Fisheries and Oceans Canada | Pêches et Océans Canada
Ecosystems and Oceans Science | Sciences des écosystèmes et des océans

Canadian Science Advisory Secretariat (CSAS)

Research Document 2022/015

Quebec Region

Details of `catchR`, an *R* package to estimate the age and length composition of fishery catches, with an application to 3Pn4RS Atlantic cod

Jordan Ouellette-Plante, Elisabeth Van Beveren, Hugues P. Benoît and Claude Brassard

Fisheries and Oceans Canada
Maurice Lamontagne Institute
P.O. Box 1000, 850 Route de la Mer
Mont-Joli (Quebec) G5H 3Z4

April 2022

Canada

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

© Her Majesty the Queen in Right of Canada, 2022
ISSN 1919-5044
ISBN 978-0-660-42532-0 Cat. No. Fs70-5/2022-015E-PDF

Correct citation for this publication:

Ouellette-Plante, J., Van Beveren, E., Benoît, H.P. and Brassard, C. 2022. Details of `catchR`, an *R* package to estimate the age and length composition of fishery catches, with an application to 3Pn4RS Atlantic cod. DFO Can. Sci. Advis. Sec. Res. Doc. 2022/015. iv + 69 p.

Aussi disponible en français :

Ouellette-Plante, J., Van Beveren, E., Benoît, H.P. et Brassard, C. 2022. Détails de `catchR`, un paquet R pour estimer la composition en âge et en longueur des captures de la pêche, avec une application à la morue franche de 3Pn4RS. Secr. can. des avis sci. du MPO. Doc. de rech. 2022/015. iv + 71 p.

TABLE OF CONTENTS

ABSTRACT

Estimating the age and length composition of catches is common practice in stock assessment, as such information on population structure improves our understanding of stock productivity and the impacts of fishing on it. Since aging fish is costly, labour intensive and requires experienced staff, age composition is often inferred from stratified subsampling of age at length and the length composition of the original sample. Although the underlying principle is relatively straightforward, gaps and inconsistencies in the sampling process have led to multiple proposed algorithms in the literature. Importantly, these algorithms are often applied annually in a manner that can be ad hoc and undocumented. Replicating estimates of catches at age from source data can therefore be difficult and can impact the credibility of assessments. For this work, we employed the commonly used forward age-length key coupled to a hierarchical stratified procedure based algorithm for filling gaps in sampling. The *R* package `catchR` resulted from this work, and proposes a new, fully transparent and automated framework to estimate catch-at-age. Using the NAFO 3Pn4RS Atlantic cod (*Gadus morhua*) stock as an example, this research document provides a description of the main `catchR` functions, `get.samples` and `get.caa`, used to calculate catch-at-age. For the 3Pn4RS Atlantic cod stock, the automated algorithm, as currently defined, exposes the shortcomings in the estimation of catch- and weight-at-age, in part because the result of each year is now fully comparable and because no subjective adjustments can be applied to correct for example a poor sampling coverage for any given combination of year, month, NAFO unit area and gear.

1. INTRODUCTION

Estimating the age and length composition of catches is common practice in stock assessment, as such information on population structure improves our understanding of the stock's productivity (growth, natural mortality and recruitment, Ailloud and Hoenig 2019) and the impact of fishing on it. Since aging fish is costly, labour intensive and requires experienced staff, the approach used since almost a century now (Fridriksson 1934) is to initially obtain from a given catch a sample from which individual lengths are measured. Then, using a length stratified subsample from which specimens are also aged, the proportion of fish at each age for individual length classes, commonly termed an age-length key (ALK) in fisheries, will be obtained. The ALK is finally used to convert a sample of fish lengths into an estimate of the frequency distribution of ages. For example, estimating catch-at-age (CAA) for the commercial fishery will typically involve collecting a large random sample of that catch for its length composition (length frequency samples, hereinafter termed LF data) from which a much smaller subsample of randomly selected fish will be obtained using a length-stratified sampling[1].

Since fish[2] do not grow at the same pace, all fish from a given length class will not be of the same age and an ALK will then be needed to convert length to age (Gulland and Rosenberg 1992). Although the underlying principle is relatively straightforward, gaps[3], (sub)sampling error and stochasticity associated with small sample size[4] have led to the proposition of multiple algorithms by various authors. For instance, rather than applying a classical forward ALK, which involves a probabilistic description of age given length, some authors proposed inverse (probabilistic description of length given age) or combined forward-inverse ALKs (Ailloud and Hoenig 2019). Forward ALKs can also be predicted by a variety of models to mitigate the impact of data gaps and inconsistencies (Berg and Kristensen 2012; Babyn et al. 2021).

For this work, we went with the commonly used forward ALK coupled to a hierarchical stratified procedure based algorithm for filling gaps in sampling. We propose a new, transparent and automated framework to estimate CAA and which we implemented in a new package for the *R* statistical software (R Core Team 2020) called `catchR`. This new framework is flexible enough to allow future improvements if desired.

Using the 3Pn4RS Atlantic cod (Gadus morhua) stock as an example, this research document provides a description of the main `catchR` functions, `get.samples` and `get.caa`, used to calculate CAA, and compares the results with previously published values. Other functions are described in Appendix A. This research document is the result of the Canadian Science Advisory Secretariat (CSAS) peer review of the assessment framework for Atlantic cod in the Northern Gulf of St. Lawrence (NAFO[5] Subdivision 3Pn and Divisions 4RS), held on April 21–22 and May

[1] Introduced by Ketchen (1950), the length stratified sampling approach's first objective is to get a better representation for lengths that are at the far ends of the length distribution.

[2] As a matter of fact, pretty much all living organisms.

[3] A gap arises when a length class from the LF sample does not have a correspondance in the ALK. This will prevent obtaining age proportions (using the classical forward ALK).

[4] Ex: all fish from a particular length class will be presumed age *x* because only one fish was aged for that particular length class, and was age *x*.

[5] Northwest Atlantic Fisheries Organization.

12 of 2021 and focused on available data to be used in a forthcoming assessment model for the stock.

2. METHODS

2.1. PORT SAMPLING PROGRAM

Fisheries and Oceans Canada (DFO) has a port sampling program for commercial landings that was specifically designed to estimate catch composition (Lambert and Ménager 1998; see Daigle and Benoît 2007). Each year the catch of a representative sample of fishing trips is sampled to obtain data on the length and possibly age composition of the catches of targeted species. The following descriptions are presented in more details in Lambert and Ménager (1998). For 3Pn4RS Atlantic cod, a maximum of 150[6] fish are sampled for length measurements[7] over the entire catch. When the catch is already sorted into size classes by members of the fishing industry, sampling and data recording is done by size class. This following information is recorded for each sample or subsample:

- Species landed
- Sample ID
- Date
- Port sampler ID
- Vessel name
- vessel ID
- Landing district
- Fishing zone
- Weight landed
- Gear used
- Depth
- Number of categories
- Sampled category
- Weight landed for the category
- Sampled weight
- State at landing
- Group
- Measure type

[6] Although the publication from Lambert and Ménager (1998) mentions 250 fish, the current protocols for cod (not published) use 150 fish.
[7] Fork length, rounded to the nearest cm.

Then, fish length measurements are performed and otoliths from three fish per 3 cm length increment are collected (hereinafter termed age data). For each pair of otoliths collected from a fish, the following information is recorded:

- Sample ID
- Date
- Species
- District
- Gear
- Fishing zone
- Vessel name
- Vessel ID
- Fork length
- Otolith ID

Otoliths are then sent to DFO science for age readings. Finally, these data are entered into an *Oracle* database.

2.2. CONTEXT FOR ATLANTIC COD CAA

CAA for 3Pn4RS cod has been estimated for about five decades starting in 1973 (Wells 1979). From the stock assessment of Fréchet and Gascon (1986) to that of Fréchet et al. (2009)[8], CAA has been estimated with the `CATCH.AWS` software, written in STSC APL language (Anon. 1986) and following the methodology described in Gavaris and Gavaris (1983). With the arrival of computers equipped with 64-bit processors, this software was rewritten in the more modern Visual Basic language during 2011–2012 and was renamed `CATCH` (Doniol-Valcroze et al. 2019). `CATCH` was first adopted for estimating CAA for the February 2015 cod stock assessment (Brassard et al. 2016) and again in the 2017 and 2019 assessments (Brassard et al. 2018, 2020). Over time, some drawbacks of the software were however noted, mostly related to transparency and reproducibility. Specifically, `CATCH` is a black box system for which the internal workings are not understood by all current users, as the source code and detailed documentation are missing. This precludes a thorough understanding of the results and excludes the possibility to expand the methodology (by allowing a + group, automating sample selection, etc.). So far, it has also been impossible to precisely reproduce results because `CATCH` has an error-prone data entry system and requires the user to make subjective manual decisions that are not always recorded and reported. This contributes to an important lack of reproducibility.

Port sampling data were extracted[9] using an in-house SAS program (`autoexec_peche.sas`), described in Appendix B. Unfortunately, this program only permits access to data that begin in

[8] Some research documents in this series (Fréchet and Schwab 1998; i.e. Fréchet et al. 2002, 2003, 2005) do not specifically mention the program. However, it is reasonable to assume it was used given its mention in subsequent research documents.
[9] Both LF and age data were extracted on March 29, 2021.

1981. However, it is clear that there are port sampling data going back at least since 1974[10] based on published research documents (Table 1).

For the years for which the extraction provided data, there were strong differences in the number of cod available compared to previously reported values (Table 1). This situation is abnormal since for a given year, it would have been expected to obtain either the same data, or an increased number of values, given that some data may have been added following a research document publication. However, this was not the case and large differences were observed, especially for LF data where published numbers sometimes were 3.5 times higher (ex: 1991) than what could be extracted.

A reason that could explain these missing data could be that the data may be spread into different databases. Newfoundland-and-Labrador (NL) region was originally responsible for scientific research and advice in the estuary and northern Gulf of St. Lawrence (nGSL) until 1983. A copy of the data covering the 1974–1982 period was acquired from the NL region in 1984 (Fréchet and Gascon 1986), but could not be found for the present work.

Another reason for these differences could be caused by samples originating from foreign fleets. Several countries fished in the 3Pn4RS stock before the Canadian fishing zone was extended to 200 nautical miles from the coast in 1977 (Table 2, Sanguin 1980). Even after 1977, France (mainland and/or Saint Pierre and Miquelon) continued to exploit this stock until 1992. According to Fréchet and Gascon (1986), data from foreign fleets were available either from NAFO or by the international at-sea observer program operated by DFO, but we were unable to find these data from either of these sources.

Taking into account that we wanted to compare our results of CAA to the ones already published and that the data available for the present work were not identical to the one used in previous publications for the 1974–1992 period, it was decided to subset the available data and keep only the 1993–2020 period for the analyses. However, even with this small period, it was not possible to validate if the extraction provided similar data to those used in the calculations of catches at age after 1994 since this information was no longer provided in research documents (Table 1).

2.3. LF DATA

LF data were imported into *R* using the `read.lf`[11] function:

```
read.lf(file, year = NULL, language = "en", ...)
```

where:

`file` .dat file to read. This file is the output of the SAS program `autoexec_peche.sas` described in Appendix B.

`year` Vector of years to read (ex: 2015:2020). NULL by default, meaning the function will keep data from all available years.

`language` Language to use for column names, either “en” (default) or “fr.”

`...` Optional arguments used with `readr::read_fwf()`.

[10] Fréchet and Gascon (1986) mentioned that poor sampling in the cod targeted commercial fisheries was explaining the exclusion of previous years.
[11] Unless stated otherwise, the *R* functions mentioned come from the catchR package.

A description of the columns of the dataset created by this function is provided online.

Length validation was performed on the anterior portion of the overall length distribution in order to detect outliers. Only one cod was considered to be an outlier as it was a 7 cm long and originated from a sample obtained from longline activities with significantly larger values for the other lengths recorded for the activity. This observation was removed prior to CAA calculations.

Since cod are not always landed whole (round, Table 3), conversion factors are required to produce round-weight equivalents for fish that are gutted (dressed) and/or head-off (STACAC 1984a). For example, a catch of 100 kg of cod landed in the eviscerated head on form will be adjusted to 120 kg of whole fish using a conversion factor of 1.2. However, using such conversion factors regardless of the season implies that eviscerated head on cod of identical lengths and weights caught at different times of the year will be assumed to have a constant weight of viscera during the whole year, which is incorrect given seasonal variation in somatic condition. For instance, the Fulton condition index is at its lowest in the spring before spawning and then increases in the fall (Brassard et al. 2020). One might therefore expect an increased conversion factor for eviscerated cod in the breeding season (Kulka 1981). With regard to the LF data, we also noted that it was sometimes impossible to use conversion factors since sample weights were at times not reported (sometimes up to almost 100% of the samples in a given year, Table 3). For these reasons, a recalculation of the total mass w_m of each sample m[12] was performed using the different length-weight relationships available (see section 2.5).

In addition, commercial landings generated by the use of shrimp trawls (see section 2.6) were reported during the series. Unfortunately, these landings were never sampled by port samplers. What is more, the Nordmore grate, which limits the capture of cod and other groundfish to smaller-sized specimens, has been employed in the northern shrimp fishery since 1993 (Savard et al. 2013). Its implementation therefore did not allow the different trawl types to be grouped together since fish caught by shrimp trawls are in proportion much smaller than for other trawl types. For this reason, it was decided to fill this data gap by using data from the at-sea observer program. Of these data, only fishing activities carried out in NAFO Subdivision 3Pn and Divisions 4RS and for which the target species was northern shrimp were selected. The different fishing gears used with LF data were aggregated into seven gear categories based on potential differences in selectivity (Table 4, Figure 1). The same categories were also used for *age* data (see next section).

2.4. *AGE* DATA

The *age* data were imported into *R* using the `read.bio` function:

```
read.bio(file, year = NULL, sp, language = "en", ...)
```

where:

`file`	.dat file to read. This file is the output of the SAS program `autoexec_peche.sas` described in Appendix B.
`year`	Vector of years to read (ex: 2015:2020). NULL by default, meaning the function will keep data from all available years.
`sp`	Species. Either “cod” or “mackerel.”
`language`	Language to use for column names, either “en” (default) or “fr.”
…	Optional arguments used with `readr::read_fwf()` (e.g., `progress`, `skip_empty_rows`, `n_max`).

A description of the columns of the dataset created by this function is provided online.

By comparing the number of samples[12] available for the 1993–2020 period between the *age* and *LF* datasets (see previous section), it became evident that a problem existed since there were more *age* than *LF* samples for some years. From the port sampler protocol presented in section 2.1, an equal or reduced number of age samples relative to LF samples would have been expected. Digging through the data, it became evident that 1,478 *age* samples, all from years 1994–1997, did not have their equivalent in the *LF* dataset. For these samples, the fact that thousands of cod had their length recorded but no age reading supported this idea that these samples should also have been included in the *LF* dataset, which was corrected afterwards. For these samples, certain information normally found in the *LF* dataset could not be found (e.g., sample and landing weights, state at landing) since the *age* dataset did not provide this information (Table 3).

After having resolved this discrepancy between the two datasets, the *age* dataset was filtered to keep only the observations where both a length and an age value were provided. Then, a validation of the age-length relationship was performed. Outliers were identified by age class as observations with lengths outside three times the interquartile range (Figure 2). Four outliers were identified this way. The only 1-year-old cod recorded in the 1993–2020 period was also identified as an outlier since its length did not match with the general growth increments observed as age increases (Figure 2). These five outliers were removed prior to CAA calculations.

Moreover, to overcome the problem of the lack of data for fishing activities using shrimp trawls mentioned above, it was decided to use data from the at-sea observer program. This alternative worked well for LF data. However, no aged cod from this program was found, and it was decided to use the age data from the DFO August multispecies bottom trawl survey performed each year in the nGSL as an alternative (see Bourdages et al. 2020 for a description). For all years where such landings were present, cod individuals caught in NAFO Subdivision 3Pn and Divisions 4RS and having both a valid length and age were used from this survey. These data were added to the *age* dataset as if all individuals from a given year were a sample collected in August. A visual inspection of the ALKs created from the two data sources was done (an example with years 2008–2010 is given in Figure 3), and a good overlap was observed. Also, the use of data from this survey allowed the aging of the smaller-sized individuals caught in shrimp trawls, which were rarely observed in the DFO port sampling program database (Figures 1 and 3).

2.5. LENGTH-WEIGHT RELATIONSHIPS

For this study, length-weight relationships from research vessel surveys conducted in the summer (1993–2020) and winter (1993–1994) were used. For the two years when both surveys were performed (1993–1994), period-specific relationships could be calculated (October–March from the winter survey and April–September from the summer survey). For years for which no winter survey was performed (1995–2020), the length-weight relationship obtained from the summer survey was used for the full year.

The annual (y) and periodic (p) specific relationships between length and weight were modelled as follows, with estimates of the parameters α and β provided in Table 5.

---

[12] For both the *LF* and *age* datasets, a shared variable, *sample.id* was available. For the *LF* dataset, the variable identifies each unique length frequency. For the *age* dataset, the variable *sample.id* makes it possible to know from which LF samples age readings were performed.

$$log_e(mass_{yp}) \sim \alpha_{yp} + \beta_{yp} \cdot log_e(length_{yp}) \quad (1)$$

Each relationship was modelled twice. After a first run with all data available (see column *Available* of Table 5), a data cleansing of the outliers was carried out to remove values for which corresponding studentized residuals were > 1.5 (absolute value, Bourdages and Ouellet 2011). Then, using only the remaining data (column *Used* of Table 5), the length-weight relationship was finally modelled. The relationships were all verified visually to remove outliers that could contribute leverage.

The weight of fish recorded in LF data was calculated using the length-weight relationship corresponding to the year and period from which the associated sample was taken from.

2.6. LANDINGS DATA

3Pn4RS cod landings data for the 1993–2020 period were extracted[13] from ZIFF data (*Zonal Interchange File Format*, STACAC 1984b). This dataset only includes the landings from the Canadian fleet, but since no foreign fleet landings have been recorded since 1992 (Table 2), the use of NAFO data was not necessary. ZIFF data were read into *R* with the `read.ziff` function:

```
read.ziff(sp, path, year = NULL, language = "en")
```

where:

`sp`	Numeric species code (ex: 100 for Atlantic cod). Can also be a vector of multiple numeric species codes.
`path`	Directory where files to read are located. Ex: "//dcqcimlna01a/BD_Peches/Ziff/Format CSV/Fichiers de données/."
`year`	Vector of years to read (ex: 2015:2020). NULL by default, meaning the function will keep data from all available years.
`language`	Language to use for column names, either "en" (default) or "fr."

A description of the columns of the dataset created by this function is provided online.

A validation was initially done to ensure that the extracted data matched with what has been published in the past. Since the majority of published research documents presented landings in tables broken down by month and management year, we thus compared total landings in this format, despite that for the CAA calculations calendar years rather than management years are used (Table 6). Small differences can be seen at different times of the series, and most can be explained by the fact that the landings were still preliminary at the moment of publishing these previous research documents (Figure 4). Good examples are the research documents 2007/068 (Fréchet et al. 2007) and 2016/010 (Brassard et al. 2016 fig. 4b). However, we found a couple of unexplained differences between values obtained for this study and what was already published. For example, landings data used in this study were short by more than 250 t for management year 2001/2002 compared to what was published.

The different gear used in the landings were aggregated into eight gear categories based on potential differences in catchability (Table 7). Monthly landings using the calendar year for the 1993–2020 period are provided in Table 8. They are also broken down by NAFO unit area and gear category in Tables 9 and 10, respectively.

[13] Extracted on December 16, 2020.

2.7. CALCULATIONS

2.7.1. Search for samples

Landings from the 3Pn4RS cod stock were divided into strata defined by year, month, NAFO unit area (hereafter referred to as *NAFO*) and gear category (hereafter referred to as *gear*). Each stratum k (k = 1, 2, …, K) had a corresponding cod landing W_k. Most of the time, it was not possible to associate even just one sample of each type (either LF or age) per landing W_k. To infer the LF and age composition of these landings, values were imputed using samples from other strata. In the previously used `CATCH` software, this process was done manually and imputation decisions were made by individual analysts, potentially leading to different results despite using the same underlying data. Failure to document imputation decisions have affected the replicability of results.

With the `catchR` package, we developed an automated and reproducible way of estimating the length and age composition of catches in strata for which there were data and strata requiring imputation. For each type of sample (LF or age), the selection of the M samples (m = 1, 2, …, M) to be associated with the landings W_k was done using the `get.samples` function:

```
get.sample(  catch, lf, al = NULL, min.al.samples = 2, min.lf.samples = 2,
             min.al.fish = 5, period.unit = c("month", "quarter"),
             prob.al = 0.95, subsample = TRUE)
```

where:

`catch` Landings dataset. `catch` is an aggregated version of the dataset created by the `read.ziff` function, where each row represents a stratum. The variables `year` (year of landing), `period` (period, i.e., the month or quarter of the landing), `region` (NAFO unit area of landing), `gear` (gear category associated with the landing) and `catch` (the round-weight equivalent landing in kg) are required.

`lf` Length frequency dataset. Each line represents a length found in a given length frequency sample. The variables `year` (sample year), `period` (period, i.e., the month or quarter in which the sample was taken), `region` (NAFO unit area from which the sample originates), `gear` (gear category associated with the sample), `length` (length measured in the length frequency sample, in cm), `weight.unit` (weight of a fish given the length in column `length`, in kg), `n` (number of fish having this length in the sample given the length in column `length`), and `sample.id` (unique identifier of the sample) are required.

`al` Age dataset. Each line represents an aged fish from a given age sample. The variables `year` (sample year), `period` (period, i.e. the month or quarter in which the sample was taken), `region` (NAFO unit area from which the sample originates), `gear` (gear category associated with the sample), `length` (length of the fish, in cm), `age` (age of the fish, in years) and `sample.id` (unique identifier of the sample) are required. NULL by default, meaning the function will only get samples in order to compute the length composition.

`min.al.samples` Minimum number of samples required for age-length keys. By default = 2.

`min.lf.samples` Minimum number of samples required for length frequencies. By default = 2

`min.al.fish`	Minimum number of fish required for age-length keys. By default = 5.
`period.unit`	Whether `catch`, `lf` and `al` are grouped by "month" (default) or "quarter."
`prob.al`	Maximum probability (0–1, default = 0.95) allowed for any length of a stratum specific LF distribution to be of an age not included in the age-length key. If it is at least `prob.al` likely that an age is missing, the algorithm will continue adding age samples. This addition criteria for age sample selection can be effectively removed by setting `prob.al` to 1.
`subsample`	Logical. True if `al` is composed of length-stratified subsamples of `lf` from which fish were aged. False if all fish from `lf` were always aged.

This function is based on a 12-level hierarchy in which the first level is based on stratum-level estimates using data for that stratum, and subsequent levels employed imputation using values from increasingly dissimilar strata.

$$g_k = \begin{cases} 1 & = \text{year + month + NAFO + gear (corresponds to } k\text{) = if } s \text{ met, else } \downarrow \\ 2 & = \text{year + neighbouring months + NAFO + gear = if } s \text{ met, else } \downarrow \\ 3 & = \text{year + month + gear = if } s \text{ met, else } \downarrow \\ 4 & = \text{year + neighbouring months + gear = if } s \text{ met, else } \downarrow \\ 5 & = \text{year + gear = if } s \text{ met, else } \downarrow \\ 6 & = \text{year = if } s \text{ met, else } \downarrow \\ 7 & = \text{neighbouring years + month + NAFO + gear = if } s \text{ met, else } \downarrow \\ 8 & = \text{neighbouring years + neighbouring months + NAFO + gear = if } s \text{ met, else } \downarrow \\ 9 & = \text{neighbouring years + month + gear = if } s \text{ met, else } \downarrow \\ 10 & = \text{neighbouring years + neighbouring months + gear = if } s \text{ met, else } \downarrow \\ 11 & = \text{neighbouring years + gear = if } s \text{ met, else } \downarrow \\ 12 & = \text{neighbouring years} \end{cases} \quad (2)$$

For both LF and age data, a set of criteria s were postulated as determining factors to set at which level of aggregation the `get.samples` function stopped looking for other samples by continuing on at higher aggregation levels. Depending on the type of data, s had either one or three criteria:

$$s = \begin{cases} \text{LF data: } \mathtt{min.lf.samples} \\ \text{age data: } \mathtt{min.al.samples, min.al.fish, prob.al} \end{cases} \quad (3)$$

For LF data (see Figure 5A for a general view of the dataset), the sole criterion to stop the `get.samples` function from going up the aggregation levels shown in Equation 2 was to have at least `min.lf.samples` (set to 2 for this analysis) independent samples. The decision to go with a minimum threshold in terms of the number of samples rather than the number of fish used in a sample was based on the fact that the sampling unit was the sample and not the individual fish. After having run the decision scheme for LF data, each stratum k was then associated with a subset of the LF data (LF_k) so that all samples m retained for LF represented as closely as possible the associated W_k landings.

For age data, two additional criteria were added: `min.al.fish` and `prob.al` (Equation 3). Firstly, we wanted for this analysis to have at least 100 available aged fish (`min.al.fish`) for each stratum's ALK. This additional criterion prevented the `get.samples` function from stopping aggregating at higher levels after reaching the minimum number of age samples set for this analysis (`min.al.samples`, set to 2 for this analysis). Without the `min.al.fish` argument, it would have been possible to get an ALK created with only a few aged fish (ex: 25), providing they came from at least two samples. Secondly, the `prob.al` criterion was added in the search algorithm to ensure that all lengths present in LF_k had an associated probability as a function of age in order to know if additional ages should be found in the ALK used for stratum k.

The approach is illustrated using stratum *2005–June–4Rb–Gillnet*. For this stratum, the search for LF samples reached the unique criterion `min.lf.samples` at aggregation level 2. As a result 2,306 specimens from 15 samples were used for length frequency calculations. For age data, the search at aggregation level 1 was unsuccessful. At level 2, however, both the criteria for the number of samples (`min.al.fish`, set to 2) and the number of specimens used in the ALK (`min.al.fish`, set to 100) were met.

The last criterion to be met was `prob.al`. In order to evaluate if the data found at aggregation level 2 met this criterion, a global a (alk_{global}) was first required, created by transforming the whole age dataset into an ALK and then adding lengths from the LF dataset that were absent from this ALK. The `multinom` function from the `nnet` *R* package (Venables and Ripley 2002) was then used to construct multinomial logistic regressions to impute ages to all lengths found in alk_{global}. Estimated probabilities were rounded to three decimals in order to avoid values < 0.001, which were assumed unlikely. This rounding corresponded to the maximum level of accuracy one can generally expect from ALKs: to know more precisely the probability of a fish of a given length to be of a certain age, an exceptionally large number of fish would need to have been aged. With this in mind, a probability < 0.00001 of having a cod of a certain length and age was then rounded to zero. From alk_{global} it was then possible to calculate the overall probability of any length to be of a certain age. For all lengths in LF_k, the `get.samples` function verified that it was not extremely likely (`prob.al`, here 95%) that they were of an age not yet included in the ALK, and continued searching otherwise. For the stratum used as an example, the LF data (LF_k) included fish of 27 cm length that, at this stage of the search for age data, were extremely likely to be of an age not yet included in the age data used for this stratum (age_k, Table 11). Without a continued search, these fish would have likely been attributed to a wrong age class (in this case the lowest value in the actual state of ALK_k, age 4), creating a bias in the proportion of fish by age and especially their mean lengths and weights, which are more susceptible to extremes. In Table 11, we observe that ages 2–3 and 15–18 are missing from age_k. One thing to keep in mind is that the decision scheme will, in the worst-case scenario, only look for samples in neighbouring years (Equation 2). So, by comparing these ages to all the ones observed in years 2004–2006, the stratum used in this example being from 2005, only ages 3 and 15 would make sense to keep looking for. In other words, even if a 27 cm cod had a fair probability of being age 2 (0.427 according to alk_{global}), it would not make sense to keep looking for samples in order to get that age since we knew we would never find one age 2 cod (in fact, none have been recorded since 1997, Figure 5B). That leaves us with only age 3 that could be worth continuing the search for, age 15 having a null probability. For this study, the `prob.al` argument from the `get.samples` function was set to 0.95 (default value), and the function would have kept continuing its search if the age 3 probability was ≥ $prob.al$, which was not the case here. For this example, the sample search for age data would then have stopped to aggregation level 2.

After having run the decision scheme for LF and age data, each stratum k was then associated with a subset of the LF (LF_k) and age (age_k) data so that all samples m retained for LF and age represented as closely as possible the associated W_k landings. Samples associated with each type of sample (LF or age) were then combined (see next two sections). It was assumed that all samples used to infer length frequencies or ALKs in the landings associated with stratum k had the same importance, regardless of the number of specimens comprising each sample (n_{km}) or of the magnitude of the landings from which each sample was collected. This is a major difference with the method shown in Gavaris and Gavaris (1983), and the absence of reported landing weights in the DFO port sampling data explains this choice (Table 3).

In a situation where the data associated with a stratum k did not provide a NAFO unit area, for example if only the division was known, the sample search was extended for the spatial component of stratum k and the NAFO division was used in the search instead of the NAFO unit area. For example, if a stratum *year-month-region-gear* was 1997-01-4Ru-OTB, where the *u* in *4Ru* stands for undetermined, the `get.samples` function would then search for samples throughout the division 4R (i.e., it would take into account data from 4Ra, 4Rb, 4Rc and 4Rd).

2.7.2. Length frequencies

The M_{LF} samples associated with each stratum were then combined. The proportion of fish of length j in sample m is given by:

$$p_{kmj} = \frac{n_{kmj}}{n_{km}} \tag{4}$$

where n_{kmj} is the number of length j fish in sample m. At the stratum level, this proportion (p_{kj}) corresponded to the average value of the M samples:

$$p_{kj} = \frac{\sum_{m=1}^{M} p_{kmj}}{M} \tag{5}$$

For a given stratum k, $\sum_j^J p_{kj}$ summed to one. The weight of a specimen of length j from sample m, ω_{kmj}, was predicted from a length-weight relationship (see section 2.5). At the stratum level, the average weight ($\overline{\omega_{kj}}$) corresponded to:

$$\overline{\omega_{kj}} = \frac{\sum_{m=1}^{M} \omega_{kmj}}{M} \tag{6}$$

2.7.3. Age-length keys

The number of aged cod of length j and age i used to infer an ALK for stratum k, n'_{kji}, was:

$$n'_{kji} = \sum_{m=1}^{M} n'_{kmji} \tag{7}$$

The number of aged cod of length j was:

$$n'_{kj} = \sum_{m=1}^{M} n'_{kmj} \tag{8}$$

The proportion of length j and age i cod in the ALK at the stratum level, p'_{kji}, was:

$$p'_{kji} = \frac{n'_{kji}}{n'_{kj}} \tag{9}$$

2.7.4. Combination of *LF* and *age* data

Ideally, for every length class present in a certain stratum k, the probability of fish therein being of a certain age is given by the ALK. However, a specific length j is sometimes observed in the length frequency data (LF_k), but not in the ALK (age_k) associated with the same stratum k. In order to infer ages of these cod, the multinom function was again used to infer ages to lengths for which there were no associated age values.

The `get.samples` function returns a dataset with for each stratum the recorded landings and the corresponding length distribution and ALK. Columns are detailed in Table 12. Due to the use of at-sea observer and of the nGSL multispecies survey data for strata associated with cod landed by shrimp trawlers, the search for samples using the `get.samples` function was carried out in two stages:

1. For strata associated to shrimp trawls. For these strata, the argument `min.al.samples` from the `get.samples` function was set to 1, since here a given sample was in fact the whole annual age dataset from the DFO August multispecies bottom trawl survey performed in the nGSL.

2. For the other strata, i.e., those where the associated gear category was not shrimp trawls.

This way, the at-sea observer and the nGSL multispecies survey data had no impact on the strata for which their use was not initially required.

2.7.5. CAA calculations

The calculations of CAA in numbers (*caan*), CAA in weight (*caaw*), mean length-at-age (*laa*) and mean weigh-at-age (*waa*) were done using the `get.caa` function:

```
get.caa(x, plus = Inf)
```

where:

`x` Output from function `get.samples`.

`plus` Value of + group (numeric). By default, it is set to infinity (Inf), so that no group+ isdefined. By setting the plus argument to 13, for example, all ages ≥ 13 would then be set to 13 ahead of the calculations, and results for age 13 would in fact be for age 13+.

The landed weight of fish of length j and age i at stratum k, W_{kji}, was calculated as:

$$W_{kji} = W_k \cdot p_{kmj} \cdot p'_{kji} \tag{10}$$

The corresponding number of specimens, N_{kji}, was calculated as:

$$N_{kji} = W_{kji} / \overline{\omega_{kj}} \tag{11}$$

Then, *caan* and *caaw* could be calculated by year as:

$$caan_i = \sum_{j}^{J} \sum_{k}^{K} N_{kji} \tag{12}$$

$$caaw_i = \sum_j^J \sum_k^K W_{kji} \quad (13)$$

For *laa* and *waa*, the calculation method corresponded to a weighted arithmetic mean of all fish lengths (*laa*) and weights (*waa*), weighted by the number of fish landed for age *i*:

$$laa_i = \sum_j^J \sum_k^K j \cdot \frac{N_{kji}}{\sum_j^J \sum_k^K N_{kji}} \quad (14)$$

$$waa_i = \frac{caaw_i}{caan_i} = \sum_j^J \sum_k^K \left(\overline{\omega_{kj}} \cdot \frac{N_{kji}}{\sum_j^J \sum_k^K N_{kji}} \right) \quad (15)$$

Their corresponding weighted standard deviation, $\sigma_{laa,i}$ and $\sigma_{waa,i}$, were calculated as:

$$\sigma_{laa,i} = \sqrt{\frac{\sum_j^J \sum_k^K N_{kji} \cdot (j - laa_i)^2}{\sum_j^J \sum_k^K N_{kji} - 1}} \quad (16)$$

$$\sigma_{waa,i} = \sqrt{\frac{\sum_j^J \sum_k^K N_{kji} \cdot \left(\overline{\omega_{kj}} - waa_i\right)^2}{\sum_j^J \sum_k^K N_{kji} - 1}} \quad (17)$$

More details on how to run the `catchR` *R* package and scripts showing how CAA calculations were performed for this study are given in Appendix A.

2.8. COMPARISON WITH PREVIOUS RESEARCH DOCUMENTS

We postulated that the practice of updating past CAA calculations during each new assessment only began recently. Consequently, we reviewed all of the 3Pn4RS Atlantic cod stock research documents published since the early 2000s (Fréchet et al. 2002, 2003, 2005, 2007, 2009; Brassard et al. 2016, 2018, 2020) to assess which research document provided the last modified value for any given year-age cell. The same was done with regard to landings data. Published values were extracted from the original pdf versions of the research documents using the *R* package `datapasta` (McBain et al. 2020).

Results obtained using the new `catchR` package were compared with the last two research documents published for the 3Pn4RS Atlantic cod stock (Brassard et al. 2018, 2020). Only the years covered by both datasets were used in comparisons.

Standardized proportions by age across years (SPAY) plots were created to see if the new approach allowed easier tracking of cohorts compared to estimated catch-at-age from the last two research documents. To do so, the proportions at age i (p_{yi}) in each year y ($y = 1,2, \ldots, Y$) were first calculated:

$$p_{yi} = \frac{caan_{yi}}{\sum_i^I caan_{yi}} \quad (18)$$

Then, p_{yi} was standardized ($SPAY_{yi}$) by subtracting the mean proportion at each age $\overline{p_i}$ and dividing by the standard deviation σ_i of the proportions computed across years:

$$\overline{p_i} = \frac{\sum_y^Y p_{yi}}{Y} \tag{19}$$

$$\sigma_i = \sqrt{\frac{\sum_y^Y \left(p_{yi} - \overline{p_i}\right)^2}{Y}} \tag{20}$$

$$SPAY_{yi} = \frac{p_{yi} - \overline{p_i}}{\sigma_i} \tag{21}$$

3. RESULTS

3.1. AVAILABLE DATA FOR CAA CALCULATIONS

Since no description of port sampling of 3Pn4RS Atlantic cod catches has been published to our knowledge since 1994 (Fréchet and Schwab 1995), a brief description of the data used is presented here. Since 1993, 2,423 LF samples were collected by DFO port samplers in the 3Pn4RS Atlantic cod fisheries, resulting in length measurements for more than 190,000 specimens (Table 13). Over that period, years 1993 and 1997 were by far the most intensely sampled years. Spatially, even though NAFO unit area 4Sw represented only 8.9% of landings reported for the 1993–2020 period, LF samples collected in this unit area represented > 20% of all LF samples collected (Tables 9, 14). The important gear in terms of cod landings has been sampled by the dockside samplers (Tables 7, 10, 15).

Overall, more than 27,000 age readings were performed over the 1993–2020 period, for a mean of almost 1,000 age readings each year (Table 16). With the exception of 2003 when no age reading was done because the stock was under a one-year moratorium and the two most recent years (2019–2020), there have always been at least 500 age readings performed annually since 1993. Similar to LF samples, most fish used for age readings originated from NAFO unit area 4Rb and were caught using either gillnets or longlines (Tables 17–18).

Only four samples were collected in 2020, with a total of 610 cod measured for length (Tables 13). Even though these results could still be preliminary, this is the smallest sample obtained since the moratorium year 2003 (Figure 6). This situation could be explained by the Covid-19 pandemic and the possible delay in adding new data to the database used for this work. According to the 1974 ICNAF[14] recommendations on sampling rates (Anon. 1974), the number of LF samples for years 2019 and 2020 should have been similar to what was observed during the 2015–2018 period. The same situation prevailed for age readings (Table 16).

3.2. `CATCHR` BEHAVIOUR

Landings from the 1993–2020 period were divided into 4,591 *year-month-NAFO-gear* strata. Of them, only 310 (6.8%) had the minimum number of LF samples (2) at the target stratum level (Figure 7A). This means that the `get.samples` function had to search for samples originating from other strata to satisfy the set threshold. The majority of strata reached their threshold at or

[14] International Commission for the Northwest Atlantic.

before aggregation level 4 (see Equation 2), which means that most of the data used to infer length frequencies originated from the specific gear category targeted, from either the specific month targeted (levels 1 and 3) or neighbouring months (levels 2 and 4) and from either the specific NAFO unit area (levels 1–2) or all of which were present for the year of interest (levels 3–4). Aggregation levels 7–12, which used neighbouring years as search criteria, were only used four times and always for strata associated to shrimp trawls.

Regarding selected age data used for each stratum, only 65 (1.4%) of strata met criteria s (see Equation 3) at the stratum level (Figure 7B). Similar to the length frequencies, the majority of strata reached their threshold at or before aggregation level 4. Aggregation levels 7–12, which used neighbouring years as search criteria, were used in 5% of strata, all of which were from the moratorium year 2003, a year when no otolith samples were collected (Table 16), or from strata associated to shrimp trawls. Age and/or length composition for most landings were calculated using samples obtained from aggregation levels ≤ 4 (Figure 8).

Each stratum had on average 18 samples to perform LF calculations (min = 2, med = 6, max = 778, Figure 7C). In terms of total specimens used, each stratum had on average 1,881 cod from which to perform calculations (min = 2, med = 880, max = 26,468, Figure 7E).

For ALKs, each stratum had on average 23 samples with which to run calculations (min = 2, med = 10, max = 522, Figure 7D). In terms of total specimens used, each stratum had on average 432 aged cod with which to perform calculations (min = 3[15], med = 260, max = 2,887, Figure 7F).

Looking at specific years, it is not surprising that for years with great coverage by the DFO port sampling program, most strata were assigned samples from a low aggregation level (Figures 9–13). Year 1993 is a good example where most of the strata had combinations of aggregation levels for LF and age data of either 3–3, 3–4 or below, respectively (Figure 9). For 2020, despite the fact that available data could still be preliminary, aggregation level 1 was not used in a single stratum for either LF_k or age_k data preparation, a first in the 1993–2020 series (Figure 13).

The correspondence between annual landings reported in the ZIFF database and what was obtained from the CAA calculations is shown in Figure 14. The mean annual difference between the two metrics is negligible, with a value of 0.01 t (min = < 0.001 t, max = 0.1 t).

Mean weight-at-age and length-at-age are shown in Figures 15 and 16, respectively. Some year-age combinations are missing, meaning that some ages were not observed in the available data during certain years (Figure 5).

3.3. COMPARISON WITH PUBLISHED VALUES

When comparing the `catchR` results with the ones presented in the last two research documents (Brassard et al. 2018, 2020), we observed that CAA numbers (*caan*), mean length-at-age (*laa*) and mean weigh-at-age (*waa*) were generally similar (Figures 17–19). However, the new approach resulted in peaks/dips for some year-age combinations which were not reported in previous research documents. Peaks/dips were also observed between the last two research documents (ex: age 12 in 2001, Figure 17).

[15] The fact that we see here a value below the one assigned to argument `min.al.fish` (100 fish) from the `get.samples` function in order to get a minimum of 100 fish to use in the ALK is generated by the combination of the LF_k and age_k datasets that ended up not using certain rows of the ALK obtained using age_k.

SPAY plots for the last two research documents and the new approach did not show much difference as far as it was possible to follow cohorts relatively well for all three data sources (Figure 20). It is worth noting, however, that research document 2019/075 (Brassard et al. 2020) did not employ an age 13 group, but rather a 13+ group.

By looking at the evolution of corrections made in each year-age cell through the different research documents published after 2000 in regard to CAA numbers (Figure 21), it can be observed that the last published research document (2019/075, Brassard et al. 2020) modified many results in a significant way back until 1989. The same observations prevailed for mean weight-at-age and length-at-age and are indicative of the difficulties with the actual tools at hand to reproduce results from the past even while working with the same data inputs (Figures 22–23).

4. DISCUSSION

4.1. DATA QUALITY AND AVAILABILITY

To our surprise, we were not able to run comparative analyses for the full series (1974–2018) currently published in the latest research document (Brassard et al. 2020). We were only able to obtain data that seemed equivalent to what was used in previous research documents starting in the early 1990s (Table 1). This same problem occurred in the February 2019 assessment when the CAA was modified to include a 13+ age group (Brassard et al. 2020, see Figure 21). Indeed, instead of running the analyses from scratch, they had to use published values from Fréchet and Schwab (1992, see their Table 6) and simply added together *caan* values of age ≥ 13 (Equation 12).

Another surprise was to find that several historical CAA values changed considerably across the different assessments published since the beginning of the 2000s (Figures 21–23). Since the published values have changed over time, it is unclear with which values to compare those obtained using `catchR`.

4.2. COMPARISONS WITH PREVIOUS PUBLISHED VALUES

Looking at the results and assuming that the postulate that both the new (`catchR`) and old (`CATCH`) approaches used the same inputs for their calculations, it is legitimate to argue that the observed differences were caused by decisions that at first glance might seem very trivial. For example, different length-weight relationships were used for the 3Pn4RS cod stock assessments performed over the years. From 1974 to 1989, the relationship used was the one provided in Minet (1978)[16]. From 1990 to 1993, two length-weight relationships calculated annually from the winter and summer research surveys were used (Fréchet et al. 1991, 1994; Fréchet and Schwab 1992; Fréchet and Gagnon 1993). For the year 1990, the relationship obtained from the winter survey conducted aboard MS[17] *Gadus Atlantica* was used for landings recorded over the January–June period, while that from the summer survey conducted aboard CCGS *Alfred Needler* was used from July to December. From 1991 to 1993, the winter survey relationship was applied to landings from January to March and October to December, while that of the summer survey was applied for the remainder of the year, like we did in the present work. However, even if a winter survey was conducted in January of 1994, only the relationship

[16] See research documents from Gascon and Fréchet (1985), Fréchet (1986), Fréchet and Gascon (1986), Fréchet (1987), Fréchet (1988), Fréchet and Schwab (1989) and Fréchet and Schwab (1990).
[17] Motor ship.

resulting from the summer survey was used in Fréchet and Schwab (1995), without further explanation.

In addition, the gear categories used are another factor that can lead to considerable differences in the results obtained. From what we understood, Fréchet and Gascon (1986) set the reference in terms of gear categories by proposing the following:

- Fixed
    - Traps (FPN-FIX)
    - Gillnets (GN)
    - Longlines (LL-TRL)
    - Handlines (LHP-LMP)
- Mobile
    - Bottom trawls (OTB)
    - Danish seines (SDN)
    - Shrimp trawls (ST)

These categories are generally the same as used in this study (see Tables 4 and 7), except that more gear types had to be classified in the present work. For strata associated to cod landings generated by fishing activities using shrimp trawls, the lack of data for this specific gear from the DFO port sampling program made it necessary to use other data sources, here those of the at-sea observer program and of the nGSL multispecies bottom trawl survey. Their use in the analyses should not have greatly modified the outcome of CAA calculations since cod landings using shrimp trawls never represented more than 1% of annual landings since 1993 (Table 7). Both these additional data sources have been used previously (Currie and Sinclair 1996).

By comparing the final results from `catchR` with what was extracted from the previous research documents (Figures 17–19), it became clear that there are differences, related to the approach used (e.g., algorithm, length-weight relationships, etc.) as well as past decisions on the association of samples with strata. Results from `catchR` generally followed well the patterns observed in the past, except for a couple of years, for unknown reasons. For example, *caan* for age 3 in 1993 was twice the value previously presented (Figure 17). When breaking down this value by stratum, we found that the large majority of these age 3 fish were associated with only two strata: *1993-July-4Ra-Traps* and *1993-July-4Rb-Traps*. Both these strata had low aggregation levels for either LF and age data[18], meaning that the data used were representative of the associated landings. Both strata used the *traps* gear category and were in the top five strata in terms of landings in 1993, so it made sense to have such a *caan* value for age 3 in that particular year. Moreover, the outputs of the `CATCH` software[19] created while preparing the last research document (Brassard et al. 2020) could be found and it appears that samples collected from trapping activities were combined with those from gillnets.

Another important difference observed between the `catchR` results and the two last research documents is the *caan* value obtained for age 3 in 2009 (> 195,000 individuals with `catchR`, 3,000 individuals in the research documents, Figure 17). An analysis of the catch by stratum did not indicate any issues with incorrect attribution of samples to strata (reasonable match, same

[18] Stratum *1993-July-4Rb-Traps* used aggregation level 1 for both LF and age data, while stratum *1993-July-4Ra-Traps* used aggregation level 3 for both LF and age data. In fact, the second stratum used the exact same data as the first one.

[19] This program only works one year at a time. Hence, one output equals 1 year.

pattern in multiple strata). Moreover, the `catchR` estimated value is not in contradiction with expected catch at age patterns for the cohort over time (2006 cohort, Figure 20).

4.3. CATCHR LIMITATIONS AND FUTURE IMPROVEMENTS

The automated algorithm, as currently defined, exposes perhaps more clearly the shortcomings in the estimation of catch- and weight-at-age, in part because the result of each year is now fully comparable and because no subjective adjustments can be applied, for example, to correct for poor sampling coverage in particular strata. For instance, using the presented algorithm, a large and probably unrealistic peak in the weight of age 13 fish in the 2010 catch was estimated (Figure 15). Abrupt changes in weight-at-age can have multiple causes. A mismatch between the LF and age data used could first be expected as the culprit. This situation could happen when age data were not acquired directly from a subsample of the specific LF data or because the length-stratification of the LF subsample for age determination was poorly performed. A simplified example of this would be a bimodal length composition with an ALK (e.g., from samples gathered at a different aggregation level) comprising only the lengths from the second mode. In such a case, the smaller fish comprising the first mode would all be classified, potentially wrongly, as the lowest observed age. This can occur when `prob.al` is set too high (so that no additional samples will be added to address this issue) or when no samples are available during any given (neighbouring) year that include the true age. Clearly, the proportion of as well as the averaged length and weights of the relevant ages will be biased for such a stratum. In this particular example, the length distribution for age 13 was based on a single individual cod measuring 98 cm. The typical average length for a 13-year-old cod is 84 cm, based on port sampling data for 1993–2020. Given the allometric relationship between length and weight, the weight of the 13-year-old fish in 2010 was clearly overestimated. Depending on the number of specimens assigned to a given age at a given year, it is therefore not surprising to be faced with significant variations in the average waa.

The quality of the CAA estimates is mostly dependent on how representative the LF_k and age_k data are of a certain landing W_k. The proposed algorithm to connect the samples with the reported landings can be extended or alternative options could be added. Specifically, we used a fixed set of aggregation levels and all samples retained were subsequently given the same weight, despite the fact that some samples might be more reflective of the catch of stratum k than others (e.g., when one sample is available from k and a second is from a differing period, gear or period, see Equation 2). Future work could focus on determining similarity between strata in terms of length frequencies and ALKs, so that the applied levels of aggregation are of optimal suitability and samples are weighted according to their representativeness. Also, it might be worth exploring other criteria for determining a sufficient sample size. A minimum of 2 samples is currently targeted, but this number might need to be adapted as a function of the catch associated with the stratum.

Furthermore, we are aware of the risks associated with using adjacent year data in length frequency and ALK calculations and the hierarchy in Equation 2 was chosen to reflect this concern. For LF samples, using LF data from neighbouring years incorrectly removes cohort dynamics that should otherwise be reflected in the CAA. For ALKs, Gulland and Rosenberg (1992, see their section 3.4.1) pointed out that mortality rates and cohort strength will vary from year to year, making the estimation of an age composition from another year's length composition inappropriate.

For this work, we wanted to use the same data source as available while using the `CATCH` software, hence the port sampling data. This work made it possible to raise questions as to how the CAA calculations were carried out for the moratorium years in particular. Since `catchR` makes it easy to assess the levels of aggregation used in the construction of the length frequency and ALK data at the level of each stratum, it may be possible to identify problematic strata (i.e., those using aggregation levels > 6, see Equation 2) in order to potentially include other data sources to better represent their landings. This approach has already been applied elsewhere (Currie and Sinclair 1996).

Unlike the classic approach proposed by Gavaris and Gavaris (1983), we assumed that all samples used to infer length frequencies or ALKs to a given stratum k had the same importance, regardless of the number of specimens comprising each sample (n_{km}) or of the magnitude of the landings from which each sample was collected. This was motivated by the absence of reported landing weights in the DFO port sampling data (Table 3). We believe that this lack of landing weight in the port sampler data is a limitation of the tools currently used to access it. In fact, it seems impossible that port samplers have omitted to take this crucial data and we therefore think that an improved query to the Oracle database could make it possible to find these missing values, and thus possibly allow improvements in the way that `catchR` does the calculations.

Several new data sources have emerged since the first CAA calculations of the 3Pn4RS cod stock. In particular, the data from the Sentinel fisheries program (fixed and mobile) could be used in particular to improve:

- Length-weight relationships. Since most cod are not landed round (Table 3), the refinement of length-weight relationships by taking into account spatial (i.e., NAFO unit area) and/or temporal (i.e. month) components could improve CAA results.
- Imputation to data poor strata. Even though landings from this program are recorded in ZIFF data, we do not think that these landings are sampled by port samplers since the fishing activities upstream of these landings are already monitored as part of the Sentinel fisheries program. Therefore, the data already exist, but is stored in another database. For the mobile survey, however, the protocol is a stratified random sampling design, which means that commercial fishermen carry out their fishing activities at locations that would not necessarily be used in a commercial fishing context. Moreover, the mobile survey uses a trawl that is not used in the commercial fishery and that has short tows, all of which can affect size and age-dependent catchability. In contrast, even though fishermen from the fixed survey do not have that much latitude in the selection of fishing sites, it remains that these sites were chosen to be on known fishing grounds and the gear used is similar to the commercial fishery. For the fixed survey, it is hence probably acceptable to use this data source. A suggestion for future years could therefore be to use this data for strata where it will be observed that the aggregation levels are considered too high (e.g., strata where the aggregation levels are > 6), similar to the approach carried out in the present work for strata where the gear category was shrimp trawls.

Finally, it is now reasonable to ask whether it could be wise to use all available at-sea observer program data for CAA calculations and not only a fraction of it to fill data gaps from the DFO port sampling program, as it was performed in this work for fishing activities involving shrimp trawls. This was not done previously because there were sufficient samples and there were concerns about the representativeness of fishing activities monitored by this program. However, a re-examination of the usefulness of these data is required for this type of analysis.

4.4. FUTURE 3PN4RS ATLANTIC COD ASSESSMENTS

For the next 3Pn4RS Atlantic cod assessments, we propose the following:

1. As it was not possible to get hold of the data allowing CAA calculations for the period 1974–1992, we simply propose to copy these data from Tables 4–6 of Brassard et al. (2020). These will therefore be fixed in time until, and if, the original data can be found.
2. For the period 1993–…, we suggest to always run the analyses over with the complete series in order to ensure that the same results are always obtained and calculations over this period remain consistent (reproducibility). Any subsequent discrepancies in results for the historical period should be investigated carefully as these could indicate unintended changes to the database.
3. The data used in constructing the CAA should be documented in the assessment document, perhaps in an appendix, to ensure reproducibility.
4. As `catchR` is destined to be improved over the years, it would be imperative to always mention the version (i.e., commit id) used for the calculations. Since the package is already on GitHub, this will allow anyone who wants to redo an analysis with an older version to do so, as long as they have the data (transparent, reproductible science). See Appendix A for an example.

5. AKNOWLEDGEMENTS

Our sincere gratitude is extended to Denis Bernier for his help with the port sampling data. We also thank Jean-Martin Chamberland and Caroline Senay for reviewing this document.

6. REFERENCES CITED

Ailloud, L.E., and Hoenig, J.M. 2019. A general theory of age-length keys: Combining the forward and inverse keys to estimate age composition from incomplete data. ICES J. Mar. Sci. 76(6): 1515–1523.

Anon. 1974. Redbook 1974. International Commission for the Northwest Atlantic Fisheries.

Anon. 1986. CAFSAC Assessment Software Catalog. CAFSAC Res. Doc. 86/96: i + 24 p.

Babyn, J., Varkey, D., Regular, P., Ings, D., and Mills Flemming, J. 2021. A gaussian field approach to generating spatial age length keys. Fish. Res. 240:105956.

Berg, C. W., and Kristensen, K. 2012. Spatial age-length key modelling using continuation ratio logits. Fish. Res. 129–130: 119–126.

Bourdages, H., and Ouellet, J.-F. 2011. Geographic distribution and abundance indices of marine fish in the northern Gulf of St. Lawrence (1990–2009). Can. Tech. Rep. Fish. Aquat. Sci. 2963:vi + 171 p.

Bourdages, H., Brassard, C., Desgagnés, M., Galbraith, P., Gauthier, J., Nozères, C., Scallon-Chouinard, P.-M., and Senay, C. 2020. Preliminary results from the ecosystemic survey in August 2019 in the Estuary and northern Gulf of St. Lawrence. DFO Can Stock. Assess. Sec. Res. Doc. 2020/009: iv + 93 p.

Brassard, C., Gauthier, J., Schwab, P., Le Bris, A., Way, M., and Collier, F. 2016. The status of the Northern Gulf of St. Lawrence (3Pn, 4RS) cod (*Gadus morhua*) stock in 2014. DFO Can. Sci. Advis. Sec. Res. Doc. 2016/010: xi + 120 p.

Brassard, C., Gauthier, J., Lussier, J.-F., Way, M., and Collier, F. 2018. The status of the Northern Gulf of St. Lawrence (3Pn, 4RS) cod stock (*Gadus morhua*) in 2016. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/039: xi + 116 p.

Brassard, C., Lussier, J.-F., Benoît, H., Way, M., and Collier, F. 2020. The status of the northern Gulf of St. Lawrence (3Pn, 4RS) Atlantic cod (*Gadus morhua*) stock in 2018. DFO Can. Sci. Advis. Sec. Res. Doc. 2019/075: x + 117 p.

Currie, L.G., and Sinclair, A. F. 1996. Catch-at-age for southern gulf of St. Lawrence cod: 1971 to 1995. Can. Manuscr. Rep. Fish. Aquat. Sci. 2389:iii + 80 p.

Daigle, D., and Benoît, H.P. 2007. Procedures for commercial catch sampling of finfish and shrimp in the southern Gulf of St. Lawrence (Fisheries and Oceans Canada, Gulf Region). Can. Manuscr. Rep. Fish. Aquat. Sci. 2833: iv + 63 p.

Doniol-Valcroze, T., Van Beveren, E., Légaré, B., Girard, L., and Castonguay, M. 2019. Atlantic mackerel (*Scomber scombrus* L.) in NAFO Subareas 3 and 4 in 2016. DFO Can. Sci. Advis. Sec. Res. Doc. 2018/062: v + 51 p.

Fréchet, A. 1986. Assessment of the 4RS, 3Pn cod stock. CAFSAC Res. Doc. 86/105: 62 p.

Fréchet, A. 1987. Assessment of the northeastern Gulf of St. Lawrence (Subdiv. 3Pn and Divs. 4RS) cod stock – 1986. CAFSAC Res. Doc. 87/65: 50 p.

Fréchet, A. 1988. Évaluation du stock de morue de la subdivision 3Pn et des divisions 4R et 4S de l'OPANO en 1987. CSCPCA Doc. de Rech. 88/24:66 p.

Fréchet, A., and Gagnon, Y. 1993. Assessment of the biomass of the northern gulf of St. Lawrence cod stock (3Pn, 4RS) in 1992. DFO Atlantic Fisheries Research Document 93/68:42 p.

Fréchet, A., and Gascon, D. 1986. Une nouvelle matrice de capture a l'âge pour le stock de morue des divisions 4R, 4S et 3Pn. CSCPCA Doc. de Rech. 86/104: 34 p.

Fréchet, A., and Schwab, P. 1989. Évaluation du stock de morue de 3Pn, 4RS en 1988. CSCPCA Doc. de Rech. 89/55: 62 p.

Fréchet, A., and Schwab, P. 1990. La pêche de la morue du nord-ouest du golfe du St. Laurent en 1989 (3Pn, 4RS), l'analyse des prises, de l'effort de pêche et l'estimation de la biomasse. CSCPCA Doc. de Rech. 90/60:53 p.

Fréchet, A., and Schwab, P. 1992. Assessment of the northern gulf of St. Lawrence cod stock (3Pn, 4RS) in 1991. CAFSAC Res. Doc. 92/77:50 p.

Fréchet, A., and Schwab, P. 1995. Review of the status of the northern gulf of St. Lawrence cod stock (3Pn, 4RS) in 1994. DFO Atlantic Fisheries Research Document 95/53:20 p.

Fréchet, A., and Schwab, P. 1998. Assessment of the northern Gulf of St. Lawrence cod stock (3Pn, 4RS) in 1997. DFO Can. Sci. Advis. Sec. Res. Doc. 98/127: 57 p.

Fréchet, A., Schwab, P., and Gagnon, P. 1991. Évaluation du stock de morue du nord du golfe du Saint-Laurent (Divisions de l'OPANO 3Pn, 4R et 4S) en 1990. CSCPCA Doc. de Rech. 91/43: 55 p.

Fréchet, A., Gagnon, Y., Schwab, P., D'Amours, D., Dutil, J.-D., Lambert, Y., Bourassa, L., and Rollet, C. 1994. Review of the status of the northern gulf of St. Lawrence cod stock (3Pn, 4RS) in 1993. DFO Atl. Fish. Res. Doc. 94/29:50 p.

Fréchet, A., Gauthier, J., Schwab, P., Moreault, G., Pageau, L., Spingle, J., and Collier, F. 2002. The status of cod in the Northern Gulf of St. Lawrence (3Pn, 4RS) in 2001. DFO Can. Sci. Advis. Sec. Res. Doc. 2002/083: 54 p.

Fréchet, A., Gauthier, J., Schwab, P., Bourdages, H., Chabot, D., Collier, F., Grégoire, F., Lambert, Y., Moreault, G., Pageau, L., and Spingle, J. 2003. The status of cod in the Northern Gulf of St. Lawrence (3Pn, 4RS) in 2002. DFO Can. Sci. Advis. Sec. Res. Doc. 2003/065: i + 25 p.

Fréchet, A., Gauthier, J., Schwab, P., Pageau, L., Savenkoff, C., Castonguay, M., Chabot, D., Tournois, C., Lussier, J.-F., Spingle, J., and Collier, F. 2005. The status of cod in the Northern Gulf of St. Lawrence (3Pn, 4RS) in 2004. DFO Can. Sci. Advis. Sec. Res. Doc. 2005/060: i + 71 p.

Fréchet, A., Gauthier, J., Schwab, P., Bourdages, H., Tournois, C., Spingle, J., Way, M., and Collier, F. 2007. The status of cod in the Northern Gulf of St. Lawrence (3Pn, 4RS) in 2006. DFO Can. Sci. Advis. Sec. Res. Doc. 2007/068: iii + 85 p.

Fréchet, A., Gauthier, J., Schwab, P., Lambert, Y., Le Bris, A., Tournois, C., Way, M., and Collier, F. 2009. The status of cod in the Northern Gulf of St. Lawrence (3Pn, 4RS) in 2008. DFO Can. Sci. Advis. Sec. Res. Doc. 2009/090: iv + 104 p.

Fridriksson, A. 1934. On the calculation of age-distribution within a stock of cod by means of relatively few age-determinations as a key to measurements on a large scale. Rapp. et Proc.-Verb. Cons. Int. Explor. Mer 86(6): 1–14.

Gascon, D., and Fréchet, A. 1985. An assessment of the cod stock in NAFO Divisions 4R 4S and in Subdivision 3Pn. CAFSAC Res. Doc. 85/107: 48 p.

Gavaris, S., and Gavaris, C.A. 1983. Estimation of catch at age and its variance for groundfish stocks in the Newfoundland region. *In* Sampling commercial catches of marine fishes and invertebrates. *Edited by* W. G. Doubleday and D. Rivard. pp. 178–182.

Gulland, J. A., and Rosenberg, A.A. 1992. A review of length-based approaches to assessing fish stocks. FAO Fisheries Technical Paper 323:100 p.

Ketchen, K.S. 1950. Stratified subsampling for determining age distributions. Trans. Am. Fish. Soc. 79(1): 205–212.

Kulka, D.W. 1981. An examination of the product form to whole weight conversion used by the foreign fleets in 1980 (Nfld. Observer Program). CAFSAC Res. Doc. 81/60:38 p.

Lambert, J.-D., and Ménager, M. 1998. Protocoles d'échantillonnage des captures commerciales de poissons et d'invertébrés marins du golfe du Saint-Laurent. Rapp. tech. can. sci. halieut. aquat. 2208: x + 246 p.

McBain, M., Carroll, J., Gelfand, S., Owlarn, S., and Aden-Buie, G. 2020. datapasta: R Tools for Data Copy-Pasta. R package version 3.1.0.

Minet, J.P. 1978. Dynamics and yield assessment of the northeastern gulf of St. Lawrence cod stock. ICNAF Sel. Pap. 3:7–16.

R Core Team. 2020. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Sanguin, A.-L. 1980. La zone canadienne des 200 milles dans l'Atlantique, un exemple de la nouvelle géographie politique des océans. Études internationales 11(2): 239–251.

Savard, L., Gauthier, J., Bourdages, H., and Desgagnés, M. 2013. Bycatch in the Estuary and Gulf of St. Lawrence Northern shrimp fishery. DFO Can. Sci. Advis. Sec. Res. Doc. 2012/151:ii + 56 p.

STACAC. 1984a. Standard conversion factors: All species. STACAC – Standards Document 2:7 p.

STACAC. 1984b. Catch & effort computer file structure for zonal interchange - 1984. STACAC Standard Document 3:13 p.

Venables, W. N., and Ripley, B.D. 2002. Modern Applied Statistics with S. Springer, New York.

Wells, R. 1979. Status of the cod stock in Div 4RS 3Pn. CAFSAC Res. Doc. 79/23:11 p.

7. TABLES

Table 1. Annual commercial sampling inventory (number of cod sampled) for the 3Pn4RS Atlantic cod stock as reported in various research documents, compared to what the presented extraction obtained. Commercial sampling information has ceased to be published in research documents from 1995 onwards.

Year	LF		Age		Source
	Published	Available	Published	Available	
1974	11,783	-	1,509	-	Fréchet and Gascon (1986)
1975	13,243	-	1,345	-	Fréchet and Gascon (1986)
1976	18,008	-	1,574	-	Fréchet and Gascon (1986)
1977	20,221	-	3,283	-	Fréchet and Gascon (1986)
1978	17,688	-	2,149	-	Fréchet and Gascon (1986)
1979	73,087	-	3,724	-	Fréchet and Gascon (1986)
1980	78,063	-	4,270	-	Fréchet and Gascon (1986)
1981	79,839	291	5,584	21	Fréchet and Gascon (1986)
1982	58,525	-	5,600	-	Fréchet and Gascon (1986)
1983	75,247	76,379	5,821	5,861	Fréchet and Gascon (1986)
1984	104,703	92,173	8,272	7,947	Fréchet and Gascon (1986)
1985	79,052	62,503	8,357	7,655	Fréchet (1986)
1986	84,174	67,053	6,733	6,334	Fréchet (1987)
1987	62,944	53,175	6,574	6,169	Fréchet (1988)
1988	50193	49,417	4,944	4,622	Fréchet and Schwab (1989)
1989	76,315	63,314	7,856	7,638	Fréchet and Schwab (1990)
1990	76,476	41,504	8,054	5,428	Fréchet et al. (1991)
1991	138,614	38,844	4,507	3,491	Fréchet and Schwab (1992)
1992	116,910	53,793	2,868	2,633	Fréchet and Gagnon (1993)
1993	41,177	41,451	2,906	2,906	Fréchet et al. (1994)
1994	1,387	2,151	180	913	Fréchet and Schwab (1995)

Table 2. Annual 3Pn4RS cod landings reported by country/entity other than Canada since 1960. Source: NAFO 21B data.

Year	Landings (t)	Country/entity*
1960	49,871	1, 2, 3, 4, 5, 6
1961	61,784	1, 3, 5
1962	39,259	1, 3, 5
1963	25,006	1, 3, 5, 7
1964	42,251	1, 3, 4, 5, 6, 7, 8, 9
1965	31,481	1, 3, 4, 5, 6
1966	28,327	1, 3, 4, 5, 6, 8, 10
1967	41,948	1, 3, 4, 5, 6, 7
1968	39,102	1, 3, 4, 5, 6
1969	20,703	1, 3, 4, 5
1970	57,243	1, 3, 4, 5, 9
1971	48,474	1, 3, 4, 5, 9
1972	27,385	1, 3, 4, 5, 6, 11
1973	37,236	1, 3, 4, 5, 11
1974	32,265	1, 3, 4, 5, 6, 11, 12
1975	31,644	1, 3, 4, 11
1976	34,275	1, 3, 4, 11
1977	18,138	1, 4
1978	15,771	1, 4
1979	13,769	1, 4
1980	9,396	1, 4
1981	12,508	1, 4
1982	12,013	1, 4
1983	10,684	1, 4
1984	11,623	1, 4
1985	9,185	1, 4
1986	13,122	1, 4
1987	1,535	4
1989	2,587	4
1990	2,485	4
1991	2,447	4
1992	2,333	4

* 1 = France Mainland, 2 = Italy, 3 = Portugal,
4 = Saint Pierre and Miquelon, 5 = Spain,
6 = United Kingdom, 7 = Soviet Union,
8 = Iceland, 9 = United States of America,
10 = Poland, 11 = Faroe Islands, 12 = Norway

Table 3. Inventory of length frequency samples according to the state at landing and the year of collection, for 3Pn4RS cod. The annual number of samples where either the sample weight or the landing weight is not reported are provided. Source: DFO port sampling program data.

	State at landing				Unweighted samples		Unweighted landings	
Year	**Whole**	**Gutted head on**	**Unknown**	**Total**	**Nb**	**% of total**	**Nb**	**% of total**
1993	24	145	-	169	15	8.9	1	0.6
1994	-	6	72	78	72	92.3	72	92.3
1995	2	4	772	778	772	99.2	774	99.5
1996	4	7	407	418	407	97.4	407	97.4
1997	74	82	227	383	299	78.1	298	77.8
1998	-	21	-	21	8	38.1	-	-
1999	2	40	-	42	2	4.8	1	2.4
2000	7	29	-	36	3	8.3	1	2.8
2001	-	36	-	36	-	-	-	-
2002	-	28	-	28	3	10.7	-	-
2003	-	3	-	3	-	-	-	-
2004	3	21	-	24	3	12.5	-	-
2005	29	5	-	34	4	11.8	1	2.9
2006	28	14	-	42	5	11.9	1	2.4
2007	9	34	-	43	4	9.3	3	7.0
2008	8	24	-	32	5	15.6	-	-
2009	-	31	-	31	6	19.4	1	3.2
2010	-	28	-	28	4	14.3	-	-
2011	-	26	-	26	6	23.1	1	3.8
2012	1	19	-	20	4	20.0	-	-
2013	7	12	-	19	8	42.1	1	5.3
2014	5	12	-	17	-	-	-	-
2015	2	20	-	22	8	36.4	2	9.1
2016	9	16	-	25	4	16.0	-	-
2017	6	23	-	29	5	17.2	-	-
2018	9	16	-	25	5	20.0	-	-
2019	-	10	-	10	-	-	-	-
2020	-	4	-	4	-	-	-	-
Total	**229**	**716**	**1,478**	**2,423**	**1,652**	**68.2**	**1,564**	**64.5**

Table 4. Gear codes used in port sampling data. Only those that generated 3Pn4RS Atlantic cod landings during the 1993–2020 period are shown. The category variable shows how the different gears were grouped together for the analyzes.

Category	Code	Description
Shrimp trawls[1]	GRL	Shrimp trawl (unspecified) with grid
	GRL1	Shrimp trawl (side) with grid
	GRL2	Shrimp trawl (stern) with grid
	OTB2	Stern otter trawl
	TT	Twin trawl
Other trawls	OTB2	Stern otter trawl
	OTM	Midwater trawl
Seines	PS	Purse seine
	SDN	Danish or Scottish or Pair seine (2 boats)
Gillnets	GN	Gillnets (not specified)
	GNS	Set gillnets (anchored)
Longlines	LL	Longlines (not specified)
	LLS	Set longlines
Hand gear	LHP	Handlines and pole-lines (hand-operated)
	LLG	Handline
	LX	Hooks and lines (not specified)
Traps	FIX	Traps (not specified)
	FPN	Stationary uncovered pound nets

[1] Since no sampling was done by port samplers for cod caught using shrimp trawls, the at-sea observer program database was used in order to get length frequency data. Age data from the nGSL multispecies bottom trawl survey database were used to produce annual ALKs.

Table 5. Details of the length-weight relationships (log_e ($mass_{yp}$ in kg) ~ $\alpha + \beta \cdot log_e$ ($length_{yp}$ in cm)) used for each available year-month combination. The number of cod initially available and used to estimate the coefficients is provided, as is the data source (SS = summer survey, WS = winter survey).

		Number of cod					
Year	**Month**	**Available**	**Used**	**α**	**β**	**R^2**	**Source**
1993	1–3, 10–12	1097	1000	-12.193	3.127	0.995	WS 1993
1993	4–9	477	449	-12.003	3.095	0.990	SS 1993
1994	1–3, 10–12	1600	1483	-12.123	3.107	0.994	WS 1994
1994	4–9	890	841	-11.892	3.058	0.993	SS 1994
1995	1–12	493	470	-11.854	3.066	0.986	SS 1995
1996	1–12	1009	928	-11.764	3.034	0.987	SS 1996
1997	1–12	1046	998	-11.775	3.043	0.987	SS 1997
1998	1–12	1016	973	-11.719	3.021	0.989	SS 1998
1999	1–12	1016	962	-11.692	3.013	0.992	SS 1999
2000	1–12	1055	1012	-11.985	3.090	0.994	SS 2000
2001	1–12	1055	988	-11.716	3.019	0.993	SS 2001
2002	1–12	577	548	-11.762	3.048	0.989	SS 2002
2003	1–12	1450	1353	-12.001	3.103	0.993	SS 2003
2004	1–12	2312	2067	-12.233	3.166	0.995	SS 2004
2005	1–12	4745	4151	-12.131	3.139	0.997	SS 2005
2006	1–12	2498	2218	-12.081	3.114	0.996	SS 2006
2007	1–12	3034	2698	-12.042	3.105	0.997	SS 2007
2008	1–12	3723	3329	-12.059	3.111	0.994	SS 2008
2009	1–12	2249	1975	-12.088	3.118	0.995	SS 2009
2010	1–12	1750	1531	-12.020	3.101	0.996	SS 2010
2011	1–12	1821	1584	-12.027	3.098	0.996	SS 2011
2012	1–12	2435	2182	-12.034	3.092	0.997	SS 2012
2013	1–12	1954	1684	-12.029	3.097	0.996	SS 2013
2014	1–12	2574	2246	-11.997	3.089	0.996	SS 2014
2015	1–12	2625	2334	-11.966	3.077	0.996	SS 2015
2016	1–12	2113	1867	-11.916	3.060	0.996	SS 2016
2017	1–12	2099	1815	-11.824	3.032	0.996	SS 2017
2018	1–12	1837	1614	-12.151	3.139	0.998	SS 2018
2019	1–12	1195	1044	-11.998	3.091	0.997	SS 2019
2020	1–12	1801	1645	-12.069	3.118	0.996	SS 2020

Table 6. Historical monthly landings statistics (t) for 3Pn4RS cod for the period 1964–2020. Until 1998, the management year corresponded to the calendar year. For 1999, landings included those from January 1, 1999 to May 14, 2000. Since then, the management year corresponds to the period going from May 15 to May 14 of the following year. Unk. = unknown. '0' values indicate landings ≤ 0.5 t. Source: NAFO 21B and ZIFF data.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Unk.	Total
1964	1,104	24,423	15,760	6,059	3,106	10,349	12,526	5,853	2,154	1,385	864	651	-	84,234
1965	791	12,577	21,171	3,698	2,146	5,267	10,421	5,945	3,636	1,359	927	990	-	68,928
1966	1,965	22,817	8,929	2,516	1,638	8,371	7,483	4,740	2,493	1,146	1,779	1,208	-	65,085
1967	7,873	7,028	14,792	8,448	2,017	7,524	12,665	5,232	7,154	3,314	1,352	1,912	1	79,312
1968	725	7,980	22,799	9,060	3,087	10,719	17,214	9,400	4,913	1,784	1,171	819	-	89,671
1969	875	4,654	9,675	4,220	5,192	10,958	12,103	8,639	7,866	3,557	2,035	1,366	-	71,140
1970	1,635	25,494	18,223	27,886	4,816	6,017	8,963	3,896	2,184	3,114	1,937	1,300	-	105,465
1971	845	44,587	7,580	5,265	2,346	5,857	8,427	3,042	2,343	1,600	1,003	915	-	83,810
1972	1,494	14,961	5,337	7,400	7,334	4,594	6,818	3,296	2,365	1,406	994	212	2,026	58,237
1973	16,472	10,556	7,586	4,826	3,235	5,860	5,125	4,145	2,365	1,459	1,016	567	2,593	65,805
1974	12,995	10,753	5,959	5,665	6,231	5,021	6,235	5,396	2,214	1,331	1,009	479	3,148	66,436
1975	8,232	19,486	2,702	2,616	5,316	5,122	5,042	4,488	2,767	1,267	819	704	1,672	60,233
1976	15,637	15,204	3,610	3,437	7,071	6,930	6,978	4,310	3,348	2,286	1,537	578	6,055	76,981
1977	11,143	8,603	3,790	11,312	10,057	7,368	8,133	5,780	3,361	1,751	1,814	454	-	73,566
1978	20,754	6,307	5,161	3,156	6,717	9,796	13,255	7,000	2,836	1,979	1,309	236	-	78,506
1979	15,543	4,273	6,475	6,647	8,517	12,890	12,085	8,660	2,971	2,449	1,816	451	-	82,777
1980	5,280	8,965	9,925	8,087	7,147	14,096	23,158	10,719	5,687	2,773	1,311	431	-	97,579
1981	9,156	15,368	3,170	3,763	12,835	17,257	16,344	10,343	5,676	2,550	1,172	277	-	97,911
1982	2,289	11,671	10,122	5,544	12,723	16,826	22,492	9,136	8,412	4,465	1,227	32	-	104,939
1983	4,152	10,213	11,335	6,251	21,049	18,341	16,228	8,173	5,698	3,956	530	154	-	106,080
1984	5,002	11,079	9,494	4,260	15,205	13,349	22,300	10,962	5,238	4,644	1,113	997	-	103,643
1985	2,416	16,369	7,661	3,407	6,904	12,612	13,874	11,414	7,730	3,130	1,005	1,959	-	88,481
1986	2,468	18,021	10,611	4,847	12,057	7,613	12,739	5,960	4,348	2,956	834	944	-	83,399
1987	8,264	7,382	5,072	3,945	6,411	8,222	9,060	7,492	5,745	2,842	1,022	1,089	-	66,545
1988	1,505	2,710	4,270	2,697	9,897	4,971	7,679	6,282	3,264	1,747	1,143	1,536	-	47,702
1989	6,198	7,511	1,982	2,048	6,520	6,229	6,306	4,797	2,080	2,189	721	181	-	46,762
1990	5,646	2,537	1,102	394	7,953	7,741	4,664	3,122	1,968	1,554	1,856	464	-	39,000
1991	1,532	2,001	3,113	3,736	4,229	4,477	5,314	2,891	3,242	2,016	1,810	121	-	34,481
1992	4,453	2,551	226	1,825	4,696	1,729	3,211	3,538	2,316	1,869	1,868	1,261	-	29,546
1993	9	51	1,255	1,244	1,489	4,350	3,811	2,234	1,119	1,088	1,173	629	-	18,452
1994	14	48	41	7	26	12	14	100	206	28	24	18	-	537
1995	-	-	-	0	12	5	26	95	25	21	-	-	-	185
1996	0	0	0	0	5	10	150	56	38	33	23	2	-	317
1997	0	1	-	2	357	255	1,189	962	815	1,038	145	27	-	4,792
1998	3	0	0	2	27	246	908	1,051	418	552	22	-	-	3,229

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Unk.	Total
1999/2000	1	51	132	107	106	870	1,985	1,458	1,031	1,014	395	39	-	7,191
2000/2001	86	72	49	33	561	907	1,251	1,533	1,087	775	398	82	-	6,833
2001/2002[1]	110	58	6	10	447	518	1,847	1,269	1,339	865	293	125	-	6,886
2002/2003[2]	0	-	0	0	146	58	2,519	1,484	843	869	393	7	-	6,320
2003/2004	0	-	0	13	8	13	118	131	48	31	39	5	-	405
2004/2005	0	-	0	14	30	25	1,887	205	537	356	207	13	-	3,274
2005/2006	-	-	0	24	44	69	2,434	628	774	473	22	3	-	4,471
2006/2007[3]	1	0	0	15	19	101	3,285	591	645	298	669	17	-	5,640
2007/2008	-	-	1	7	22	132	3,711	447	1,126	578	447	5	-	6,474
2008/2009[4]	-	-	1	3	45	117	2,973	924	1,240	551	301	2	-	6,157
2009/2010	1	0	2	2	7	176	1,691	693	690	696	687	51	-	4,696
2010/2011	0	0	0	15	10	54	1,362	882	556	499	185	1	-	3,566
2011/2012	0	0	0	12	15	50	1,006	163	315	11	193	9	-	1,773
2012/2013	0	0	0	22	16	40	671	110	296	20	131	3	-	1,310
2013/2014	-	-	-	7	11	34	699	77	220	10	147	3	-	1,208
2014/2015	-	-	-	4	18	16	644	92	344	26	100	23	-	1,266
2015/2016	-	-	-	10	12	19	731	138	187	18	137	13	-	1,264
2016/2017	-	0	0	23	9	27	840	140	156	30	154	7	-	1,387
2017/2018	-	1	7	12	16	30	903	706	637	24	206	129	-	2,672
2018/2019	-	0	0	10	18	19	861	774	323	221	268	74	-	2,570
2019/2020[5]	1	0	-	1	15	23	304	97	168	27	87	38	-	761
2020/2021[5]	-	-	-	-	5	17	382	57	63	123	8	2	-	657

[1] Excluded 253 t from recreational fishing. See Frechet et al. (2003).
[2] Excluded 34 t from recreational fishing. See Frechet et al. (2003).
[3] Excluded 75.3 t from recreational fishing. See DFO (2008).
[4] Excluded 67 t from recreational fishing. See Frechet et al. (2009).
[5] Preliminary data.

Table 7. Gear codes used in the ZIFF landing data for the 1993–2020 period. Only those that generated 3Pn4RS Atlantic cod landings during this period are shown. The category variable shows the categories of gear used for the analyzes. The importance of each gear in total landings since 1993 (Tot), and the annual variation of their input in landings during this period (Min/Max) are shown.

	Gear		% of landings		
Category	**Code**	**Description**	**Tot**	**Min**	**Max**
Shrimp trawls	19	Shrimp trawl	< 0.1	0	0.9
Other trawls	11	Bottom otter trawl (side)	< 0.1	0	0.1
	12	Bottom otter trawl (stern)	9.4	< 0.1	51.1
	13	Midwater trawl (unspecified)	< 0.1	0	4.6
	15	Midwater trawl (stern)	0.6	0	13.8
	16	Bottom pair trawl	< 0.1	0	< 0.1
	18	Semi-midwater trawl	< 0.1	0	1.4
Seines	21	Danish seine	0.7	0.2	6.4
	22	Scottish seine	< 0.1	0	< 0.1
Gillnets	41	Gillnet (set or fixed)	39.4	4.3	69.5
Longlines	50	Setheared hooks	0.7	0	3.8
	51	Longline	39.5	11.5	77.3
Hand gear	53	Jigger	0.5	0	28.1
	55	Mechanized squid jigger	< 0.1	0	< 0.1
	59	Hand line (baited)	6.4	0.1	14.5
Traps	61	Trap net	2.5	0	13.3
	62	Pot	< 0.1	0	< 0.1
	67	Rectangular trap	< 0.1	0	< 0.1
	68	Conical trap	< 0.1	0	< 0.1
Others / unknown	71	Dredge (boat)	< 0.1	0	< 0.1
	99	Unspecified fixed gear	0.1	0	47.4

Table 8. Reported landings (t) of the 3Pn4RS Atlantic cod stock in the commercial fishery for the period 1993–2020, by year and month. Source: ZIFF data.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
1993	9	51	1,255	1,244	1,489	4,350	3,811	2,234	1,119	1,088	1,173	629	18,452
1994	14	48	41	7	26	12	14	100	206	28	24	18	537
1995	-	-	-	0	12	5	26	95	25	21	-	-	185
1996	0	0	0	0	5	10	150	56	38	33	23	2	317
1997	0	1	-	2	357	255	1,189	962	815	1,038	145	27	4,792
1998	3	0	0	2	27	246	908	1,051	418	552	22	-	3,229
1999	0	0	0	5	92	870	1,985	1,458	1,031	1,014	395	39	6,890
2000	1	51	132	102	538	907	1,251	1,533	1,087	775	398	82	6,857
2001	86	72	49	33	455	518	1,847	1,269	1,339	865	293	125	6,950
2002	110	58	6	10	173	58	2,519	1,484	843	869	393	7	6,529
2003	0	-	0	0	5	13	118	131	48	31	39	5	389
2004	0	-	0	13	30	25	1,887	205	537	356	207	13	3,274
2005	0	-	0	14	43	69	2,434	628	774	473	22	3	4,461
2006	-	-	0	24	23	101	3,285	591	645	298	669	17	5,652
2007	1	0	0	15	22	132	3,711	447	1,126	578	447	5	6,483
2008	-	-	1	7	40	117	2,973	924	1,240	551	301	2	6,158
2009	-	-	1	3	13	176	1,691	693	690	696	687	51	4,700
2010	1	0	2	2	6	54	1,362	882	556	499	185	1	3,552
2011	0	0	0	15	14	50	1,006	163	315	11	193	9	1,775
2012	0	0	0	12	19	40	671	110	296	20	131	3	1,303
2013	0	0	0	22	10	34	699	77	220	10	147	3	1,222
2014	-	-	-	7	10	16	644	92	344	26	100	23	1,262
2015	-	-	-	4	21	19	731	138	187	18	137	13	1,267
2016	-	-	-	10	9	27	840	140	156	30	154	7	1,373
2017	-	0	0	23	12	30	903	706	637	24	206	129	2,671
2018	-	1	7	12	18	19	861	774	323	221	268	74	2,578
2019	-	0	0	10	20	23	304	97	168	27	87	38	775
2020	1	0	-	1	9	17	382	57	63	123	8	2	662
Total	**227**	**283**	**1,494**	**1,601**	**3,496**	**8,195**	**38,200**	**17,096**	**15,245**	**10,277**	**6,853**	**1,328**	**104,294**

Table 9. Reported landings (t) of the 3Pn4RS Atlantic cod stock in the commercial fishery for the period 1993–2020, by year and NAFO unit area. Source: ZIFF data.

Year	3Pn	4Ra	4Rb	4Rc	4Rd	4Ru	4Si	4Ss	4Su	4Sv	4Sw	4Sx	4Sy	4Sz	Total
1993	3,194	4,071	5,679	2,021	1,723	-	53	340	-	342	654	39	88	249	18,452
1994	196	67	44	60	114	-	2	5	1	1	17	22	6	3	537
1995	35	3	17	5	11	-	0	2	-	62	42	2	1	4	185
1996	72	13	22	13	31	-	1	5	0	33	123	0	0	2	317
1997	2,006	806	600	593	299	-	0	7	-	141	327	7	1	4	4,792
1998	877	387	367	316	637	-	0	13	-	77	526	27	0	3	3,229
1999	1,165	1,551	1,478	908	944	-	1	21	1	124	632	44	20	2	6,890
2000	1,478	1,215	1,469	772	865	0	1	52	0	193	714	83	13	1	6,857
2001	1,740	1,311	1,267	992	722	17	0	23	18	252	570	26	12	1	6,950
2002	1,713	1,172	1,377	795	591	3	1	28	2	123	686	31	7	0	6,529
2003	86	37	81	21	71	0	1	6	-	20	60	4	1	1	389
2004	783	595	642	297	350	1	0	10	3	98	442	43	11	1	3,274
2005	856	976	701	437	770	68	0	11	50	278	293	7	11	3	4,461
2006	1,208	1,196	680	434	1,168	193	0	12	2	297	427	11	20	2	5,652
2007	1,081	1,575	939	748	914	369	1	14	9	146	668	11	5	5	6,483
2008	1,131	1,704	973	665	829	2	1	15	3	194	610	15	14	2	6,158
2009	1,357	922	799	424	593	-	3	15	-	183	380	15	8	2	4,700
2010	705	1,135	546	270	224	1	4	14	0	276	346	13	14	4	3,552
2011	316	510	188	143	82	0	13	18	-	131	239	123	11	3	1,775
2012	187	372	188	105	67	-	8	9	-	74	181	104	5	4	1,303
2013	185	366	145	151	99	0	8	6	-	56	163	33	6	4	1,222
2014	153	492	138	105	48	-	12	11	0	81	167	42	2	10	1,262
2015	155	422	141	85	58	-	9	19	0	122	197	31	3	26	1,267
2016	174	510	130	79	81	-	9	6	-	115	150	86	4	30	1,373
2017	345	1,202	298	230	135	-	82	15	0	69	194	40	1	60	2,671
2018	439	1,239	206	109	89	-	40	5	-	91	247	31	1	80	2,578
2019	154	249	68	20	40	-	42	3	1	22	104	32	1	37	775
2020	119	183	92	46	57	-	6	13	-	32	91	13	1	10	662
Total	**21,910**	**24,281**	**19,274**	**10,843**	**11,611**	**654**	**298**	**699**	**90**	**3,632**	**9,250**	**932**	**267**	**552**	**104,294**

Table 10. Reported landings (t) of the 3Pn4RS Atlantic cod stock in the commercial fishery for the period 1993–2020, by year and gear category. Source: ZIFF data.

Year	Shrimp trawls	Other trawls	Seines	Gillnets	Longlines	Hand gear	Traps	Others / unknown	Total
1993	83	10,054	190	2,501	2,124	1,041	2,458	-	18,452
1994	5	154	9	23	98	229	-	19	537
1995	0	1	4	48	42	1	0	88	185
1996	0	4	11	193	97	0	4	6	317
1997	1	59	8	538	3,857	265	57	7	4,792
1998	0	5	30	629	2,302	261	2	-	3,229
1999	0	11	35	3,401	2,475	968	0	-	6,890
2000	0	24	36	3,046	3,167	557	26	0	6,857
2001	0	26	64	2,020	3,963	864	13	-	6,950
2002	0	34	30	2,672	3,155	634	5	0	6,529
2003	0	10	25	161	191	1	1	-	389
2004	0	29	44	1,482	1,408	310	0	-	3,274
2005	0	35	50	2,285	1,855	229	6	-	4,461
2006	0	84	38	2,685	2,655	189	1	-	5,652
2007	0	0	31	3,430	2,868	154	-	-	6,483
2008	0	2	25	3,027	2,889	215	0	-	6,158
2009	0	2	23	1,852	2,615	207	2	-	4,700
2010	0	3	13	1,710	1,563	263	0	-	3,552
2011	4	1	12	1,004	671	84	1	-	1,775
2012	0	2	4	820	403	74	-	-	1,303
2013	1	1	2	670	415	131	1	-	1,222
2014	0	2	3	785	342	129	0	-	1,262
2015	0	1	4	872	342	48	-	-	1,267
2016	0	3	5	866	433	64	1	-	1,373
2017	0	3	6	1,747	744	171	-	-	2,671
2018	0	14	7	1,793	711	54	0	-	2,578
2019	0	9	3	442	283	38	-	-	775
2020	0	3	1	422	207	28	-	-	662
Total	**95**	**10,576**	**713**	**41,124**	**41,876**	**7,211**	**2,578**	**120**	**104,294**

Table 11. Illustration of the procedure to search for age samples using the function `get.samples`. The stratum used as an example is 2005–June–4Rb–Gillnets. The results of the search for samples after having completed the search at aggregation level 2 are shown. Only a few lengths (cm) are provided (the 8 smallest and the largest). For each length, the number of specimens used for length frequencies (LF) and age readings are specified. The age columns in green (ages 4–14) are those already present in the ALK at aggregation level 2. The age columns in yellow (ages 3 and 15) are those that are still missing from the ALK, but which could be added if the search continues with coarser aggregation levels (3–12). The age columns in red (ages 2 and 16–18) are the ages which are found in the global ALK (i.e., using all data from the 1993–2020 period), but which will never be found in the ALK since these ages were never observed during the 2004–2006 period. The row highlighted in orange (27 cm length) is the only length at aggregation level 2 that was only found in the LF data.

			Probability at age																
Length	**LF**	**Age**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**
27	8	0	0.427	0.533	0.039	0.001	-	-	-	-	-	-	-	-	-	-	-	-	-
49	1	1	-	-	-	1.000	-	-	-	-	-	-	-	-	-	-	-	-	-
50	4	3	-	-	0.333	0.667	-	-	-	-	-	-	-	-	-	-	-	-	-
51	17	8	-	-	0.125	0.250	0.375	0.250	-	-	-	-	-	-	-	-	-	-	-
52	17	6	-	-	-	0.333	0.167	0.333	0.167	-	-	-	-	-	-	-	-	-	-
53	30	5	-	-	-	-	-	0.600	0.200	0.200	-	-	-	-	-	-	-	-	-
54	42	1	-	-	-	-	-	1.000	-	-	-	-	-	-	-	-	-	-	-
55	70	9	-	-	-	0.111	0.111	0.333		0.444	-	-	-	-	-	-	-	-	-
…	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-
105	1	1	-	-	-	-	-	-	-	-	-	1.000	-	-	-	-	-	-	-

Table 12. Description of the columns found in the dataset produced by the `get.samples` function from the `catchR` package.

Variable	Description
`id`	Stratum *k*
`year`	Year
`period`	Period (month for this analysis)
`region`	Region (NAFO unit areas for this analysis)
`gear`	Gear grouping
`catch`	Cod landings in stratum *k* (kg)
`length`	Length *j* (cm)
`n.lf`	Number of length *j* cod measured for LF
`lf.prop`	*n.lf* / *n.lftot*
`n.lftot`	Total number of cod measured for LF
`weight.sample`	Weight of all length *j* cod measured for LF
`weight.sample.tot`	Weight of all cod measured for LF
`weight.unit`	Weight of a length *j* cod measured for LF
`n.al`	Number of length *j* cod used in the age-length key
`n.altot`	Total number of cod used in the age-length key
`age.2`	Proportion of length *j* cod that are age 2 in stratum *k*
…	…
`age.18`	Proportion of length *j* cod that are age 18 in stratum *k*. The number of age columns will vary depending on dataset used as input.
`nsample.lengthfreq`	Number of samples used to infer length frequencies in stratum *k*
`nsample.agelength`	Number of samples used to produce an age-length key in stratum *k*
`option.lengthfreq`	Aggregation level used in the search for LF samples (1–12)
`option.agelength`	Aggregation level used in the search for age samples (1–12)

Table 13. Number of Atlantic cod samples from the 3Pn4RS stock collected for length frequencies in the commercial fishery for the period 1993–2020, by year and month. Values in brackets are the number of cod measured. Source: DFO dockside monitoring program data.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
1993	-	-	6 (1,512)	6 (1,535)	22 (5,300)	27 (7,317)	44 (10,922)	15 (3,301)	17 (3,796)	18 (4,255)	13 (3,259)	1 (254)	169 (41,451)
1994	-	-	-	-	1 (250)	2 (486)	1 (257)	1 (204)	-	-	9 (28)	64 (926)	78 (2,151)
1995	*33 (324)*	-	-	48 (396)	21 (487)	1 (251)	101 (1,051)	249 (2,859)	187 (2,887)	112 (1,485)	21 (693)	5 (117)	778 (10,550)
1996	-	-	-	-	1 (250)	14 (664)	138 (2,292)	136 (3,169)	96 (1,280)	19 (425)	12 (569)	2 (39)	418 (8,688)
1997	1 (9)	3 (59)	-	-	6 (1,094)	21 (1,131)	119 (11,114)	92 (4,366)	102 (5,857)	27 (2,030)	8 (710)	4 (98)	383 (26,468)
1998	-	-	-	-	1 (249)	4 (531)	7 (1,559)	6 (1,266)	2 (486)	1 (256)	-	-	21 (4,347)
1999	-	-	-	-	1 (270)	4 (733)	15 (2,929)	12 (2,250)	5 (1,143)	5 (700)	-	-	42 (8,025)
2000	-	-	-	-	5 (1,237)	3 (631)	8 (1,782)	12 (2,313)	6 (1,472)	2 (442)	-	-	36 (7,877)
2001	-	-	-	-	4 (935)	5 (1,097)	9 (1,974)	7 (1,320)	6 (1,088)	4 (1,020)	1 (252)	-	36 (7,686)
2002	-	-	-	-	4 (1,012)	-	8 (1,761)	9 (1,627)	5 (1,262)	2 (470)	-	-	28 (6,132)
2003	-	-	-	-	-	-	3 (605)	-	-	-	-	-	3 (605)
2004	-	-	-	-	1 (252)	2 (473)	13 (2,781)	2 (352)	5 (889)	1 (122)	-	-	24 (4,869)
2005	-	-	-	-	-	3 (507)	19 (2,954)	7 (1,167)	3 (464)	2 (259)	-	-	34 (5,351)
2006	-	-	-	-	-	1 (161)	25 (3,868)	7 (1,041)	7 (1,028)	2 (320)	-	-	42 (6,418)
2007	-	-	-	-	1 (163)	2 (312)	23 (3,584)	2 (314)	8 (1,023)	7 (1,131)	-	-	43 (6,527)
2008	-	-	-	-	1 (152)	-	13 (1,973)	6 (930)	9 (1,368)	3 (435)	-	-	32 (4,858)
2009	-	-	-	-	-	-	14 (3,217)	3 (535)	10 (1,358)	4 (701)	-	-	31 (5,811)
2010	-	-	-	-	-	1 (165)	10 (1,470)	8 (1,123)	7 (949)	2 (269)	-	-	28 (3,976)
2011	-	-	-	-	1 (117)	-	12 (1,795)	5 (780)	7 (1,168)	1 (151)	-	-	26 (4,011)
2012	-	-	-	1 (17)	-	1 (158)	9 (1,505)	3 (453)	3 (601)	1 (150)	2 (451)	-	20 (3,335)
2013	-	-	-	-	1 (154)	-	12 (1,866)	2 (309)	4 (681)	-	-	-	19 (3,010)
2014	-	-	-	-	-	-	9 (1,415)	2 (303)	5 (767)	1 (152)	-	-	17 (2,637)
2015	-	-	-	-	-	1 (150)	10 (1,535)	2 (300)	2 (300)	1 (154)	6 (950)	-	22 (3,389)
2016	-	-	-	-	-	-	13 (2,017)	5 (646)	2 (367)	1 (152)	4 (1,066)	-	25 (4,248)
2017	-	-	-	1 (185)	-	3 (481)	11 (1,859)	6 (890)	6 (856)	-	2 (349)	-	29 (4,620)
2018	-	-	-	1 (150)	-	-	12 (1,887)	4 (528)	2 (306)	6 (1,001)	-	-	25 (3,872)
2019	-	-	-	-	1 (156)	1 (148)	4 (460)	1 (152)	2 (302)	-	1 (168)	-	10 (1,386)
2020	-	-	-	-	-	1 (151)	-	1 (150)	2 (309)	-	-	-	4 (610)
Total	**34 (333)**	**3 (59)**	**6 (1,512)**	**57 (2,283)**	**72 (12,078)**	**97 (15,547)**	**662 (70,432)**	**605 (32,648)**	**510 (32,007)**	**222 (16,080)**	**79 (8,495)**	**76 (1,434)**	**2,423 (192,908)**

Table 14. Number of Atlantic cod samples from the 3Pn4RS stock collected for length frequencies in the commercial fishery for the period 1993–2020, by year and NAFO unit area. Values in brackets are the number of cod measured. Source: DFO dockside monitoring program data.

Year	3Pn	4Ru	4Ra	4Rb	4Rc	4Rd
1993	24 (5,652)	1 (242)	-	81 (20,611)	18 (3,837)	17 (4,222)
1994	36 (751)	22 (160)	-	4 (5)	1 (250)	6 (973)
1995	168 (3,412)	2 (7)	196 (2,066)	62 (583)	54 (1,042)	84 (1,311)
1996	29 (941)	99 (1,517)	16 (184)	27 (231)	2 (26)	37 (2,424)
1997	51 (5,276)	-	56 (1,524)	104 (11,554)	34 (2,184)	32 (1,909)
1998	2 (503)	-	-	7 (985)	2 (506)	1 (256)
1999	-	-	-	21 (3,004)	4 (1,024)	2 (520)
2000	4 (1,001)	-	-	18 (3,226)	3 (785)	-
2001	3 (562)	-	-	12 (2,190)	12 (3,065)	2 (286)
2002	6 (1,279)	-	-	12 (2,324)	3 (798)	1 (165)
2003	-	-	-	-	1 (248)	-
2004	2 (505)	-	-	13 (2,314)	1 (207)	-
2005	1 (199)	-	1 (157)	22 (3,386)	2 (310)	4 (643)
2006	-	-	-	21 (3,197)	-	8 (1,272)
2007	-	-	1 (172)	20 (2,941)	7 (1,100)	3 (476)
2008	-	-	2 (251)	15 (2,266)	3 (461)	4 (608)
2009	-	-	2 (219)	19 (3,291)	1 (227)	1 (199)
2010	-	3 (356)	1 (155)	12 (1,505)	1 (156)	-
2011	2 (339)	-	1 (89)	9 (1,310)	1 (161)	-
2012	3 (468)	-	-	5 (983)	1 (158)	-
2013	-	-	-	6 (966)	2 (312)	1 (159)
2014	-	-	-	5 (769)	-	-
2015	1 (148)	-	-	3 (478)	-	5 (802)
2016	-	-	-	7 (1,079)	-	4 (1,066)
2017	4 (694)	-	4 (767)	4 (625)	-	-
2018	-	-	1 (112)	10 (1,586)	-	-
2019	-	-	1 (45)	1 (115)	-	1 (168)
2020	-	-	-	-	-	-
Total	**336 (21,730)**	**127 (2,282)**	**282 (5,741)**	**520 (71,524)**	**153 (16,857)**	**213 (17,459)**

Table 14 (continued).

Year	4Su	4Si	4Ss	4Sv	4Sw	4Sx	4Sy	4Sz	Total
1993	-	2 (502)	1 (250)	4 (1,001)	11 (2,532)	5 (1,264)	1 (255)	4 (1,083)	169 (41,451)
1994	-	-	-	1 (3)	8 (9)	-	-	-	78 (2,151)
1995	-	-	1 (251)	58 (638)	153 (1,240)	-	-	-	778 (10,550)
1996	-	-	1 (121)	53 (826)	154 (2,418)	-	-	-	418 (8,688)
1997	1 (4)	-	1 (260)	30 (670)	70 (2,832)	1 (251)	-	3 (4)	383 (26,468)
1998	-	-	1 (251)	-	8 (1,846)	-	-	-	21 (4,347)
1999	-	-	-	3 (782)	12 (2,695)	-	-	-	42 (8,025)
2000	-	-	2 (484)	1 (255)	6 (1,572)	2 (554)	-	-	36 (7,877)
2001	-	-	-	2 (463)	4 (872)	1 (248)	-	-	36 (7,686)
2002	-	-	3 (752)	-	3 (814)	-	-	-	28 (6,132)
2003	-	-	2 (357)	-	-	-	-	-	3 (605)
2004	1 (252)	-	3 (572)	2 (482)	2 (537)	-	-	-	24 (4,869)
2005	-	-	-	-	4 (656)	-	-	-	34 (5,351)
2006	-	-	-	6 (907)	5 (794)	1 (161)	-	1 (87)	42 (6,418)
2007	-	-	2 (312)	6 (906)	4 (620)	-	-	-	43 (6,527)
2008	-	-	-	1 (151)	7 (1,121)	-	-	-	32 (4,858)
2009	-	-	-	2 (347)	6 (1,528)	-	-	-	31 (5,811)
2010	-	-	-	5 (764)	4 (741)	2 (299)	-	-	28 (3,976)
2011	-	-	-	3 (466)	6 (1,077)	4 (569)	-	-	26 (4,011)
2012	-	-	1 (150)	-	4 (668)	6 (908)	-	-	20 (3,335)
2013	-	-	-	3 (462)	6 (956)	1 (155)	-	-	19 (3,010)
2014	-	1 (121)	1 (152)	1 (163)	6 (982)	3 (450)	-	-	17 (2,637)
2015	-	1 (150)	3 (454)	-	6 (1,006)	1 (150)	-	2 (201)	22 (3,389)
2016	-	-	1 (45)	1 (152)	4 (700)	7 (1,053)	-	1 (153)	25 (4,248)
2017	1 (150)	2 (321)	1 (150)	1 (153)	5 (713)	4 (596)	1 (150)	2 (301)	29 (4,620)
2018	-	1 (150)	1 (155)	2 (305)	5 (802)	1 (150)	-	4 (612)	25 (3,872)
2019	-	3 (458)	-	-	-	1 (151)	2 (299)	1 (150)	10 (1,386)
2020	-	1 (151)	1 (152)	1 (157)	-	-	1 (150)	-	4 (610)
Total	**3 (406)**	**11 (1,853)**	**26 (4,868)**	**186 (10,053)**	**503 (29,731)**	**40 (6,959)**	**5 (854)**	**18 (2,591)**	**2,423 (192,908)**

Table 15. Number of Atlantic cod samples from the 3Pn4RS stock collected for length frequencies in the commercial fishery for the period 1993–2020, by year and gear category. Values in brackets are the number of cod measured. Source: DFO dockside monitoring program data.

Year	Shrimp trawls	Other trawls	Seines	Gillnets	Longlines	Hand gear	Traps
1993	59 (15,557)	10 (2,451)	33 (7,188)	30 (7,021)	5 (905)	32 (8 329)	169 (41,451)
1994	1 (190)	4 (947)	9 (12)	64 (1,002)	-	-	78 (2,151)
1995	-	4 (494)	350 (3,490)	389 (6,225)	5 (99)	30 (242)	778 (10,550)
1996	1 (121)	9 (1,958)	249 (3,809)	130 (2,382)	-	29 (418)	418 (8,688)
1997	-	2 (264)	150 (3,056)	103 (8,377)	111 (14,552)	17 (219)	383 (26,468)
1998	-	-	5 (1,081)	7 (1,766)	9 (1,500)	-	21 (4,347)
1999	-	-	30 (4,937)	6 (1,544)	6 (1,544)	-	42 (8,025)
2000	-	-	18 (3,279)	16 (4,012)	2 (586)	-	36 (7,877)
2001	-	2 (286)	14 (2,605)	18 (4,361)	2 (434)	-	36 (7,686)
2002	-	1 (165)	9 (1,659)	18 (4,308)	-	-	28 (6,132)
2003	-	-	-	3 (605)	-	-	3 (605)
2004	-	-	11 (1,834)	13 (3,035)	-	-	24 (4,869)
2005	-	1 (189)	24 (3,715)	9 (1,447)	-	-	34 (5,351)
2006	-	-	29 (4,349)	12 (1,890)	1 (179)	-	42 (6,418)
2007	-	-	27 (4,104)	16 (2,423)	-	-	43 (6,527)
2008	-	-	22 (3,308)	10 (1,550)	-	-	32 (4,858)
2009	-	-	19 (3,625)	11 (1,947)	1 (239)	-	31 (5,811)
2010	-	-	18 (2,398)	9 (1,389)	1 (189)	-	28 (3,976)
2011	-	-	15 (2,323)	10 (1,535)	1 (153)	-	26 (4,011)
2012	-	-	11 (2,129)	8 (1,075)	1 (131)	-	20 (3,335)
2013	-	-	12 (1,921)	5 (779)	2 (310)	-	19 (3,010)
2014	-	-	9 (1,443)	6 (886)	2 (308)	-	17 (2,637)
2015	-	-	10 (1,535)	12 (1,854)	-	-	22 (3,389)
2016	-	-	10 (1,565)	15 (2,683)	-	-	25 (4,248)
2017	-	-	11 (1,699)	17 (2,766)	1 (155)	-	29 (4,620)
2018	-	-	13 (1,987)	12 (1,885)	-	-	25 (3,872)
2019	-	-	3 (312)	7 (1,074)	-	-	10 (1,386)
2020	-	-	-	4 (610)	-	-	4 (610)
Total	**61 (15,868)**	**33 (6,754)**	**1,111 (69 363)**	**960 (70,431)**	**150 (21,284)**	**108 (9,208)**	**2,423 (192,908)**

Table 16. Number of Atlantic cod samples from the 3Pn4RS stock collected for age readings in the commercial fishery for the period 1993–2020, by year and month. Values in brackets are the number of aged cod. Source: DFO port sampling program data.

Year	Jan.	Feb.	Mar.	Apr.	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.	Total
1993	-	-	5 (129)	4 (99)	16 (558)	22 (450)	18 (376)	8 (286)	7 (175)	16 (440)	13 (392)	-	109 (2,905)
1994	-	-	-	-	1 (50)	2 (74)	1 (35)	1 (21)	-	-	9 (26)	63 (705)	77 (911)
1995	33 (301)	-	-	39 (237)	18 (269)	1 (41)	49 (170)	130 (424)	122 (528)	106 (685)	19 (200)	5 (32)	522 (2,887)
1996	-	-	-	-	-	13 (308)	68 (433)	57 (409)	37 (164)	15 (311)	7 (98)	-	197 (1,723)
1997	-	3 (55)		-	6 (215)	11 (265)	62 (613)	52 (331)	39 (198)	21 (430)	5 (122)	-	199 (2,229)
1998	-	-	-	-	1 (45)	4 (101)	7 (246)	6 (164)	2 (59)	1 (44)	-	-	21 (659)
1999	-	-	-	-	1 (44)	4 (142)	12 (349)	8 (204)	4 (88)	5 (157)	-	-	34 (984)
2000	-	-	-	-	5 (173)	2 (63)	8 (240)	12 (348)	6 (143)	-	-	-	33 (967)
2001	-	-	-	-	4 (153)	5 (204)	7 (251)	7 (221)	5 (175)	4 (170)	1 (35)	-	33 (1,209)
2002	-	-	-	-	4 (146)	-	8 (274)	9 (292)	3 (92)	2 (73)		-	26 (877)
2003	-	-	-	-	-	-	-	-	-	-	-	-	-
2004	-	-	-	-	1 (39)	2 (74)	13 (414)	2 (68)	5 (140)	1 (33)	-	-	24 (768)
2005	-	-	-	-	-	1 (38)	17 (359)	3 (86)	2 (67)	2 (60)	-	-	25 (610)
2006	-	-	-	-	-	1 (32)	25 (517)	6 (141)	4 (81)	-	-	-	36 (771)
2007	-	-	-	-	1 (33)	2 (70)	19 (511)	2 (53)	8 (175)	7 (276)	-	-	39 (1,118)
2008	-	-	-	-	1 (35)		13 (401)	6 (247)	9 (121)	3 (105)	-	-	32 (909)
2009	-	-	-	-	-		13 (396)	3 (120)	6 (188)	4 (150)	-	-	26 (854)
2010	-	-	-	-	-	1 (42)	9 (269)	6 (199)	3 (98)	2 (67)	-	-	21 (675)
2011	-	-	-	-	1 (31)		12 (338)	4 (132)	4 (105)	1 (28)	-	-	22 (634)
2012	-	-	-	1 (15)	-	1 (45)	9 (249)	3 (82)	1 (34)	1 (28)	-	-	16 (453)
2013	-	-	-	-	1 (35)		12 (337)	2 (77)	4 (136)	-	-	-	19 (585)
2014	-	-	-	-	-		9 (281)	2 (60)	5 (183)	1 (31)	-	-	17 (555)
2015	-	-	-	-	-	1 (33)	10 (256)	2 (79)	2 (78)	1 (33)	5 (193)	-	21 (672)
2016	-	-	-	-	-		13 (450)	4 (142)	2 (79)	1 (41)	4 (152)	-	24 (864)
2017	-	-	-	1 (43)	-	3 (104)	11 (426)	6 (200)	6 (213)	-	2 (66)	-	29 (1,052)
2018	-	-	-	1 (36)	-		12 (426)	4 (172)	2 (76)	6 (244)	-	-	25 (954)
2019	-	-	-	-	1 (43)	1 (38)	4 (156)	1 (39)	2 (76)	-	1 (24)	-	10 (376)
2020	-	-	-	-	-	1 (42)	-	1 (47)	2 (91)	-	-	-	4 (180)
Total	**33 (301)**	**3 (55)**	**5 (129)**	**46 (430)**	**62 (1,869)**	**78 (2,166)**	**441 (8,773)**	**347 (4,644)**	**292 (3,563)**	**200 (3,406)**	**66 (1,308)**	**68 (737)**	**1,641 (27,381)**

Table 17. Number of Atlantic cod samples from the 3Pn4RS stock collected for age readings in the commercial fishery for the period 1993–2020, by year and NAFO unit area. Values in brackets are the number of aged cod. Source: DFO port sampling program data.

Year	3Pn	4Ru	4Ra	4Rb	4Rc	4Rd
1993	19 (679)	-	-	50 (1,090)	10 (285)	12 (385)
1994	35 (537)	22 (152)	-	4 (4)	1 (50)	6 (156)
1995	151 (1,206)	2 (6)	101 (440)	35 (180)	19 (88)	49 (291)
1996	21 (370)	61 (526)	12 (70)	10 (52)	-	14 (354)
1997	38 (1,057)	-	30 (242)	22 (248)	12 (113)	16 (237)
1998	2 (77)	-	-	7 (199)	2 (85)	1 (43)
1999	-	-	-	16 (401)	4 (153)	2 (80)
2000	2 (63)	-	-	18 (482)	3 (114)	-
2001	2 (78)	-	-	12 (433)	12 (466)	2 (72)
2002	6 (214)	-	-	10 (302)	3 (120)	1 (50)
2003	-	-	-	-	-	-
2004	2 (71)	-	-	13 (390)	1 (37)	-
2005	1 (27)	-	1 (35)	15 (322)	1 (38)	3 (107)
2006	-	-	-	16 (245)	-	8 (219)
2007	-	-	-	17 (492)	7 (238)	3 (109)
2008	-	-	2 (15)	15 (363)	3 (124)	4 (147)
2009	-	-	1 (34)	16 (549)	1 (31)	1 (37)
2010	-	1 (31)	1 (28)	8 (261)	1 (33)	-
2011	2 (49)	-	1 (29)	6 (210)	1 (31)	-
2012	1 (15)	-	-	3 (103)	1 (45)	-
2013	-	-	-	6 (203)	2 (72)	1 (35)
2014	-	-	-	5 (186)	-	-
2015	1 (51)	-	-	3 (107)	-	4 (142)
2016	-	-	-	7 (270)	-	4 (152)
2017	4 (144)	-	4 (158)	4 (138)	-	-
2018	-	-	1 (40)	10 (407)	-	-
2019	-	-	1 (40)	1 (41)	-	1 (24)
2020	-	-	-	-	-	-
Total	**287 (4,638)**	**86 (715)**	**155 (1,131)**	**329 (7,678)**	**84 (2,123)**	**132 (2,640)**

Table 17 (continued).

Year	4Su	4Si	4Ss	4Sv	4Sw	4Sx	4Sy	4Sz	Total
1993	-	1 (27)	1 (44)	4 (108)	6 (113)	1 (43)	1 (37)	4 (94)	109 (2,905)
1994	-	-	-	1 (3)	8 (9)	-	-	-	77 (911)
1995	-	-	1 (41)	45 (169)	119 (466)	-	-	-	522 (2,887)
1996	-	-	-	20 (40)	59 (311)	-	-	-	197 (1,723)
1997	1 (3)	-	1 (55)	20 (74)	55 (157)	1 (39)	-	3 (4)	199 (2,229)
1998	-	-	1 (46)	-	8 (209)	-	-	-	21 (659)
1999	-	-	-	2 (57)	10 (293)	-	-	-	34 (984)
2000	-	-	2 (81)	1 (38)	6 (152)	1 (37)	-	-	33 (967)
2001	-	-	-	2 (49)	2 (65)	1 (46)	-	-	33 (1,209)
2002	-	-	3 (118)	-	3 (73)	-	-	-	26 (877)
2003	-	-	-	-	-	-	-	-	-
2004	1 (37)	-	3 (110)	2 (75)	2 (48)	-	-	-	24 (768)
2005	-	-	-	-	4 (81)	-	-	-	25 (610)
2006	-	-	-	6 (145)	4 (96)	1 (32)	-	1 (34)	36 (771)
2007	-	-	2 (70)	6 (127)	4 (82)	-	-	-	39 (1,118)
2008	-	-	-	1 (37)	7 (223)	-	-	-	32 (909)
2009	-	-	-	2 (85)	5 (118)	-	-	-	26 (854)
2010	-	-	-	5 (173)	3 (66)	2 (83)	-	-	21 (675)
2011	-	-	-	3 (101)	5 (99)	4 (115)	-	-	22 (634)
2012	-	-	1 (34)	-	4 (91)	6 (165)	-	-	16 (453)
2013	-	-	-	3 (109)	6 (129)	1 (37)	-	-	19 (585)
2014	-	1 (32)	1 (31)	1 (37)	6 (159)	3 (110)	-	-	17 (555)
2015	-	1 (33)	3 (113)	-	6 (141)	1 (37)	-	2 (48)	21 (672)
2016	-	-	-	1 (41)	4 (114)	7 (253)	-	1 (34)	24 (864)
2017	1 (1)	2 (69)	1 (33)	1 (43)	5 (182)	4 (170)	1 (42)	2 (72)	29 (1,052)
2018	-	1 (36)	1 (32)	2 (84)	5 (147)	1 (41)	-	4 (167)	25 (954)
2019	-	3 (122)	-	-	-	1 (39)	2 (75)	1 (35)	10 (376)
2020	-	1 (42)	1 (54)	1 (37)	-	-	1 (47)	-	4 (180)
Total	**3 (41)**	**10 (361)**	**22 (862)**	**129 (1,632)**	**346 (3,624)**	**35 (1,247)**	**5 (201)**	**18 (488)**	**1,641 (27,381)**

Table 18. Number of Atlantic cod samples from the 3Pn4RS stock collected for age readings in the commercial fishery for the period 1993–2020, by year and gear category. Values in brackets are the number of aged cod. Source: DFO port sampling program data.

Year	Other trawls	Seines	Gillnets	Longlines	Hand gear	Traps	Total
1993	46 (914)	8 (240)	19 (619)	22 (901)	1 (24)	13 (207)	109 (2,905)
1994	-	4 (130)	9 (12)	64 (769)	-	-	77 (911)
1995	-	4 (76)	219 (798)	296 (1 994)	3 (19)	-	522 (2,887)
1996	-	9 (298)	76 (332)	86 (761)	-	26 (332)	197 (1,723)
1997	-	-	113 (412)	86 (1,817)	-	-	199 (2,229)
1998	-	-	5 (162)	7 (281)	9 (216)	-	21 (659)
1999	-	-	23 (590)	6 (233)	5 (161)	-	34 (984)
2000	-	-	18 (433)	13 (479)	2 (55)	-	33 (967)
2001	-	2 (72)	14 (465)	17 (672)	-	-	33 (1,209)
2002	-	1 (50)	9 (255)	16 (572)	-	-	26 (877)
2003	-	-	-	-	-	-	
2004	-	-	11 (291)	13 (477)	-	-	24 (768)
2005	-	-	18 (376)	7 (234)	-	-	25 (610)
2006	-	-	24 (424)	12 (347)	-	-	36 (771)
2007	-	-	23 (608)	16 (510)	-	-	39 (1,118)
2008	-	-	22 (520)	10 (389)	-	-	32 (909)
2009	-	-	15 (457)	11 (397)	-	-	26 (854)
2010	-	-	12 (363)	9 (312)	-	-	21 (675)
2011	-	-	12 (338)	10 (296)	-	-	22 (634)
2012	-	-	7 (214)	8 (214)	1 (25)	-	16 (453)
2013	-	-	12 (343)	5 (181)	2 (61)	-	19 (585)
2014	-	-	9 (288)	6 (210)	2 (57)	-	17 (555)
2015	-	-	10 (256)	11 (416)	-	-	21 (672)
2016	-	-	10 (339)	14 (525)	-	-	24 (864)
2017	-	-	11 (405)	17 (617)	1 (30)	-	29 (1,052)
2018	-	-	13 (494)	12 (460)	-	-	25 (954)
2019	-	-	3 (120)	7 (256)	-	-	10 (376)
2020	-	-	-	4 (180)	-	-	4 (180)
Total	**46 (914)**	**28 (866)**	**717 (9,914)**	**785 (14,500)**	**26 (648)**	**39 (539)**	**1,641 (27,381)**

8. FIGURES

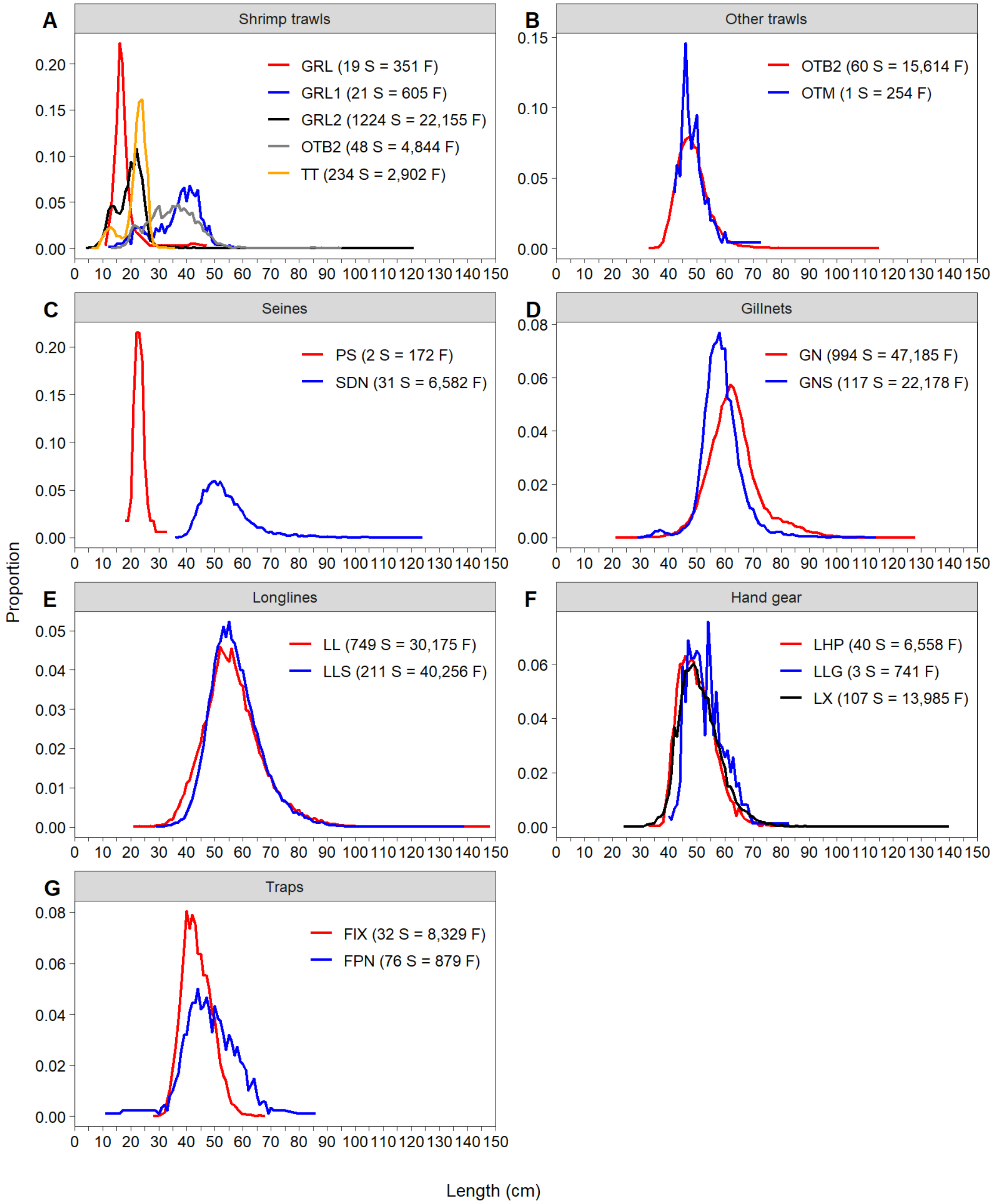

Figure 1. Length frequency distributions of the different gear from which Atlantic cod length measurements were recorded for the period 1993–2020. Each panel represents a gear category. For each gear, the number of samples (S) from which fish (F) were measured is indicated in brackets. See Table 4 for gear description.

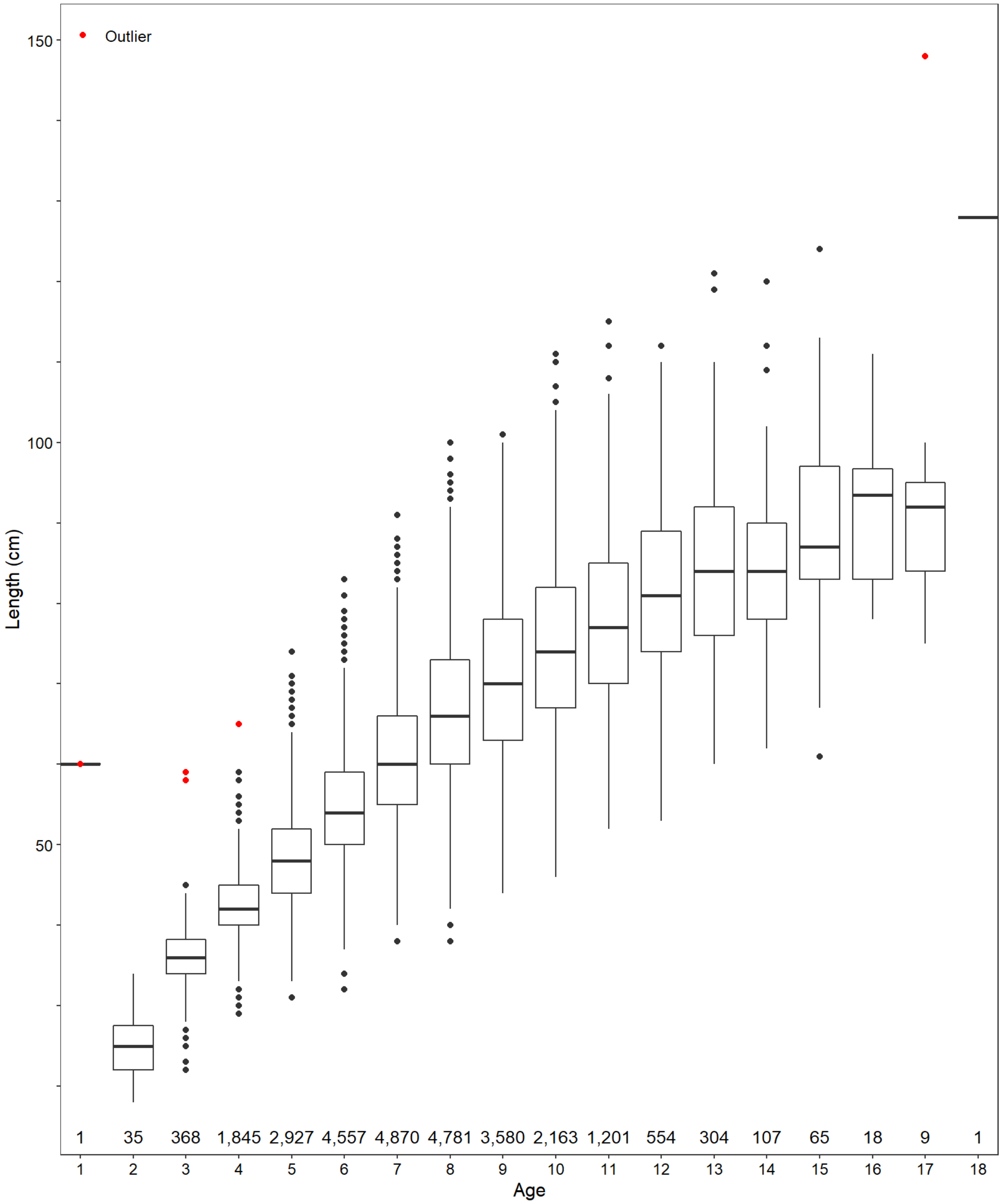

Figure 2. Boxplots showing the length distribution of aged cod for the 1993–2020 period. The bottom, middle, and top horizontal lines of the boxplots represent the 25th, 50th (median), and 75th percentiles, respectively. The upper whisker extends from the box to the highest value not exceeding 1.5 times the interquartile range. The lower whisker follows the same principle, but with the lower values. Red dots (5) correspond to cod that were categorized as outliers and not used in CAA analyses. The number of cod by age is given under each boxplot.

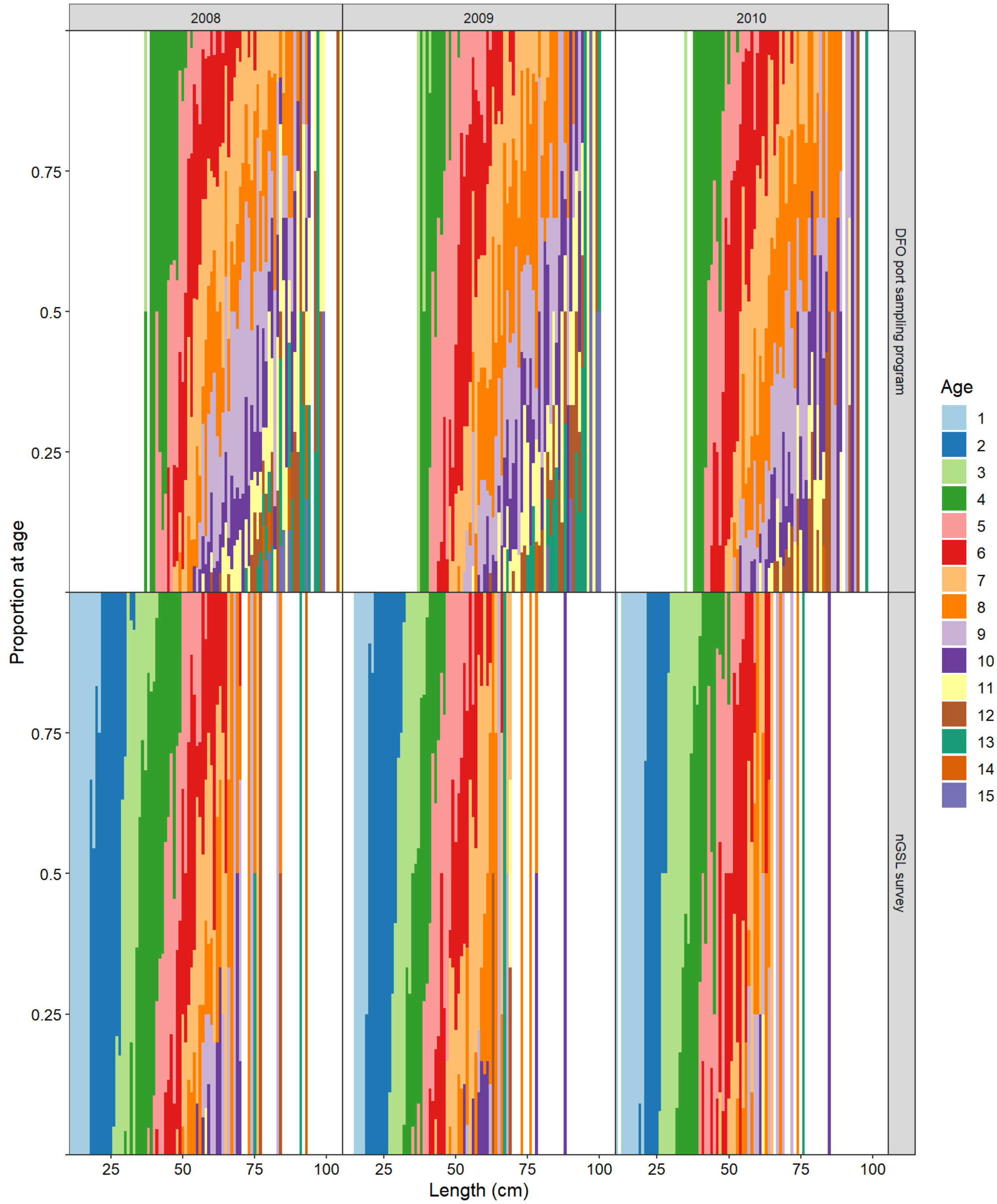

Figure 3. Annual ALKs for the years 2008–2010 based on the DFO port sampling program data and those from the nGSL multispecies bottom trawl survey.

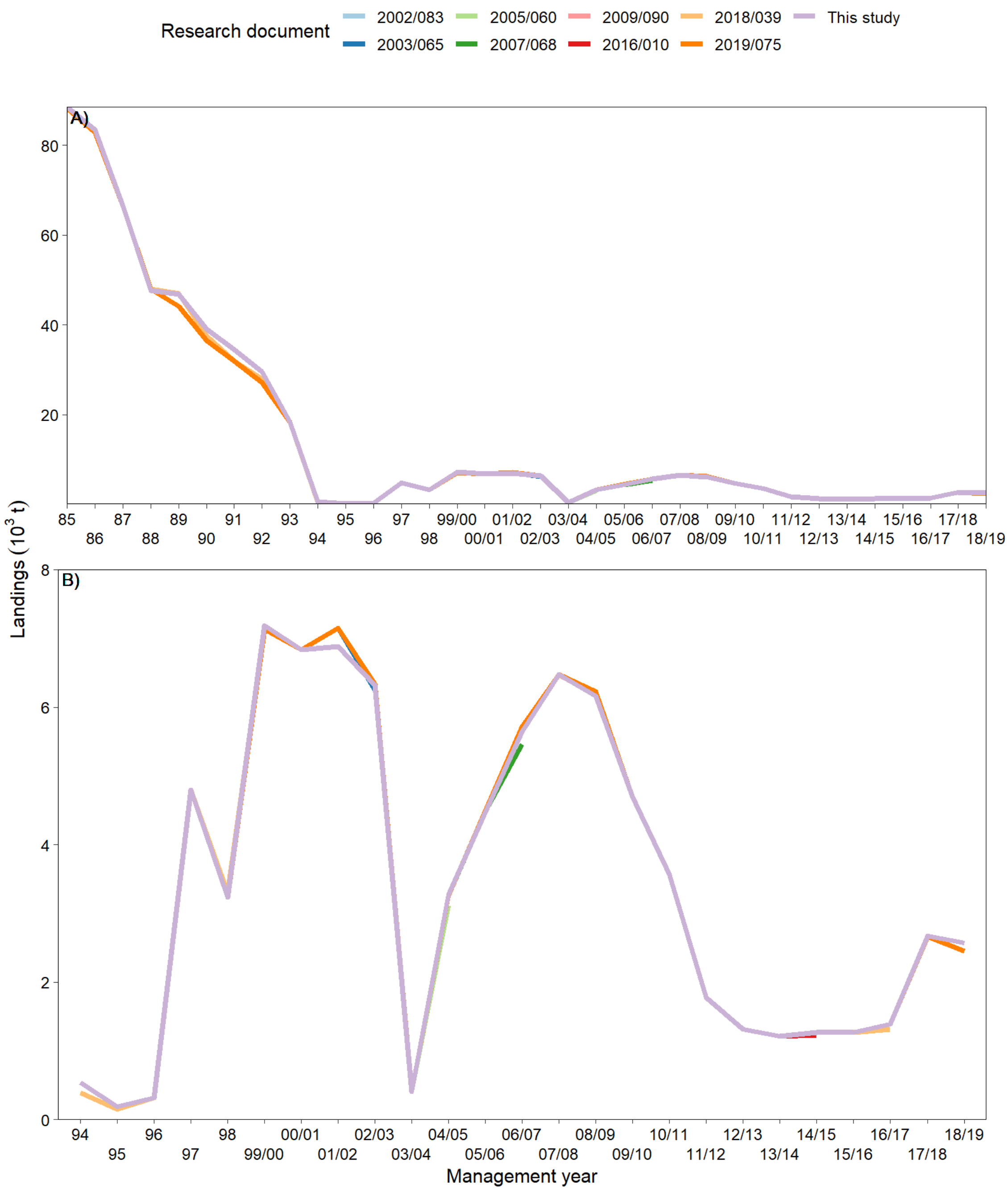

Figure 4. A) Evolution of the 3Pn4RS Atlantic cod stock landings used for this study compared to those published in the last 8 research documents. Years before 1985 did not show any differences and are therefore not shown. An enlarged inset of the post-1993 period is shown in B).

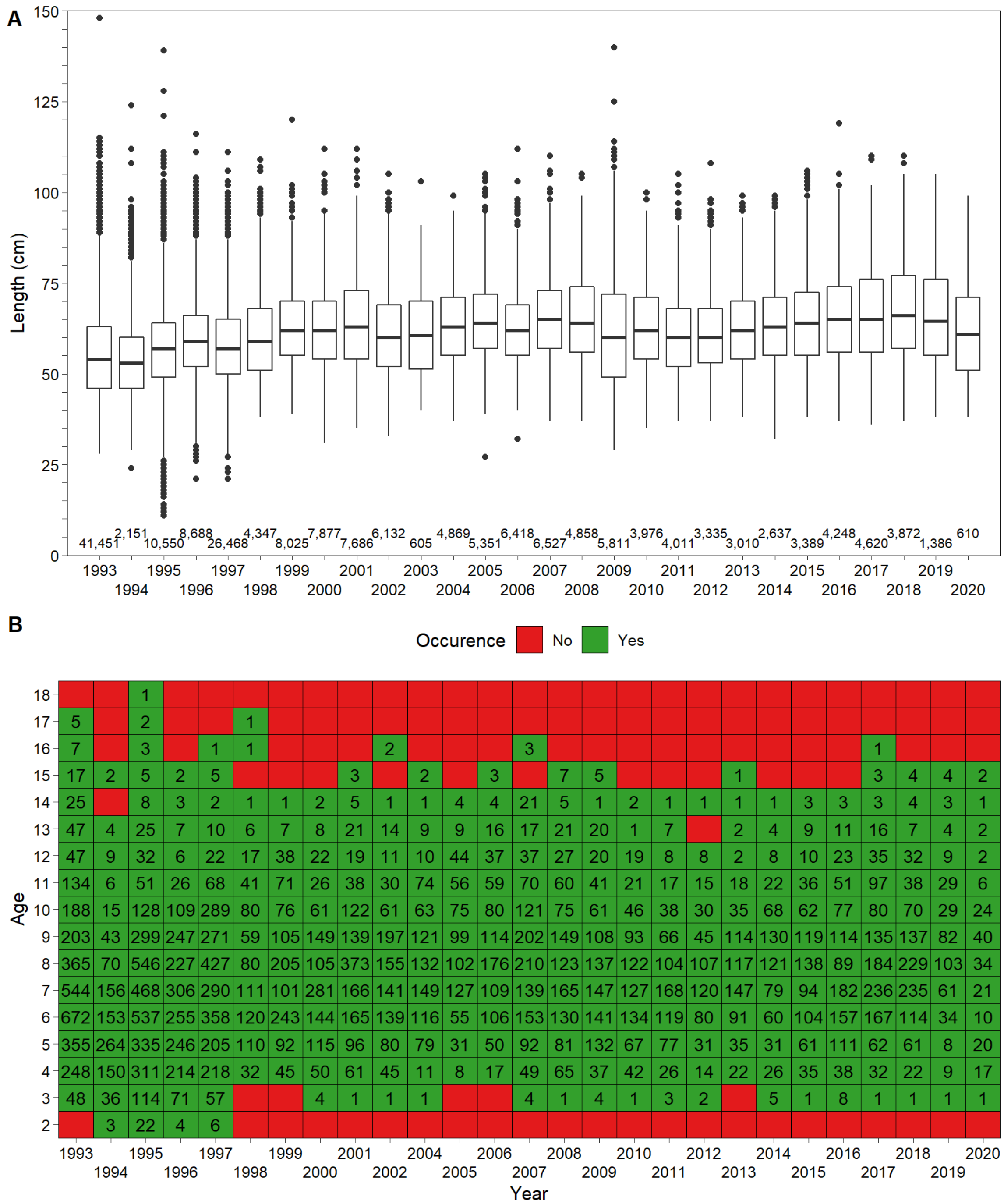

Figure 5. A) LF and B) age data used for CAA calculations of 3Pn4RS Atlantic cod over the 1993–2020 period. Outliers were initially removed. In A), the number of cod measured for length is shown under each boxplot. In B), values within cells are the number of cod of a given age and year. See Figure 2 for a boxplot description. Source: DFO port sampling program.

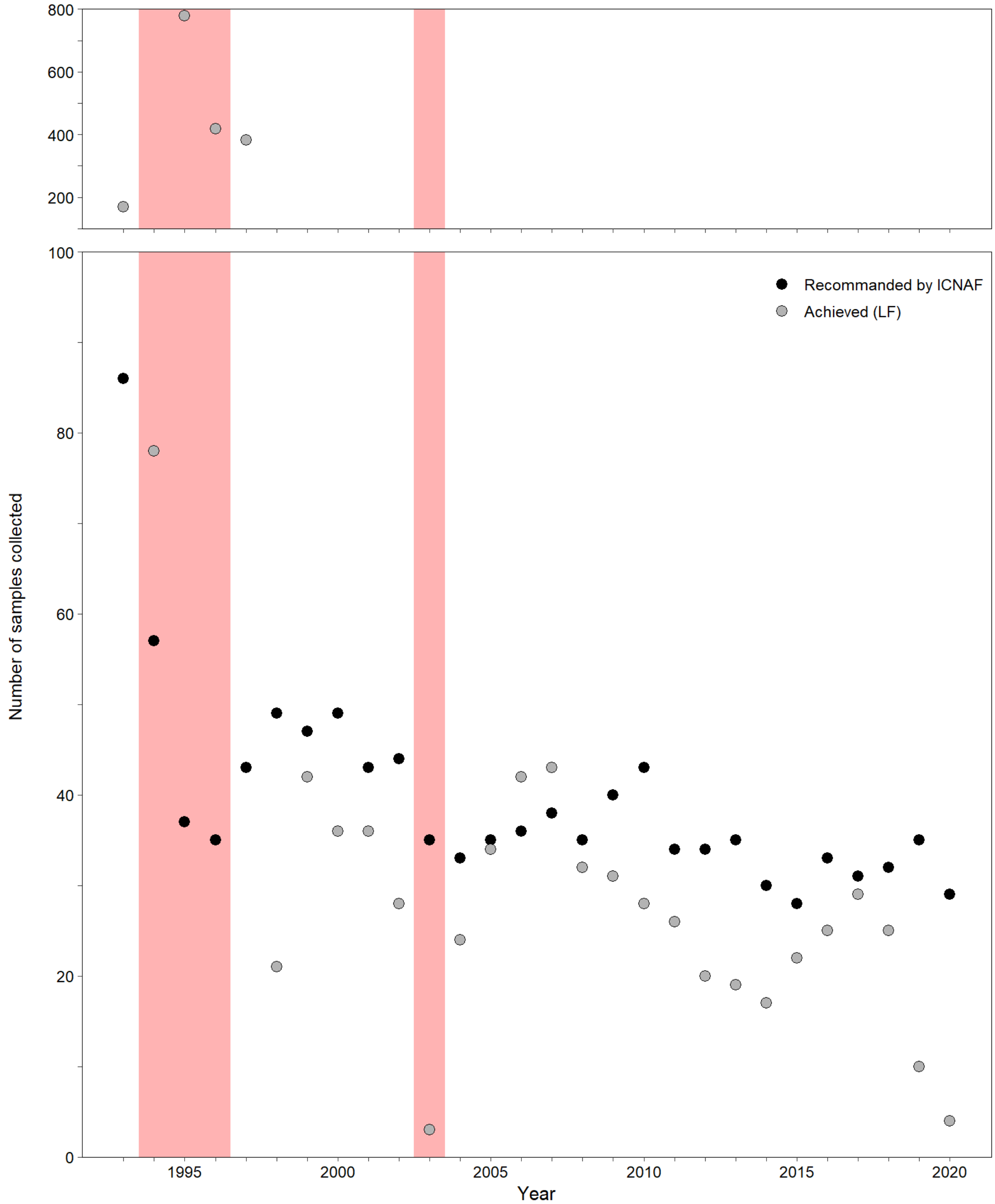

Figure 6. Comparison of the number of length frequency samples collected annually as part of the dockside monitoring program since 1993 and the ICNAF recommendation (1 sample per 1,000 tonnes landed per year, quarter, NAFO division and gear. Anon. 1974). Moratorium years are indicated with the pink background. Source: DFO port sampling program and ZIFF data.

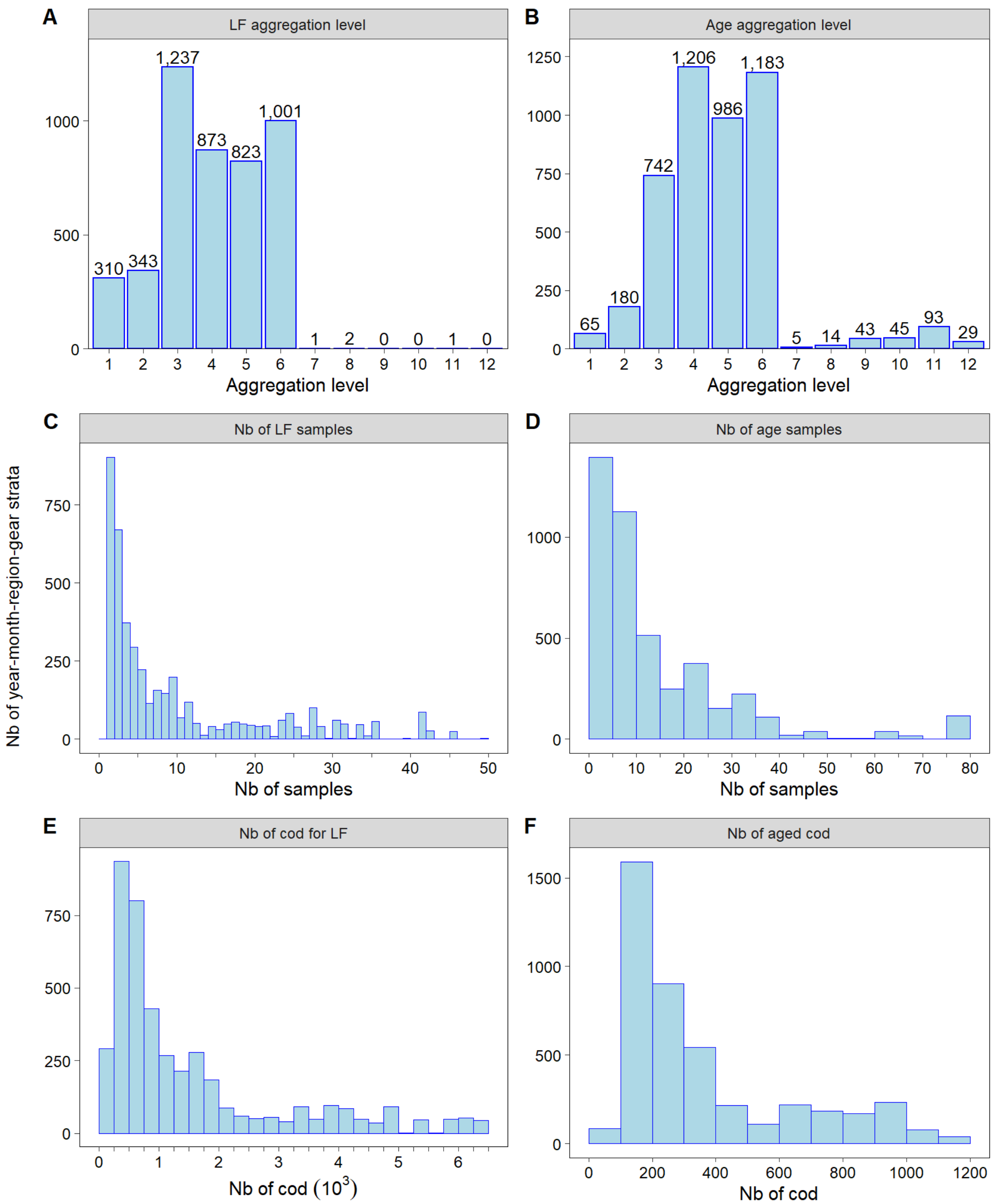

Figure 7. Aggregation levels (A-B), number of samples (C-D) and number of fish (E-F) used in the calculations of length frequencies (left) and ALKs (right) for the analyzes of age catches of commercial Atlantic cod fisheries in 3Pn4RS during the 1993–2020 period. Values above bars (A-B) indicate the number of strata. The x-axis scales in panels C–F were set to reflect the range used by 95% of all strata.

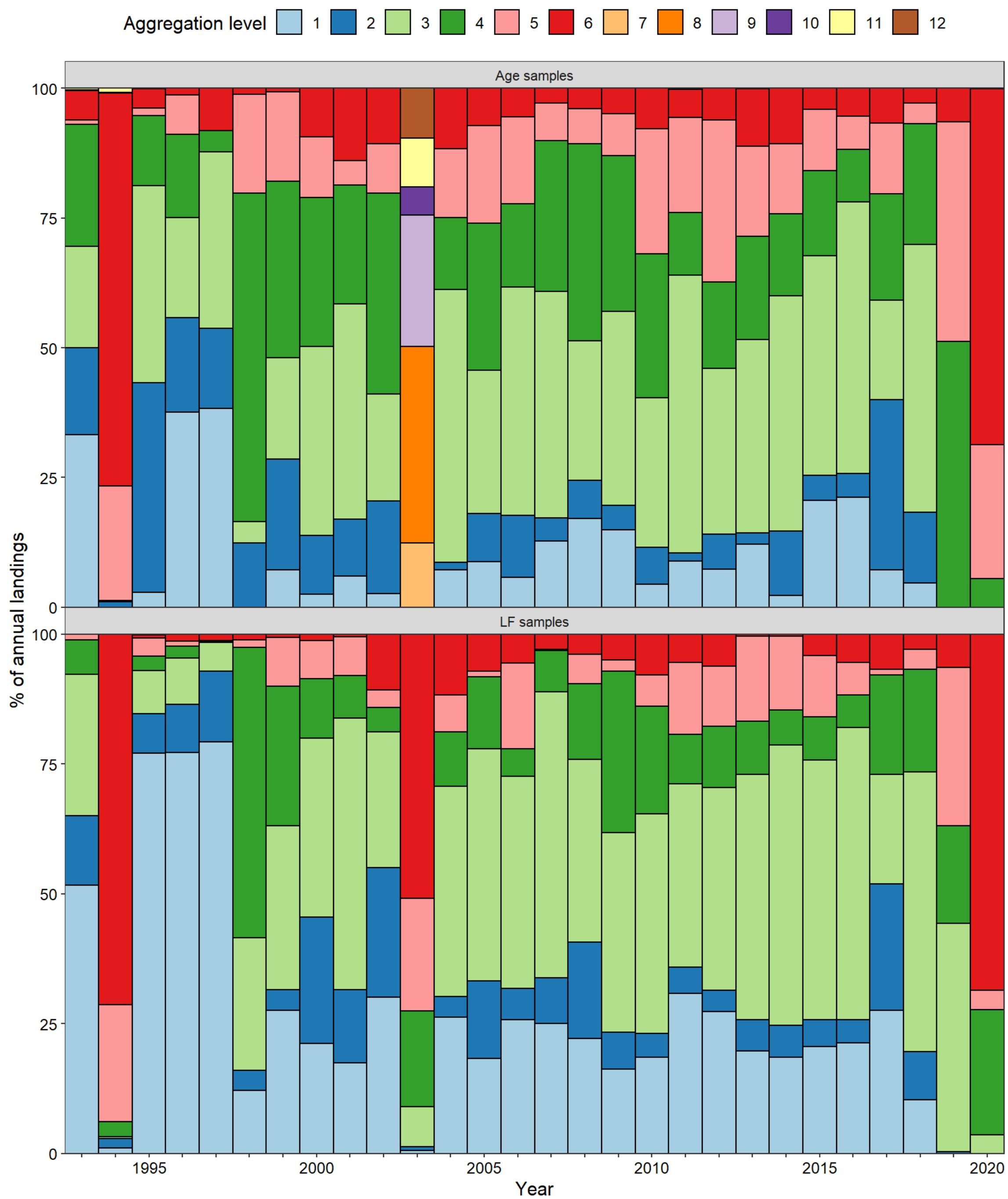

Figure 8. Percentage of annual landings for which the age (top) or length (bottom) composition was calculated from samples obtained from one of the 12 aggregation levels (see Equation 2).

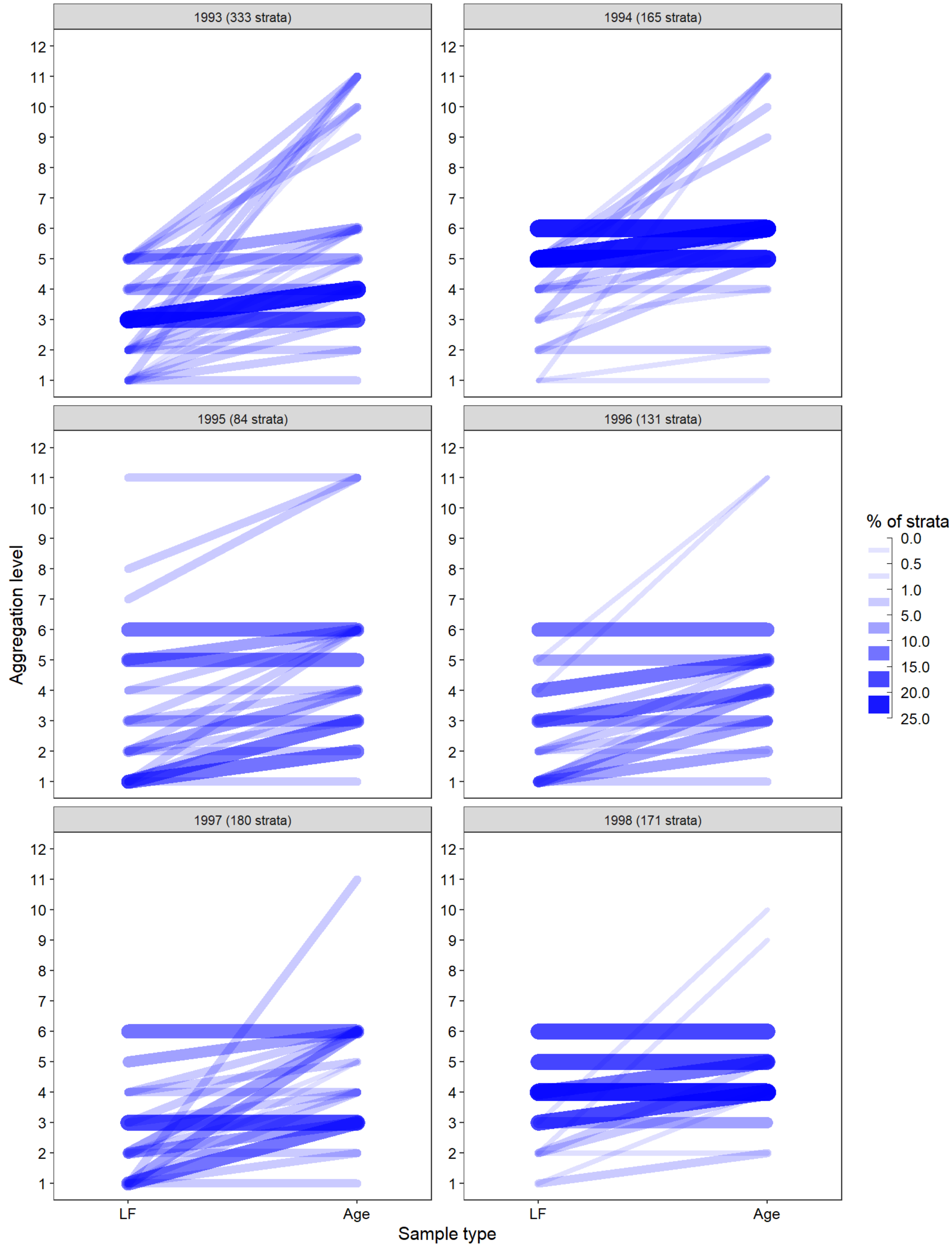

Figure 9. Percentage of annual strata that used a particular combination of aggregation levels (1–12) for length frequency (LF) and age data, for the 1993–1998 period. The size and opacity of each combination increases with its importance within the year.

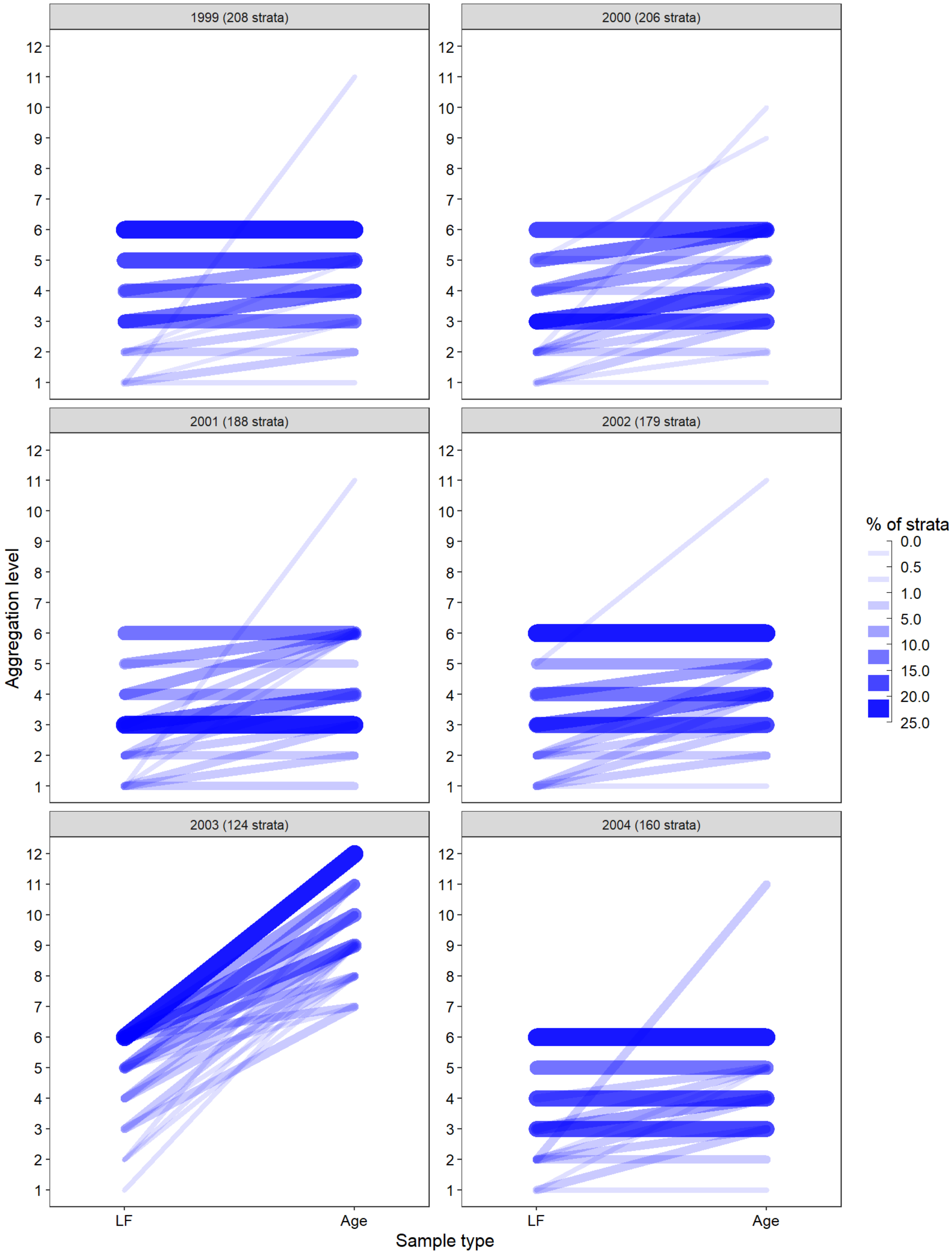

Figure 10. Percentage of annual strata that used a particular combination of aggregation levels (1–12) for length frequency (LF) and age data, for the 1999–2004 period. The size and opacity of each combination increases with its importance within the year.

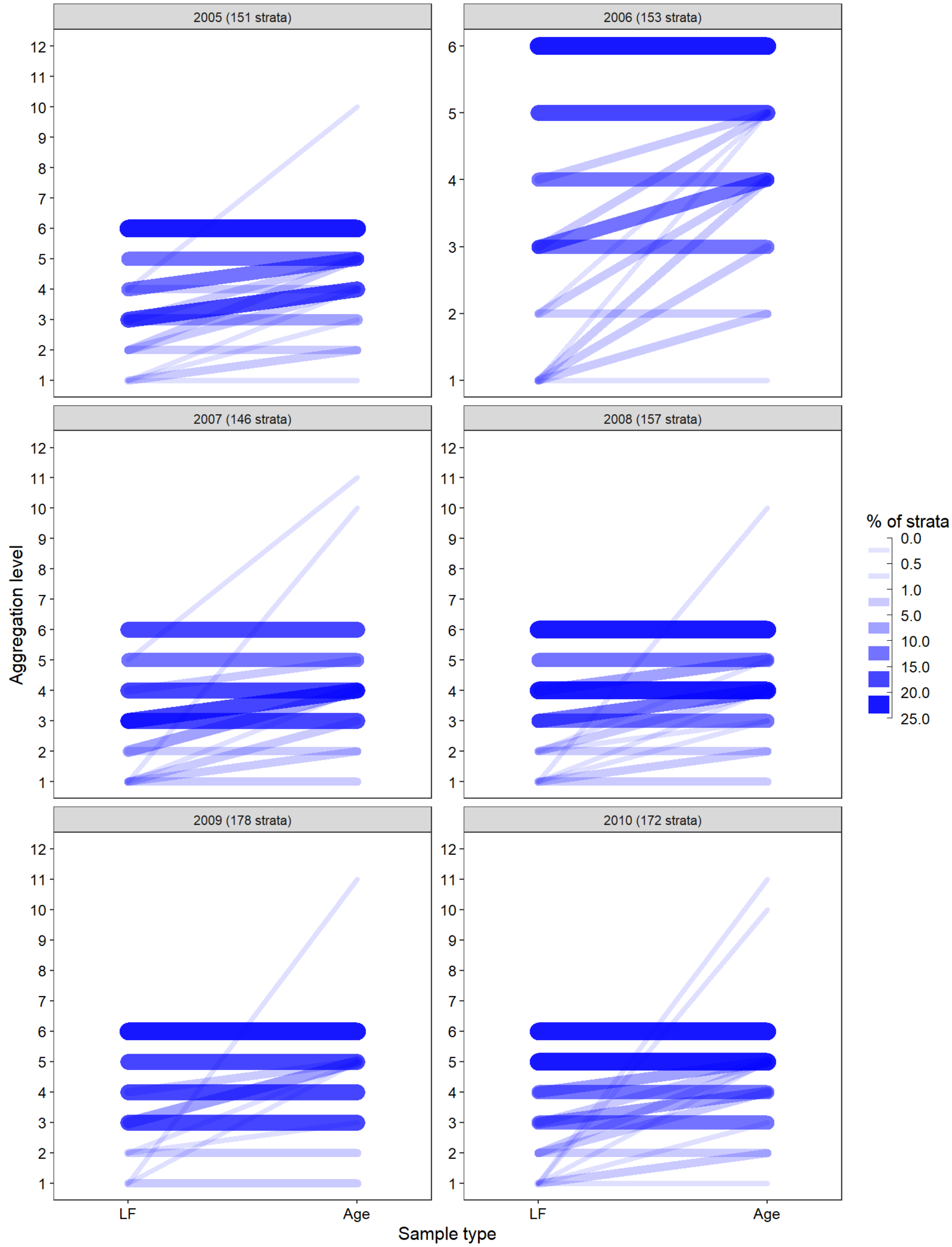

Figure 11. Percentage of annual strata that used a particular combination of aggregation levels (1–12) for length frequency (LF) and age data, for the 2005–2010 period. The size and opacity of each combination increases with its importance within the year.

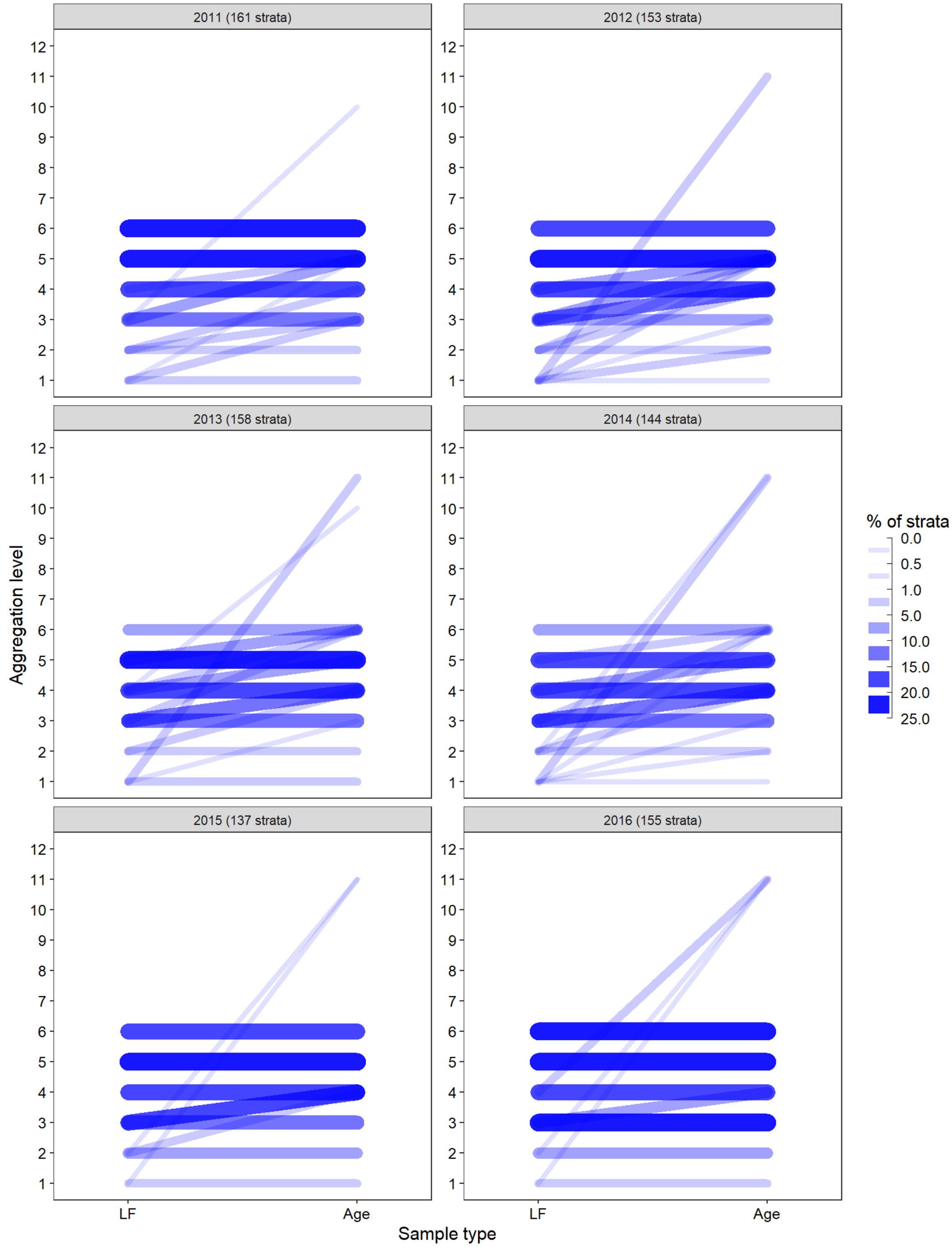

Figure 12. Percentage of annual strata that used a particular combination of aggregation levels (1–12) for length frequency (LF) and age data, for the 2011–2016 period. The size and opacity of each combination increases with its importance within the year.

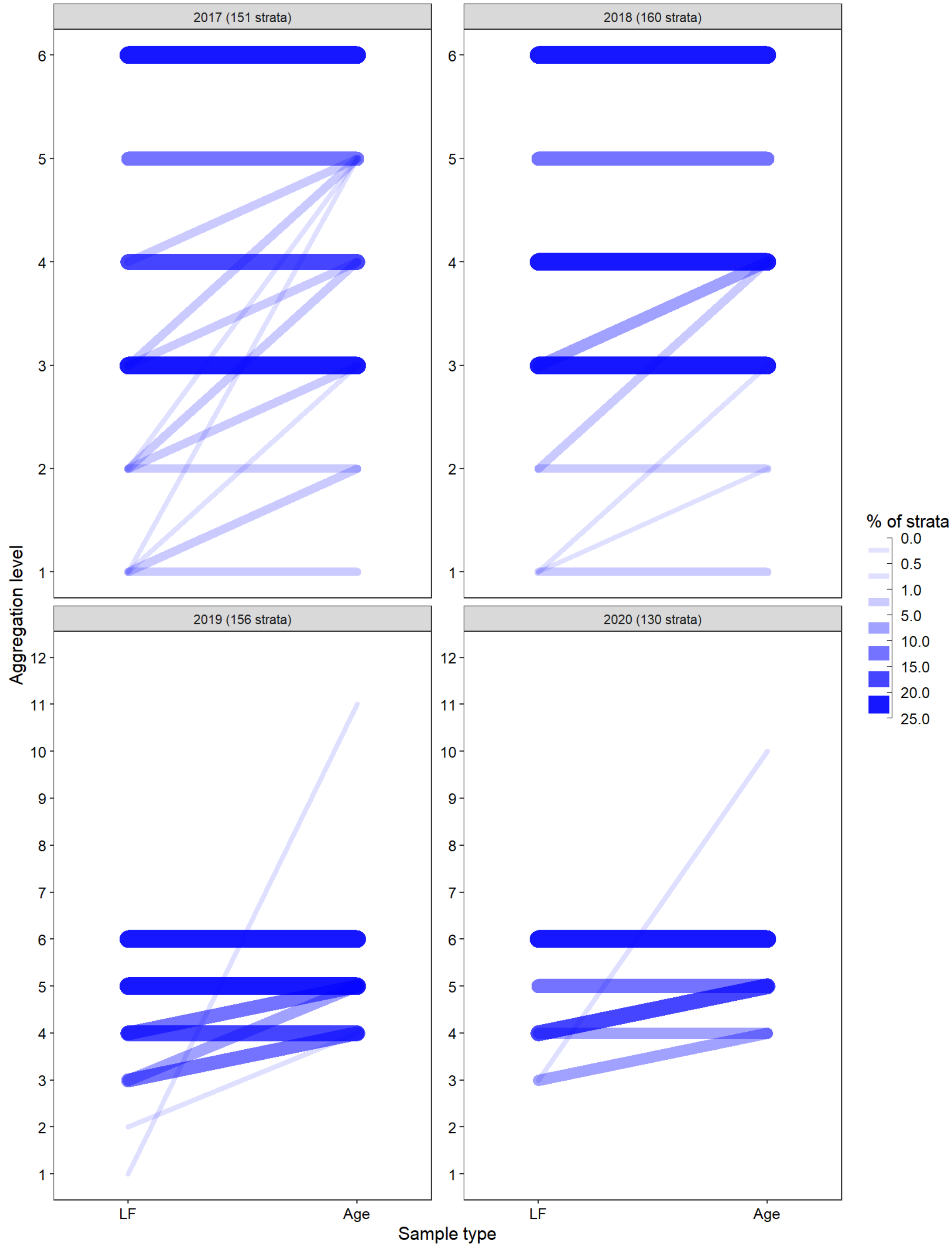

Figure 13. Percentage of annual strata that used a particular combination of aggregation levels (1–12) for length frequency (LF) and age data, for the 2017–2020 period. The size and opacity of each combination increases with its importance within the year.

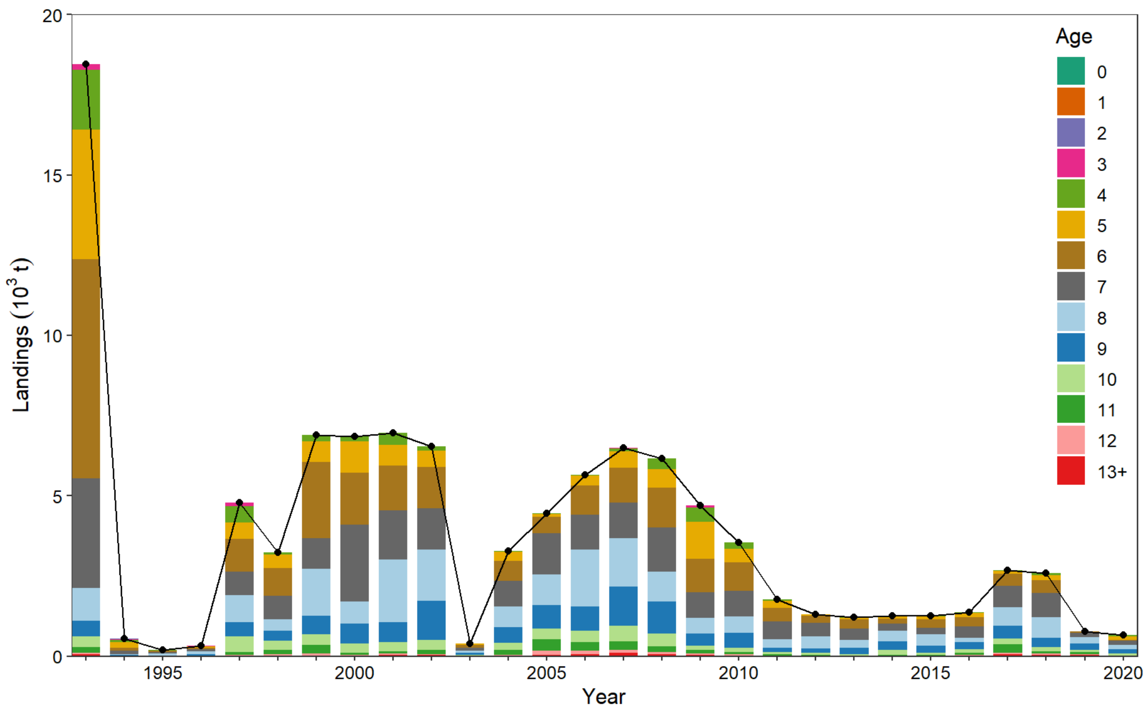

Figure 14. Correspondence between annual landings reported in the ZIFF database (black line) and annual landings at age.

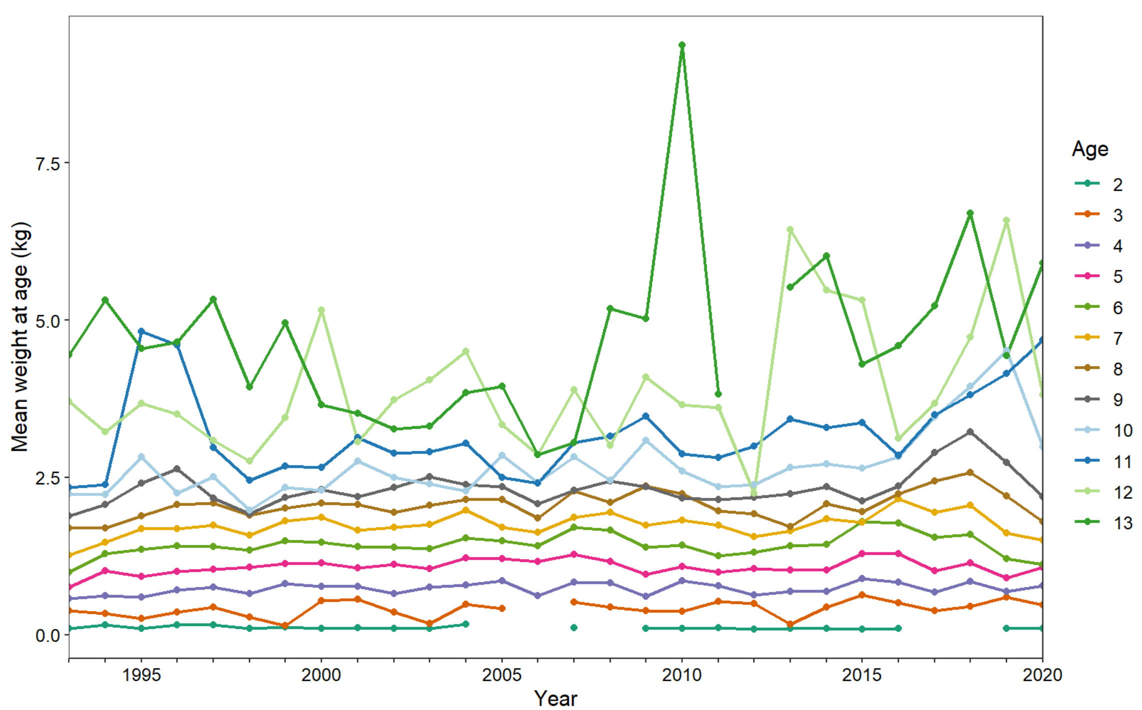

Figure 15. Mean weight at age (kg) for the 1993–2020 period, according to the new approach. Ages > 13 are not shown.

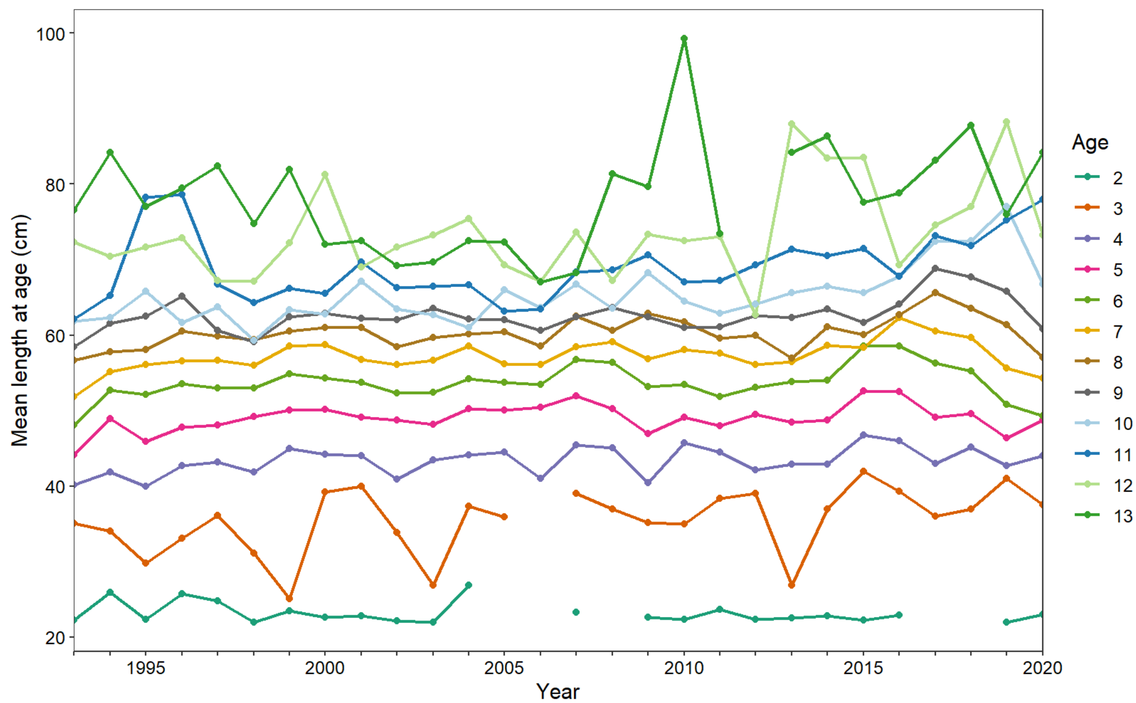

Figure 16. Mean length at age (cm) for the 1993–2020 period. Ages > 13 are not shown.

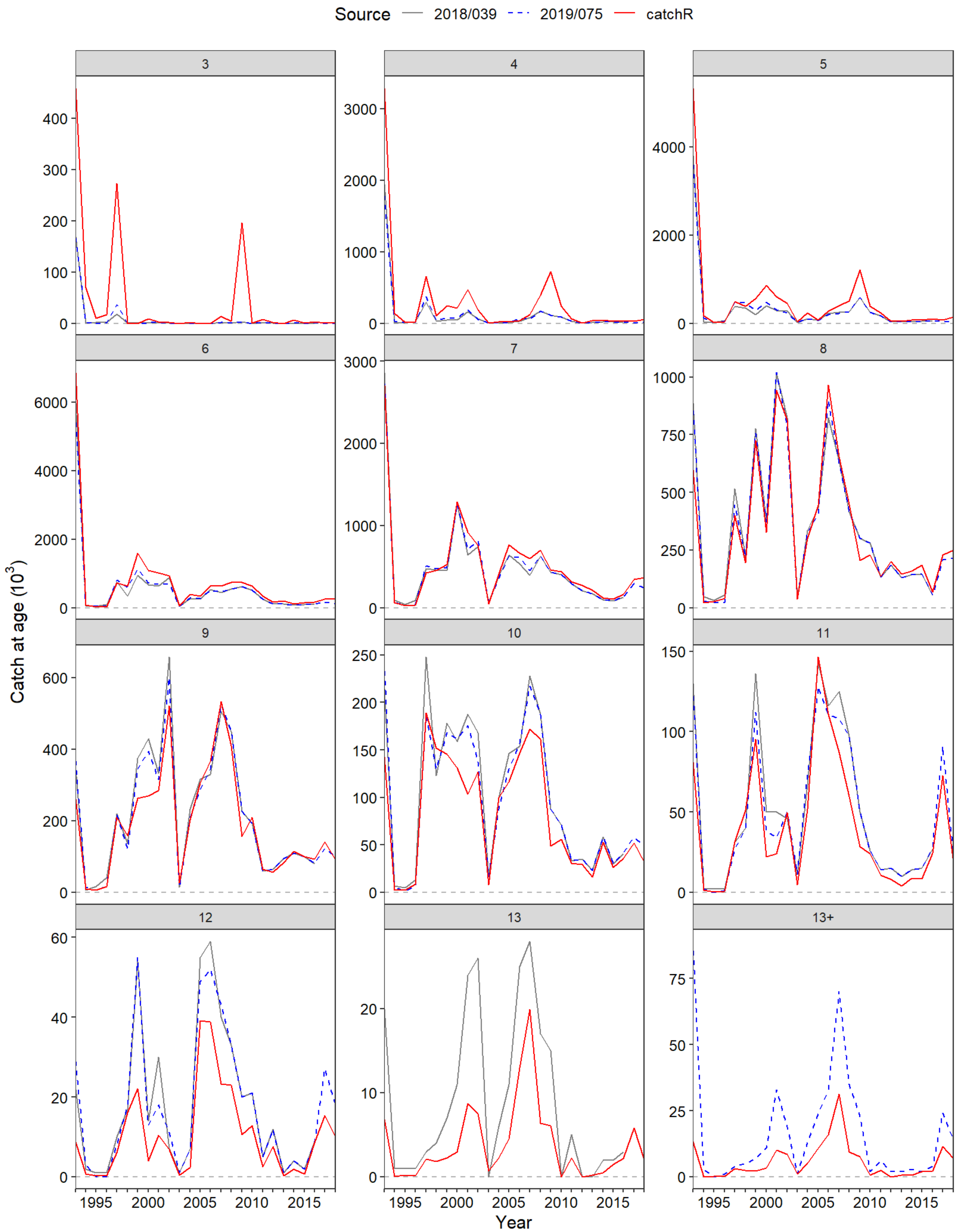

Figure 17. Catch at age numbers (in thousands) published in the last two research documents and according to the new approach, by age. Research document 2019/075 (Brassard et al. 2020) did not employ an age 13 group like the previous work (2018/039, Brassard et al. 2018), but rather a 13 + (≥ 13) group.

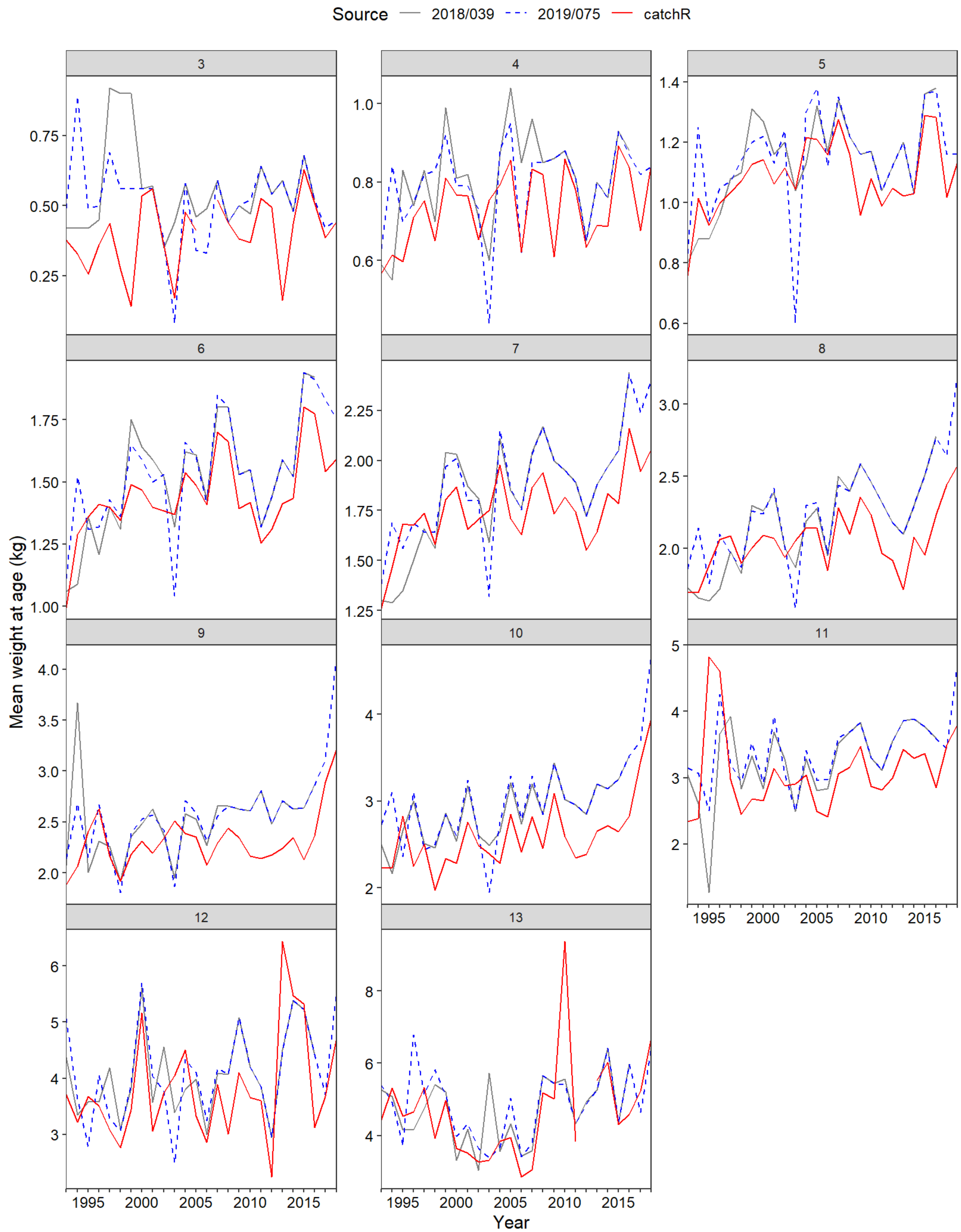

Figure 18. Mean weights at age (kg) published in the last two research documents and according to the new approach, by age.

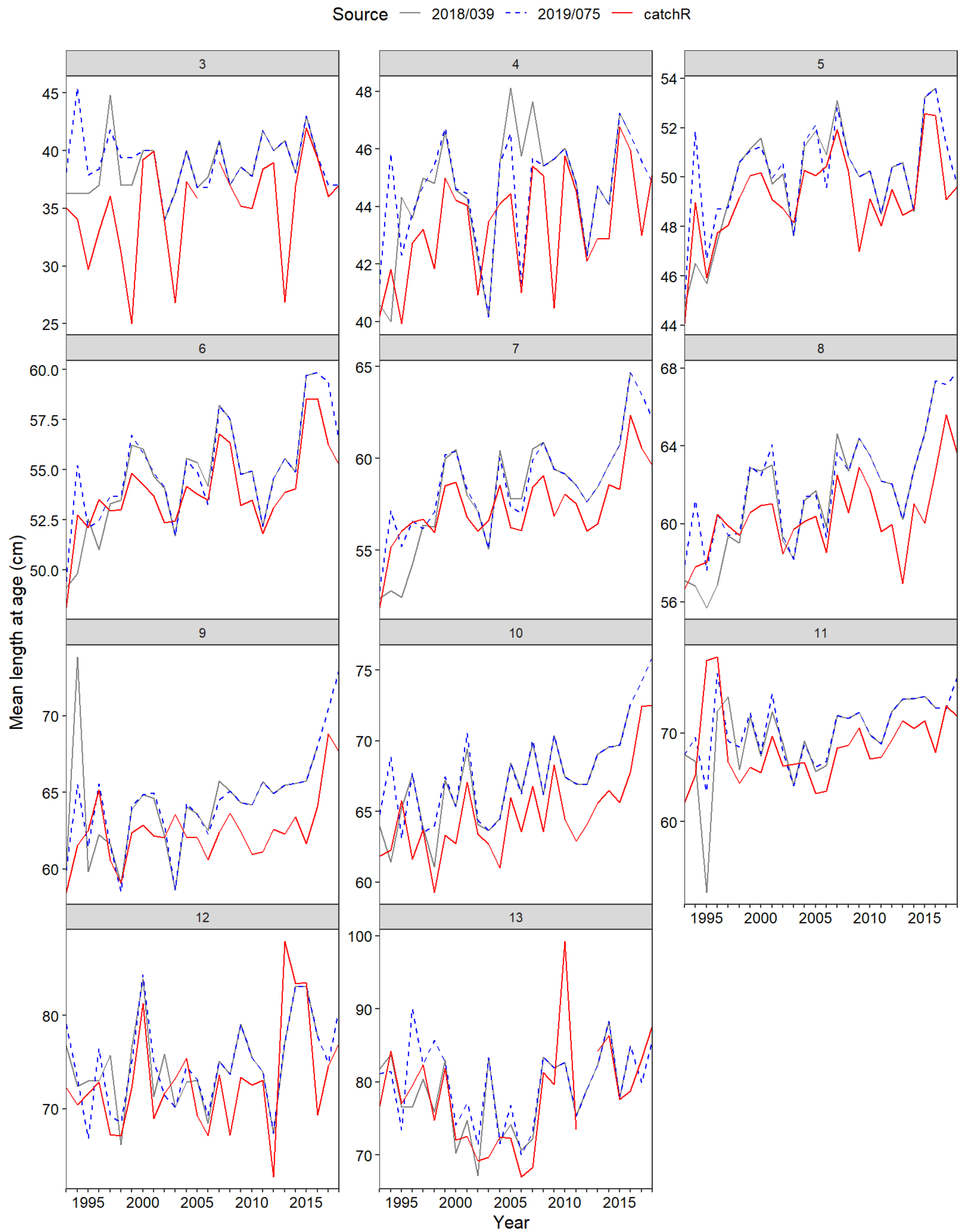

Figure 19. Mean lengths at age (cm) published in the last two research documents and according to the new approach, by age.

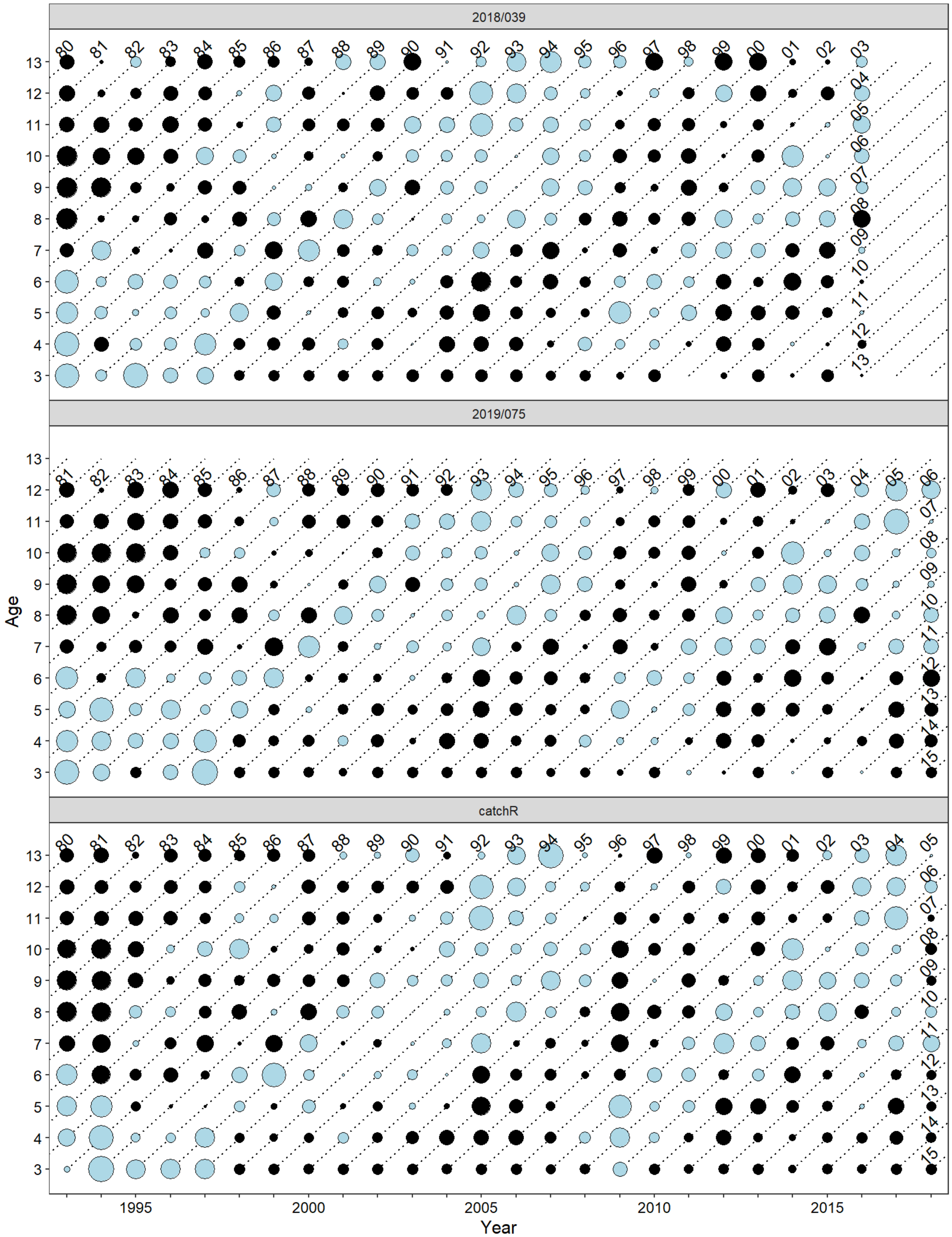

Figure 20. Standardized proportions at age and year (SPAY) according to values obtained from catchR and the two previous research documents, with blue and black bubbles indicating above and below average, respectively. The bubble size is indicative of the SPAY value. Cohort years' last two digits are indicated above bubbles from oldest ages and the ones from the most recent year.

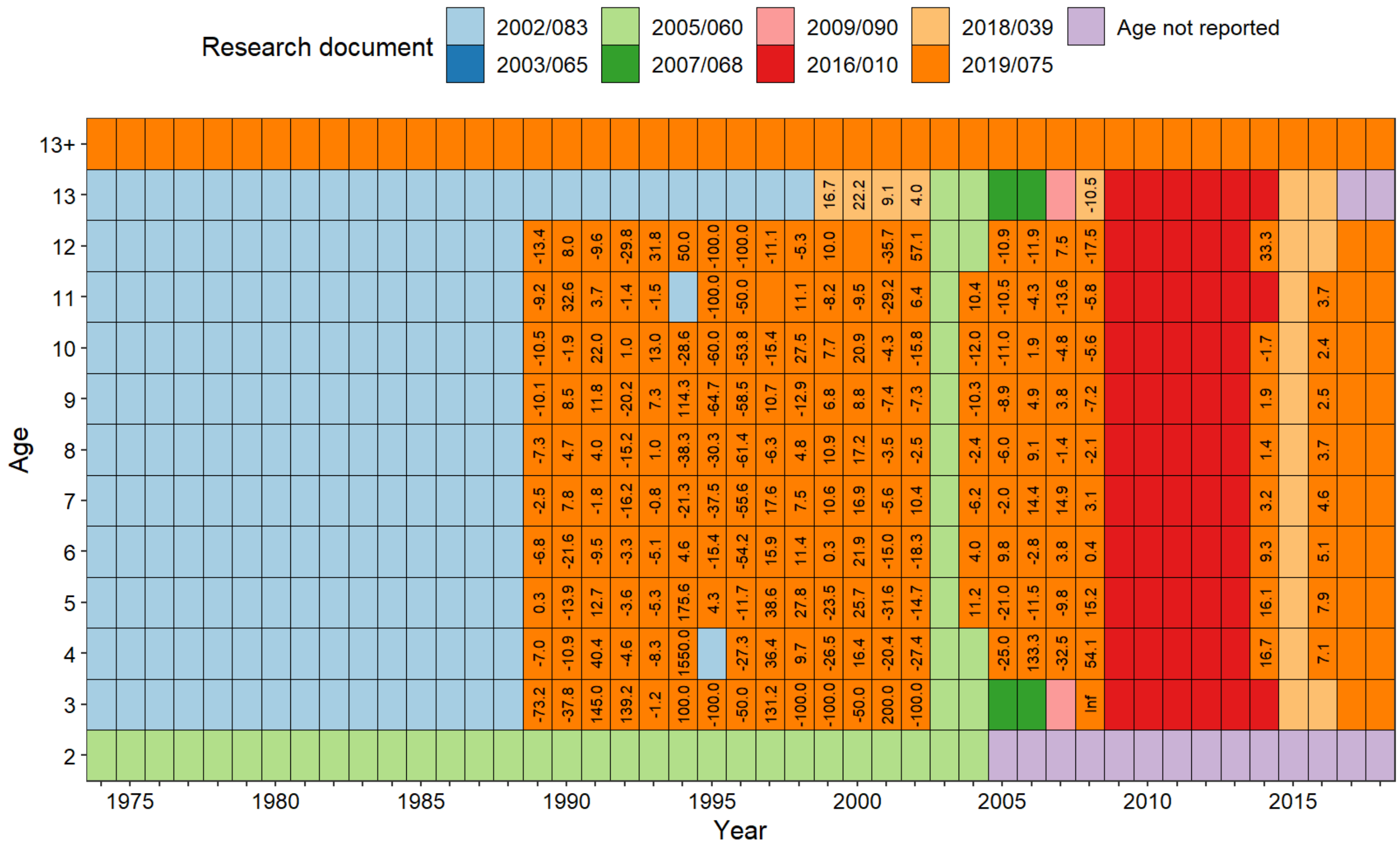

Figure 21. Evolution of the corrections made in each year-age cell in regard to catch-at-age numbers. The filling colour corresponds to the research document that last modified each cell. Values shown in cells represent the% difference between the first and last value used ((last - first) / first 100). Only research documents published after 2000 were considered.

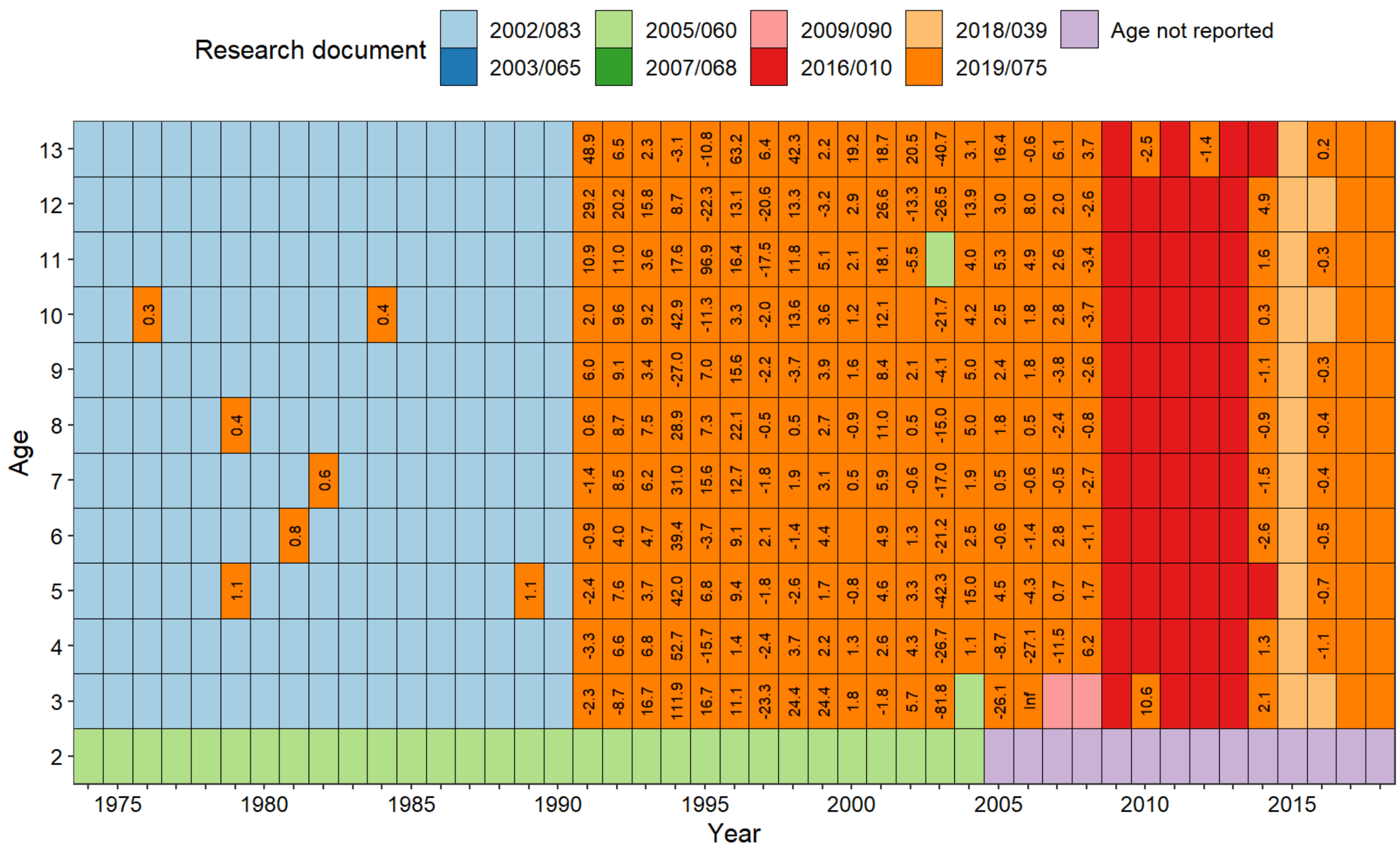

Figure 22. Evolution of the corrections made in each year-age cell in regard to mean weight-at-age. The filling colour corresponds to the research document that last modified each cell. Values shown in cells represent the% difference between the first and last value used ((last - first) / first 100). Only research documents published after 2000 were considered.

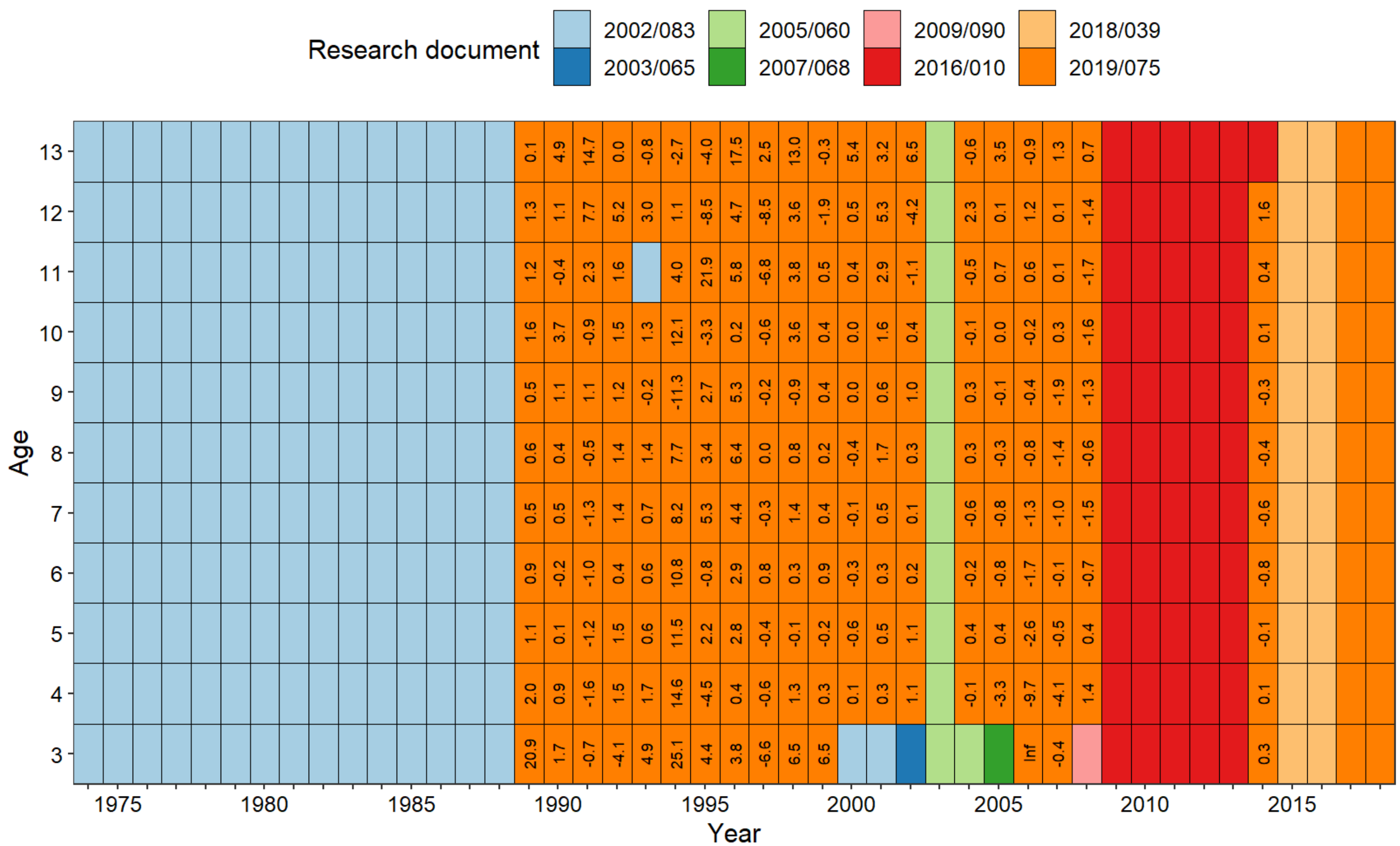

Figure 23. Evolution of the corrections made in each year-age cell in regard to mean length-at-age. The filling colour corresponds to the research document that last modified each cell. Values shown in cells represent the% difference between the first and last value used ((last - first) / first 100). Only research documents published after 2000 were considered.

APPENDIX A. *R* PACKAGE CATCHR

The *R* package `catchR` can be viewed and downloaded at:

https://github.com/iml-assess/catchR.

The `catchR` version used for this research document was version number 68ea9bf (commit id). To download the latest version of the package, type in the *R* console:

```
devtools::install_github("iml-assess/catchR")
```

To download the version used in this report, or any other version, use the argument `ref` of the function to enter the desired commit id. For example, running

```
devtools::install_github("iml-assess/catchR", ref = "68ea9bf")
```

would install the `catchR` version used for this report.

The analyses performed for this work can be viewed here:

https://github.com/iml-assess/catchR_morue

The rendered *.html* pages can be viewed here:

http://htmlpreview.github.io/?https://github.com/iml-assess/catchR_morue/blob/master/_site/index.html

APPENDIX B. LF/AGE DATA EXTRACTION USING SAS

Steps:

1. After opening the SAS program, in the menu bar, click on File, then on Open Program. Select the autoexec_peche.sas program located at S:/SAS/Peche/.
2. In the menu bar, click on Run, then on Submit. A window opens (Figure B.1).
3. Click on Répertoire de sortie pour le fichier d'extraction and choose a directory where the extracted data will be saved.
4. Click on Extraction.
5. Choose Atlantic cod as species (Figure B.2).
6. Depending on the data sought, check:
   a. LF: Entêtes-Fréq (Lignes 4 et 5).
   b. Age: Caractéristiques biologiques.
7. Click on Continuer.
8. Rename the file containing the extraction (Figure B.3). Ex: *cod_freq.dat* ou *cod_carbio.dat*.
9. For the years of interest, choose a range in années à extraire.
10. Leave the units of the variables as presented.
11. Click on Continuer. The extraction is complete.
12. Click on Sortie and then return to step 6 to extract a different type of data.

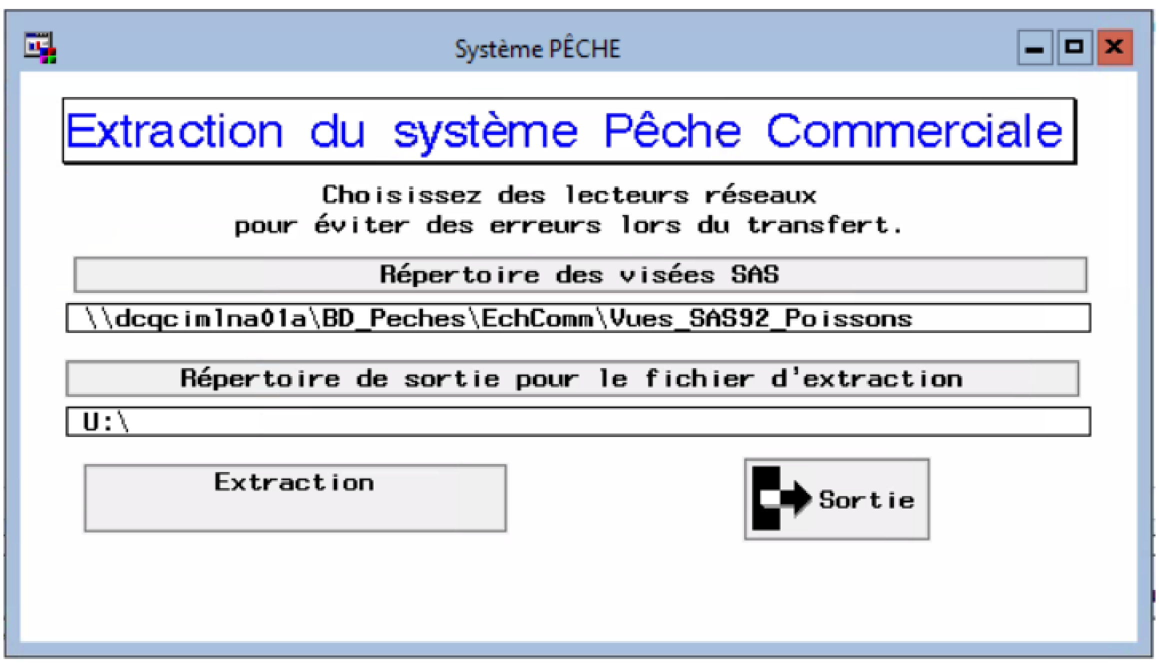

Figure B.1. Program autoexec_peche.sas. Home window.

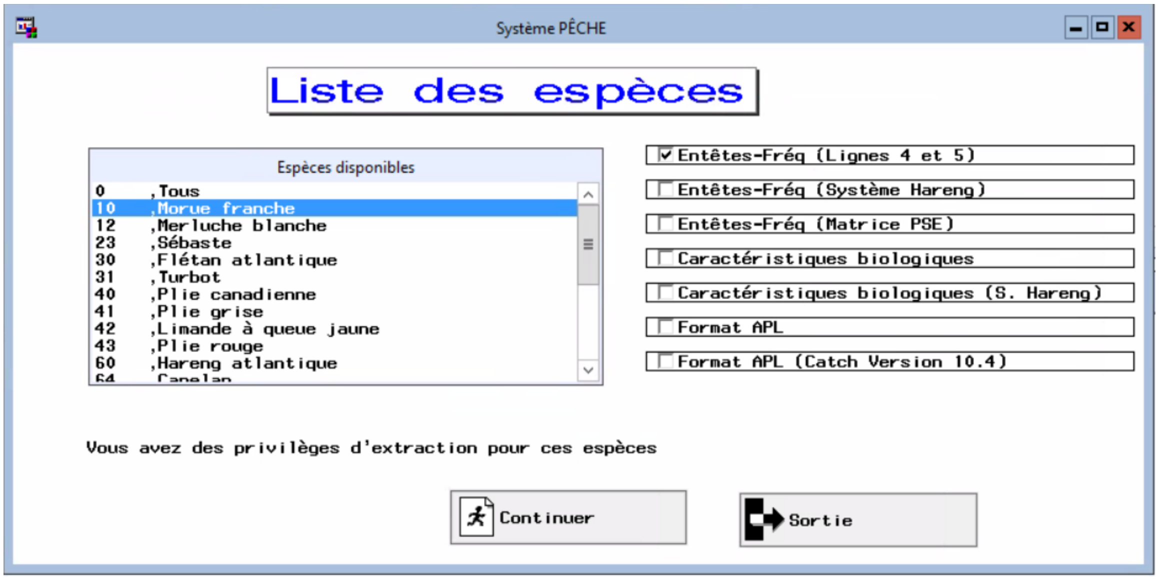

Figure B.2. Program autoexec_peche. sas. *List of species and choice of data to extract.*

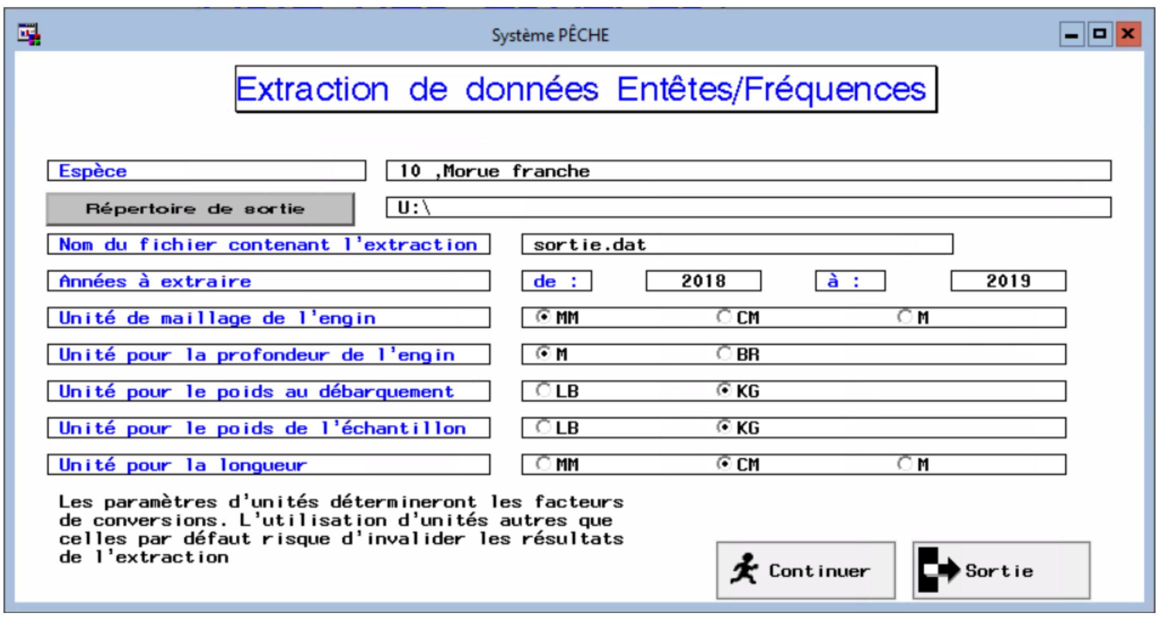

Figure B.3. Program autoexec_peche. sas. *Parameterization of the data to be extracted.*