABUNDANCE INDEX OF BLUEFIN TUNA IN THE BAY OF BISCAY: RESULTS FROM THE FIRST SEVEN SURVEYS	Onandia I., Goñi N., Uranga J., Arregui I., Martinez U., Boyra G., Melvin G.D., Godard I., Arrizabalaga H
SCRS/2022/067	DATA AND INITIAL MODEL SET-UP FOR THE 2022 VPA STOCK ASSESSMENT OF THE EASTERN ATLANTIC AND MEDITERRANEAN BLUEFIN TUNA	Rouyer T., Kimoto A., Zarrad R., Ortiz M., Palma C., Mayor C., Lauretta M., Rodriguez-Marin E., and Walter J.
SCRS/2022/068	UPDATE OF THE FRENCH AERIAL ABUNDANCE INDEX FOR 2021	Rouyer T., Derridj O., and Fromentin J.M.
SCRS/2022/069	UPDATE OF ELECTRONIC TAGGING DATA AND METHODOLOGIES FOR ATLANTIC BLUEFIN TUNA IN ORDER TO PLAN FUTURE TAGGING ACTIVITIES	Aarestrup K., Alemany F., Arregui I., Arrizabalaga H., Cabanellas-Reboredo M., Carruthers T., Hanke A., Lauretta M., Pagá A., Rouyer T., Tensek S., Walter J., and Rodriguez-Marin E.
SCRS/2022/070	DATA AND INITIAL MODEL SET-UP FOR THE 2022 ASAP STOCK ASSESSMENT OF THE EASTERN ATLANTIC AND MEDITERRANEAN BLUEFIN TUNA	Cadrin S.X., Carrano C., and Maguire J.-J.
SCRS/2022/071	RETROCALCULATED LARVAL ABUNDANCE INDEX OF ATLANTIC BLUEFIN TUNA IN THE WESTERN MEDITERRANEAN SEA, 2001-2020.	Alvarez-Berastegui D., Tugores M.P., Martín M., Lineth N., Pérez-Torres, A.P., Balbín R., and Reglero P.
SCRS/2022/072	A REVIEW OF AVAILABLE INFORMATION FOR THE EASTERN ATLANTIC BLUEFIN TUNA USING CHINESE LONGLINER OBSERVER DATA FOR THE PERIOD 2013-2019	Feng J., Zhang F., Zhu J., and Wu F.
SCRS/2022/073	THE STANDARDIZED CPUE FOR JAPANESE LONGLINE FISHERY IN THE ATLANTIC UP TO 2021	Tsukahara Y., Fukuda H., and Nakatsuka S.
SCRS/2022/074	A SIMPLE CANDIDATE MANAGEMENT PROCEDURE USING JAPANESE LONGLINE INDICES	Tsukahara Y., and Nakatsuka S.
SCRS/2022/075	DESCRIPTION OF THE ICCAT LENGTH AT AGE DATA BASE FOR BLUEFIN TUNA FROM THE EASTERN ATLANTIC, INCLUDING THE MEDITERRANEAN SEA.	Rodriguez-Marin E., Quelle P., and Busawon D.
SCRS/2022/076	REPORT OF THE MANAGEMENT STRATEGY EVALUATION TECHNICAL SUB-GROUP FEBRUARY 14-16TH, 2022	Walter J., and Peterson C.
SCRS/2022/077	A PROPOSAL FOR A BIOMASS LIMIT REFERENCE POINT (Blim) FOR ATLANTIC BLUEFIN TUNA	Walter J., Butterworth D., and Rodriguez-Marin E.
SCRS/2022/078	EFFECT OF TUNING A CMP TO EACH RECRUITMENT SCENARIO WITHIN THE ATLANTIC BLUEFIN TUNA MSE	Peterson C., Lauretta M., and Walter J.
SCRS/2022/079	DATA AND INITIAL MODEL SET-UP FOR THE 2022 STOCK SYNTHESIS STOCK ASSESSMENT OF THE EASTERN ATLANTIC AND MEDITERRANEAN BLUEFIN TUNA	Sampedro P., Kimoto A., Ortiz M., Sharma, R., Fukuda, H., Gordoa, A., Lauretta, M., Rouyer T., Sunderlöf, A., Tsukahara Y., Walter J., and Rodríguez-Marín E.
SCRS/2022/080	BFT MSE OPERATING MODEL INDEX PROJECTIONS AND QUESTIONS OF PLAUSIBILITY: ARE THESE FUTURES POSSIBLE?	Duprey N.M.T., and Hanke A.R.

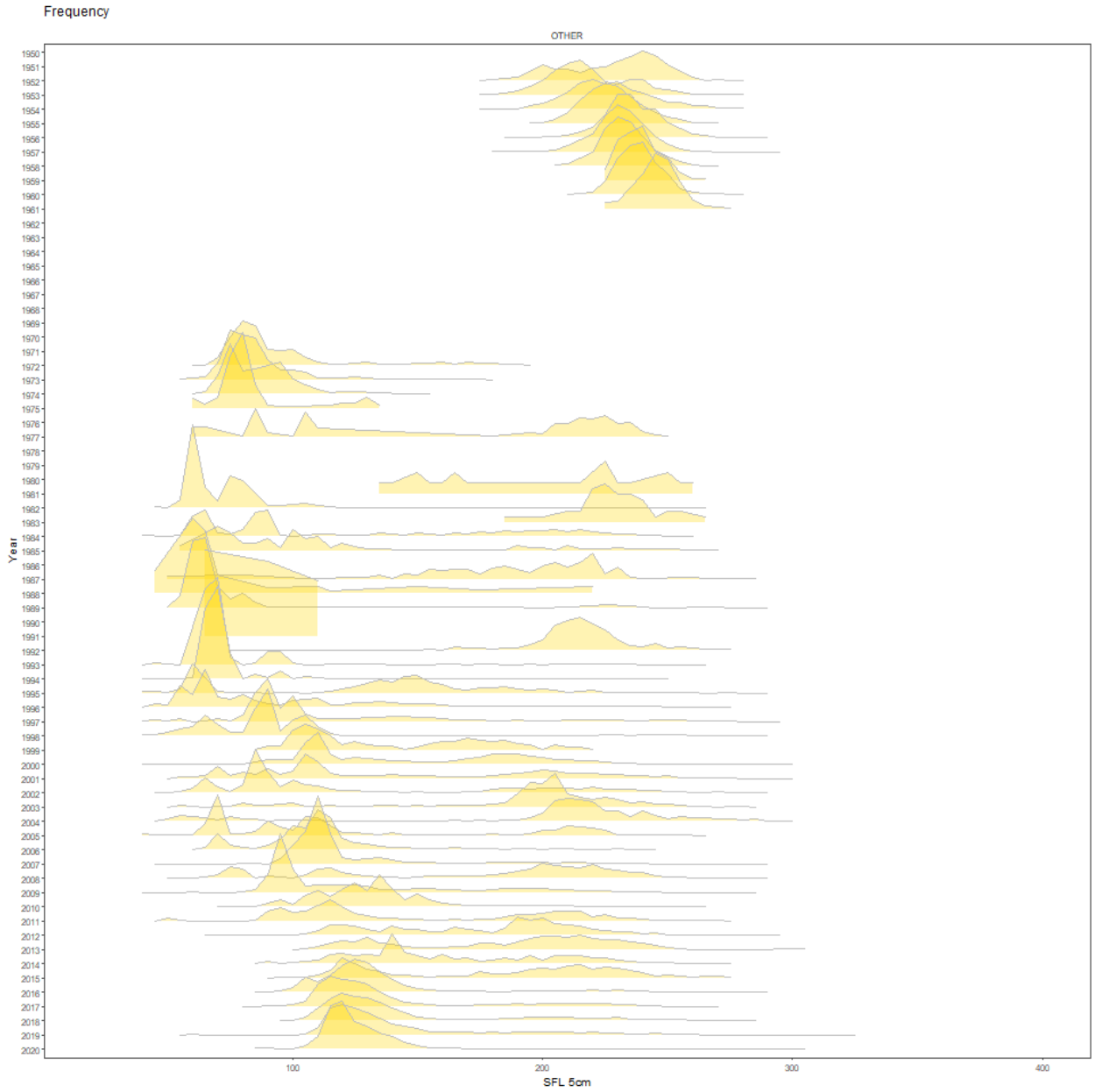
SCRS/2022/081	PUTTING CANDIDATE MANAGEMENT PROCEDURES INTO PRACTICE	Duprey N.M.T., Hanke A.R., Butterworth D. S., Rademeyer R. A., Peterson C., Lauretta M., and Walter J.
SCRS/2022/082	REFINEMENTS OF THE BR CMP AS AT APRIL 2022	Butterworth D. S., and Rademeyer R. A.
SCRS/P/2022/011	APPLYING MIXED-EFFECTS GROWTH MODELS TO BACK-CALCULATED SIZE-AT-AGE DATA FOR ATLANTIC BLUEFIN TUNA	Stewart N.D., Busawon D.S., Rodriguez-Marin E., Siskey M., and Hanke A.
SCRS/P/2022/012	ESTIMATING AGE-AT-MATURITY FROM BIPHASIC GROWTH MODELS FOR ATLANTIC BLUEFIN TUNA	Stewart N.D., Busawon D.S., Rodriguez-Marin E., Siskey M., Wilson K., and Hanke A.
SCRS/P/2022/013	PRELIMINARY CMP RESULTS APRIL 2022	Carruthers T.
SCRS/P/2022/014	AN EXPLOITATION RATE PROPOSAL FOR AN APPROPRIATE MSE PERFORMANCE METRIC RELATING TO FISHING MORTALITY	Carruthers T.

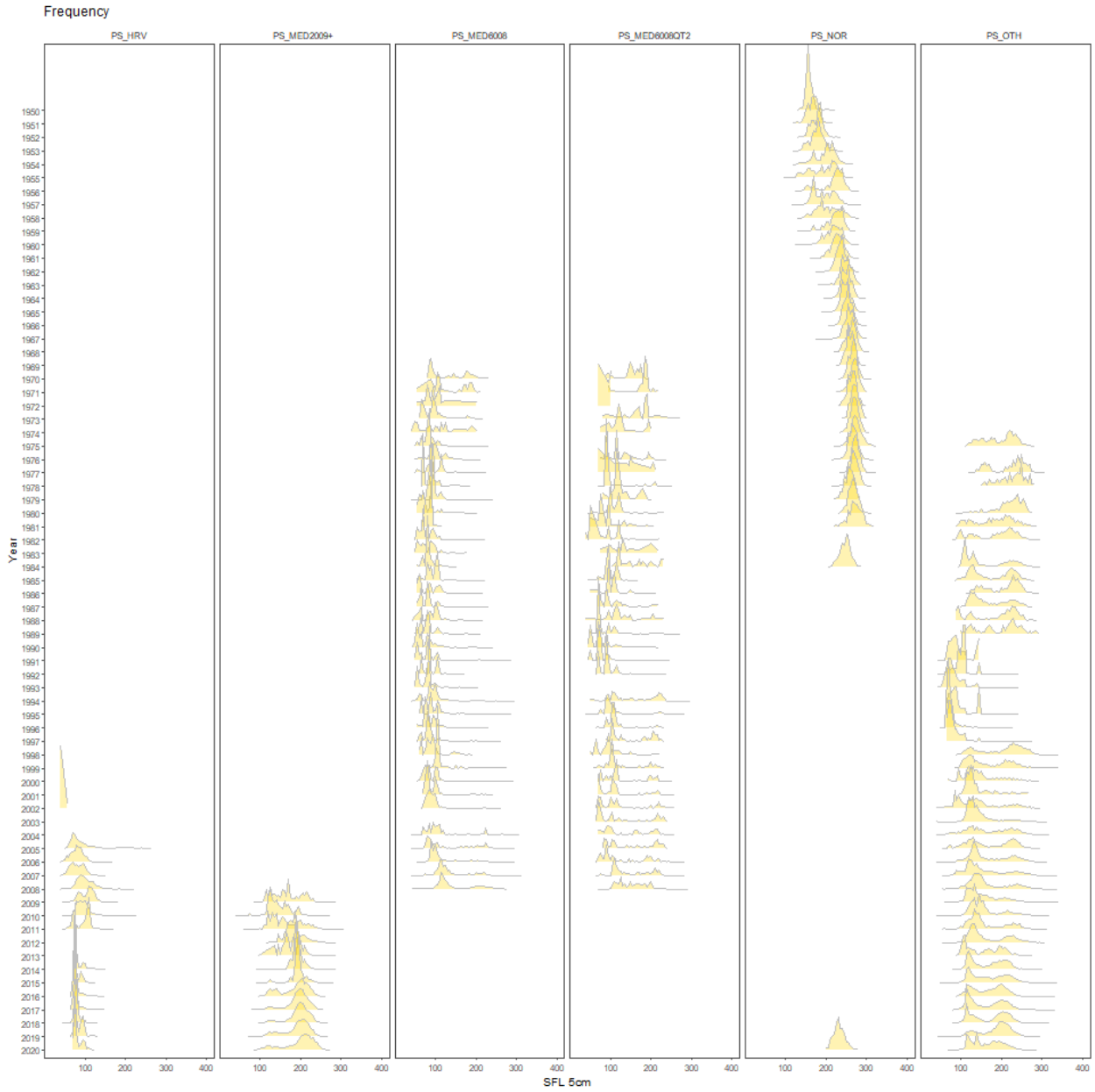
6 Length frequencies

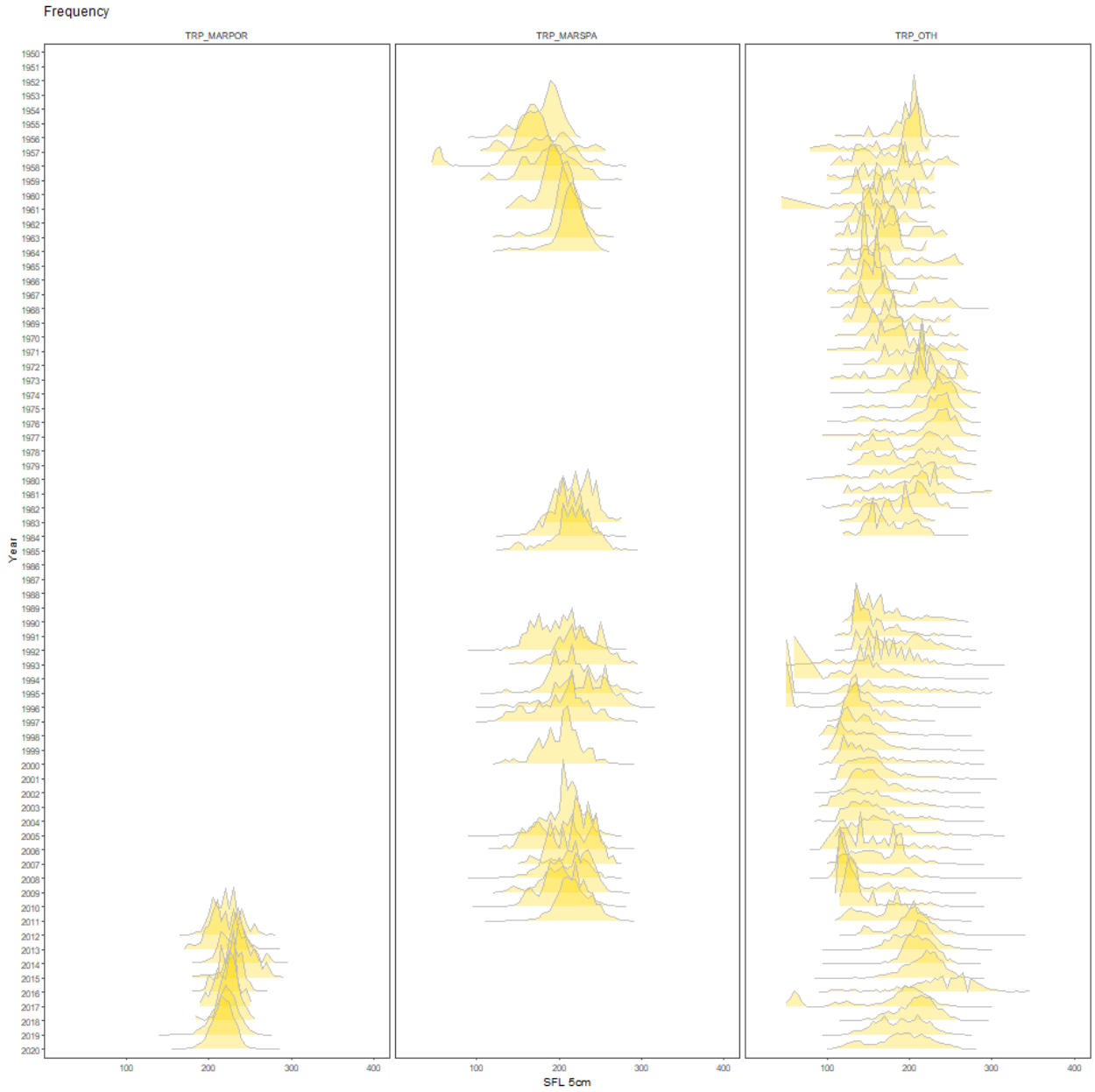
By gear type (each page) and fleet (panels within page), 1950-2020.




Frequency
LL_JPN_EM
LL_JPN_NE
LL_JPN_NE2
LL_OTH
Year
SFL 5cm


Frequency
OTHER
Year
SFL 5cm
100
200
300
400


Frequency
PS_HRV
PS_MED2009+
PS_MED6008
PS_MED6008QT2
PS_NOR
PS_OTH
Year
SFL 5cm


Frequency
TRP_MARPOR
TRP_MARSPA
TRP_OTH
Year
SFL 5cm