Age Compositions of Red Snapper (*Lutjanus campechanus*) from the Southeast Reef Fish Survey Chevron Trap Survey Obtained Using Fourier Transform-Near Infrared Spectroscopy (FT-NIRS)

Walter Bubley and Claire Murdock

SEDAR90-DW-09

25 April 2025
Updated: 5 May 2025

This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

Please cite this document as:

Bubley, Walter and Claire Murdock. 2025. Age Compositions of Red Snapper (Lutjanus campechanus) from the Southeast Reef Fish Survey Chevron Trap Survey Obtained Using Fourier Transform-Near Infrared Spectroscopy (FT-NIRS). SEDAR90-DW-09. SEDAR, North Charleston, SC. 14 pp.

Age Compositions of Red Snapper (*Lutjanus campechanus*) from the Southeast Reef Fish Survey Chevron Trap Survey Obtained Using Fourier Transform-Near Infrared Spectroscopy (FT-NIRS)

Walter Bubley and Claire Murdock

Marine Resources Research Institute
South Carolina Department of Natural Resources
P.O. Box 12259
Charleston, SC 29412

SEDAR 90-DW-09

MARMAP/SEAMAP-SA Reef Fish Survey Technical Report #2025-08

Both the National Marine Fisheries Service Southeast Region Strategic Plan and the South Atlantic Fishery Management Council Snapper-Grouper Management Plan list goals of collecting quality data in a timely matter to support the fisheries (NOAA Fisheries 2020; SAFMC 2022). Assessment techniques such as interim analyses in waters of the southeastern United States are being considered, but these methods rely on more frequent data than are currently available using standard methods.

Data for life history has often been a bottleneck in assessments and with the implementation of operational assessments in the SEDAR process and the development of interim analyses in the region, the frequency of required data will only be increasing. Fishery-independent abundance index information and length compositions are available relatively soon after each sampling season but characterizing the demographics of the population is more difficult on an annual basis. Length frequencies are not as informative as age compositions in an age-based approach due to large overlaps in lengths between multiple age classes in some species. Age composition data are ideal but can lag when utilizing the traditional approach of visually assessing hard parts to obtain age estimates due to associated processing, interpretation, and species priorities being based on the assessment schedule. On average, age sample (otolith) processing and interpretation requires approximately 10 minutes per sample. This does not include embedding material drying time and mounting medium during the processing stage. Additionally, each sample is assessed by two readers, so there can be a substantial cost associated with at least 3 people being involved with the processing and reading of otoliths. Presently, because of these extensive time and personnel effort constraints, every species is not aged every year, making annual updates of age compositions for an index unrealistic for the large suite of species assessed in the South Atlantic.

Development of Fourier Transform-Near Infrared Spectroscopy (FT-NIRS) use was the recent subject of a NOAA Fisheries 5-Year Strategic Initiative to obtain relatively fast and efficient age estimates for fish, meaning annual age compositions could be feasible using this technology (NOAA Fisheries 2018). Because whole otoliths can be scanned using FT-NIRS, there is minimal prep work prior to the scanning procedure and FT-NIRS scans generally take less than 3 minutes per sample, including calibration curve creation, sample handling, and data input. Compared to traditional techniques to process and age a sample, the FT-NIRS technique provides the same information at a comparable quality in approximately 30% of the time, increasing potential throughput. Additionally, costs of personnel associated using FT-NIRS and equipment usage fees are still less than 40% of those required for using traditional methods. When applying these time savings and cost savings to the volume of life history data the Southeast Reef Fish Survey (SERFS) collects annually (> 10,000 individuals), it would allow the SCDNR Reef Fish Survey, a primary data provider to SEDAR stock assessments and potential interim analyses, to be more responsive to assessment schedule changes and interim analyses requests.

Methods

Sample Collection and Selection

- MARMAP/SERFS historically collected Red Snapper sagittal otoliths between 1990 and 2023 were utilized.
    - All have estimated ages using traditional methods.

- Only specimens with both otoliths collected were utilized for analysis, since one otolith may be sectioned to obtain an age estimate.
- In years during which fewer than 250 Red Snapper were collected, all otoliths were scanned using FT-NIRS. A subsample of 250 Red Snapper otoliths were scanned in years during which more than 250 whole otoliths were in the collection. This subsample included random samples collected from that year.

Near Infrared Scanning

- A Bruker Fourier-Transform Near Infrared Multi-Purpose Analyzer (MPA; Bruker Scientific, Billerica, MA, USA) housed within the South Carolina Department of Natural Resources Marine Resources Research Institute was utilized.
- All spectral data were collected in diffuse reflectance mode.
- All otoliths were cleaned and air dried prior to scanning.
- Whole otoliths were positioned convex-side down in the middle of the integrating sphere with the rostral axis positioned horizontally relative to the sample window (see Robins et al. 2015 for detailed description and pictures of scanning setup).
- A gold-coated transflectance cap was placed over the top of the structure to reduce stray light entering the detector.
- OPUS software (v. 8.2, Bruker Scientific, Billerica, MA, USA) was utilized for spectral acquisition.
- Each scan consisted of 64 spectral scans acquired for each otolith at a frequency of 16 cm^{-1} along the entire NIR spectrum (3600-12,000 cm^{-1}), with scans averaged to produce a single representative spectrogram for each sample.

FT-NIRS Data Processing and Analysis

- OPUS software was used for data processing and model generation.
  - The software allows for multiple combinations of spectral regions and data preprocessing techniques to be considered, selecting the optimal model.
- Calibration (Training) model
  - Samples were evenly divided into four subsets of data created by randomly assigning it so there was no overlap between the subsets.
  - Pre-processing
    - Wave number selection and data preprocessing treatments (transformations/derivitives, etc.) were compared to determine treatments and wavenumber ranges minimizing the RMSECV of predicted ages, resulting in optimized models capable of generating FT-NIRS-predicted ages from spectral data alone.
  - Multivariate spectral data were fit to 3 of the 4 subsets (75%) of the traditionally estimated ages using partial least-squares regression (PLS; Chen and Wang 2001).
    - In PLS, the information contained in the spectral data was compared to reference values for the component of interest (i.e., age) and changes that occurred in both matrices were correlated with each other.

        - The spectral data were broken down into principal components or factors (also called loadings) and scores, which allowed the most relevant spectral information to be retained while the "noise" or nonrelevant information in the data was disregarded. This process also allowed for the identification and isolation of spectral regions with the highest correlation with changes in the reference values (i.e., age).
        - Evaluations of the loadings and regression coefficients was used to determine the optimal number of factors (or rank) to include in the model without "overfitting" the data.
    - This process was repeated 3 additional times to create 4 training models in total. This was done to ensure that ages could be estimated from all samples, without utilizing a calibration model which included that particular sample for which an age is being estimated.
    - Cross-validation
        - Models were evaluated for age prediction capability using a "leave one out" method of cross-validation using a set of representative NIR spectra (the calibration samples) and corresponding reference values (in this case age) determined using traditional methods.
            - In "leave one out" cross-validation, each calibration sample is temporarily removed from the data set, a PLS model is created from the remaining samples, and the sample that was temporarily removed is predicted as an unknown. The difference between the reference value and the predicted is determined and the sample is returned to the data set, with the "leave one out" process being repeated until each sample has been removed once.
        - Upon completion, model goodness of fit was judged based on the R^2 (coefficient of determination), root mean square error of cross validation (RMSECV), and residual prediction deviation (RPD) values. RPD values of three or higher are generally accepted as "good" from a chemometrics standpoint (Williams, 2001).

- Testing Set
    - Once a cross-validated model was obtained, we further assessed the performance (robustness) of this model using the model to predict a completely new set of representative unknown test samples (test samples = the data subset (25%) not used to create the calibration model).
        - Only chevron trap collected otoliths were utilized in the testing sets
    - In this case, the entire test set was predicted by the model and the residual (NIR prediction versus reference value) for each sample that was calculated.
    - The mean error of prediction for the external validation set is calculated as the root mean square error of prediction (RMSEP). For age data, we also calculated the percent root mean square error (% RMSE), to evaluate standardized model error in the context of the maximum age included in the model (Couture et al. 2016; Passerotti et al. 2020a; Passerotti et al. 2020b).

$$\% \text{ RMSE} = \frac{\text{RMSEP}}{\text{Maximum Age}} * 100$$

  - This process was repeated 3 additional times with the associated training and testing sets so that all historic samples were assigned a predicted age.
- Validation of FT-NIRS age estimates relative to traditional ageing
  - Percent agreement and percent agreement within 1 year were calculated.
  - Average percent error (APE, Beamish and Fournier 1981) was used to assess the precision between age estimates.
  - Annual proportion by age of the following were overlaid to compare the age compositions
    - Historically derived age estimates from the SERFS chevron trap index
    - Historically derived age estimates from the subset of otoliths used to create the calibration curve
    - The FT-NIRS predicted ages from the subset of otoliths used to create the calibration curve.

Results

<u>Sample Collection and Selection</u>

- In total, there were 2,843 otoliths from which spectra were obtained between 1990 and 2023 (Table 1) and throughout the full SERFS chevron trap survey range (Fig. 1).
- Traditionally derived ages ranged from 0 to 19 years.

<u>FT-NIRS Data Processing and Analysis</u>

- Wave numbers containing the most information were selected and the data were transformed using the first derivative and vector normalization.
- Four calibration models were developed and used for age predictions of the associated testing sets (Fig. 2)
  - The average and range of the calibration model fit metrics are:
    - R^2 (coefficient of determination) = 92.7 (range: 91.6 – 93.2)
    - Root mean square error of estimation (RMSEE) = 0.92 (range: 0.87 – 0.99)
    - Residual prediction deviation (RPD) values = 3.70 (range: 3.5 – 3.8)
  - The average and range of the testing model fit metrics are:
    - R^2 (coefficient of determination) = 93.2 (range = 92.7 – 94.0)
    - Root mean square error of prediction (RMSEP) = 0.88 (range: 0.86 – 0.91)
    - Percent of root mean square error (%RMSE) values = 4.63 (range: 4.53 – 4.76)
- FT-NIRS predicted ages compared to the historically derived ages had an APE = 10.76 but showed some bias for ages >10 years (Fig. 3). Percent agreement of the predicted and the traditionally derived ages was 51%, with 95% agreement within one year.

- Age compositions were similar between historically derived age estimates from the subset of otoliths used to create FT-NIRS calibration models and those produced from the SERFS chevron trap index (Table 2, Fig. 4).
- Age compositions were similar using FT-NIRS predicted age estimates from the subset of otoliths used to create FT-NIRS calibration models and those produced from the SERFS chevron trap index and the FT-NIRS calibration models (Table 3, Fig. 4).

Literature Cited

Beamish, R. J., and Fournier, D. A. 1981. A method for comparing the precision of a set of age determinations. Canadian Journal of Fisheries and Aquatic Sciences, 38: 982–983.

Chen, J., and Wang, X. Z. 2001. A new approach to near-infrared spectral data analysis using independent component analysis. Journal of Chemical Information and Computer Science, 41: 992–1001.

Couture, J. J., Singh, A., Rubert-Nason, K. F., Serbin, S. P., Lindroth, R. L., and Townsend, P. A. 2016. Spectroscopic determination of ecologically relevant plant metabolites. Methods in Ecology and Evolution, 7: 1402–1412.

NOAA Fisheries. 2020. Southeast Geographic Strategic Plan 2020-2023. https://media.fisheries.noaa.gov/dam-migration/noaa_southeast_spupdate_30054_pm.pdf . Accessed 14 November 2022. 22 pp.

NOAA Fisheries. 2018. “Innovative Technology Promises Fast, Cost-Efficient Age Data for Fisheries Management.” www.fisheries.noaa.gov/feature-story/innovative-technology-promises-fast-cost-efficient-age-data-fisheries-management. Accessed 14 November 2022.

Passerotti, M.S., T.E. Helser, I.M. Benson, B.K. Barnett, J.C. Ballenger, W.J. Bubley, M.J.M. Reicher, and J.M. Quattro. 2020a. Age estimation of red snapper (Lutjanus campechanus) using FT-NIR spectroscopy: feasibility of application to production ageing for management. ICES Journal of Marine Science 77(6): 2144-2156

Passerotti, M. S., Jones, C. M., Swanson, C. E., and Quattro, J. M. 2020b. Fourier-transform near infrared spectroscopy (FT-NIRS) rapidly and non-destructively predicts daily age and growth of juvenile red snapper *Lutjanus campechanus* (Poey, 1860). Fisheries Research 223: 105439.

Robins, J. B., Wedding, B. B., Wright, C., Grauf, S., Fowler, A., Saunders, T., and Newman, S. 2015. Revolutionising Fish Ageing: using Near Infrared Spectroscopy to Age Fish. State of Queensland through Department of Agriculture and Fisheries.

SAFMC. 2022. "Snapper Grouper FMP Goals and Objectives." https://safmc.net/fishery-management-plans/snapper-grouper/ . Accessed 14 November 2022.

Williams, P.C. 2001. Implementation of near-infrared spectroscopy. Pages 145-169 in Near-Infrared Technology in the Agricultural and Food Industries. 2nd ed. P.C. Williams and K. Norris, ed. American Association of Cereal Chemistry, St. Paul, MN.

Table 1. Number of otoliths scanned with FT-NIRS by year to create calibration sets.

Year	Total Scanned
1990	1
1991	5
1995	4
1996	5
1998	11
1999	17
2000	14
2001	5
2002	33
2003	6
2004	5
2005	11
2006	5
2007	28
2008	28
2009	11
2010	88
2011	111
2012	203
2013	183
2014	194
2015	249
2016	252
2017	284
2018	259
2019	254
2021	287
2022	269
2023	21
Total	2,843

Table 2. Annual age compositions from age estimates obtained using traditional methods for samples used in FT-NIRS calibration models.

Age	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Totals
0	0.000	0.000	0.000	0.017	0.000	0.038	0.004	0.000	0.000	0.000		0.000	0	**13**
1	0.011	0.027	0.045	0.222	0.151	0.328	0.058	0.189	0.086	0.106		0.169	0.145	**339**
2	0.102	0.054	0.401	0.328	0.474	0.417	0.339	0.246	0.349	0.416		0.337	0.423	**845**
3	0.636	0.333	0.134	0.117	0.219	0.106	0.442	0.336	0.409	0.332		0.259	0.212	**681**
4	0.000	0.387	0.094	0.050	0.031	0.021	0.107	0.107	0.060	0.071		0.066	0.1	**204**
5	0.205	0.144	0.193	0.089	0.016	0.004	0.025	0.070	0.034	0.044		0.045	0.021	**150**
6	0.023	0.045	0.084	0.117	0.036	0.000	0.000	0.008	0.030	0.027		0.029	0.012	**77**
7	0.000	0.000	0.035	0.039	0.042	0.026	0.000	0.004	0.013	0.000		0.021	0.008	**39**
8	0.011	0.000	0.000	0.017	0.021	0.034	0.004	0.000	0.000	0.000		0.025	0.029	**30**
9	0.000	0.009	0.005	0.000	0.005	0.017	0.008	0.004	0.000	0.000		0.012	0.012	**16**
10	0.000	0.000	0.000	0.000	0.000	0.004	0.004	0.012	0.000	0.000		0.012	0.004	**9**
11	0.011	0.000	0.000	0.000	0.000	0.004	0.004	0.016	0.000	0.000		0.000	0.004	**8**
12	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.008	0.009	0.000		0.004	0	**6**
13	0.000	0.000	0.010	0.000	0.000	0.000	0.000	0.000	0.004	0.000		0.004	0	**4**
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004		0.012	0.004	**5**
15	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000	0.000	0.000		0.000	0.008	**3**
16	0.000	0.000	0.000	0.006	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.008	**3**
17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.004	0.008	**3**
18	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0	**0**
19	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000		0.000	0	**1**
Totals	**88**	**111**	**202**	**180**	**192**	**235**	**242**	**244**	**232**	**226**	**0**	**243**	**241**	**2436**

Table 3. Annual age compositions from age estimates obtained using FT-NIRS predicted samples.

Age	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	Totals
1	0.011	0.036	0.124	0.083	0.104	0.272	0.029	0.115	0.030	0.226		0.091	0.075	**262**
2	0.102	0.045	0.356	0.389	0.484	0.443	0.446	0.324	0.573	0.385		0.370	0.415	**950**
3	0.602	0.324	0.099	0.228	0.250	0.153	0.376	0.283	0.272	0.235		0.333	0.282	**659**
4	0.114	0.396	0.149	0.028	0.031	0.034	0.099	0.189	0.065	0.080		0.082	0.095	**249**
5	0.136	0.135	0.168	0.050	0.021	0.013	0.025	0.041	0.034	0.031		0.033	0.046	**127**
6	0.011	0.036	0.050	0.083	0.010	0.013	0.000	0.008	0.009	0.035		0.016	0.004	**52**
7	0.000	0.018	0.035	0.083	0.042	0.021	0.000	0.000	0.000	0.000		0.012	0.029	**47**
8	0.000	0.009	0.010	0.044	0.031	0.021	0.008	0.004	0.000	0.000		0.012	0.012	**31**
9	0.011	0.000	0.000	0.006	0.021	0.030	0.004	0.004	0.000	0.004		0.008	0.012	**21**
10	0.000	0.000	0.000	0.000	0.000	0.000	0.008	0.016	0.004	0.000		0.025	0.000	**13**
11	0.011	0.000	0.000	0.000	0.005	0.000	0.000	0.008	0.004	0.000		0.000	0.004	**6**
12	0.000	0.000	0.010	0.006	0.000	0.000	0.000	0.004	0.004	0.004		0.008	0.004	**9**
13	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.004	0.000	0.000		0.004	0.004	**4**
14	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.008	**2**
15	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.004	**1**
16	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000		0.000	0.000	**0**
17	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.004	0.000		0.004	0.004	**3**
Totals	**88**	**111**	**202**	**180**	**192**	**235**	**242**	**244**	**232**	**226**	**0**	**243**	**241**	**2436**

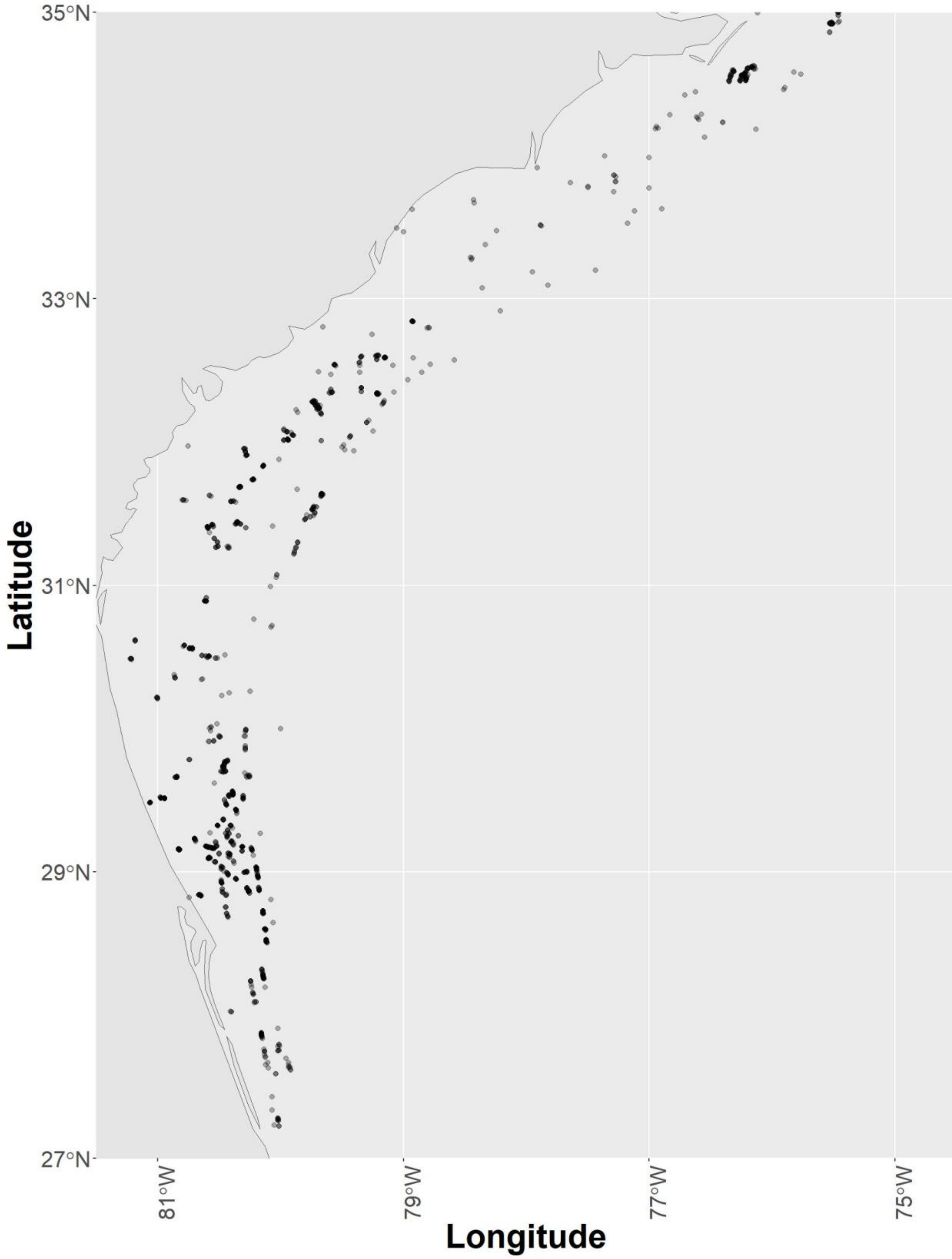

Figure 1. Map of locations from which otoliths were obtained from Red Snapper for FT-NIRS scanning.

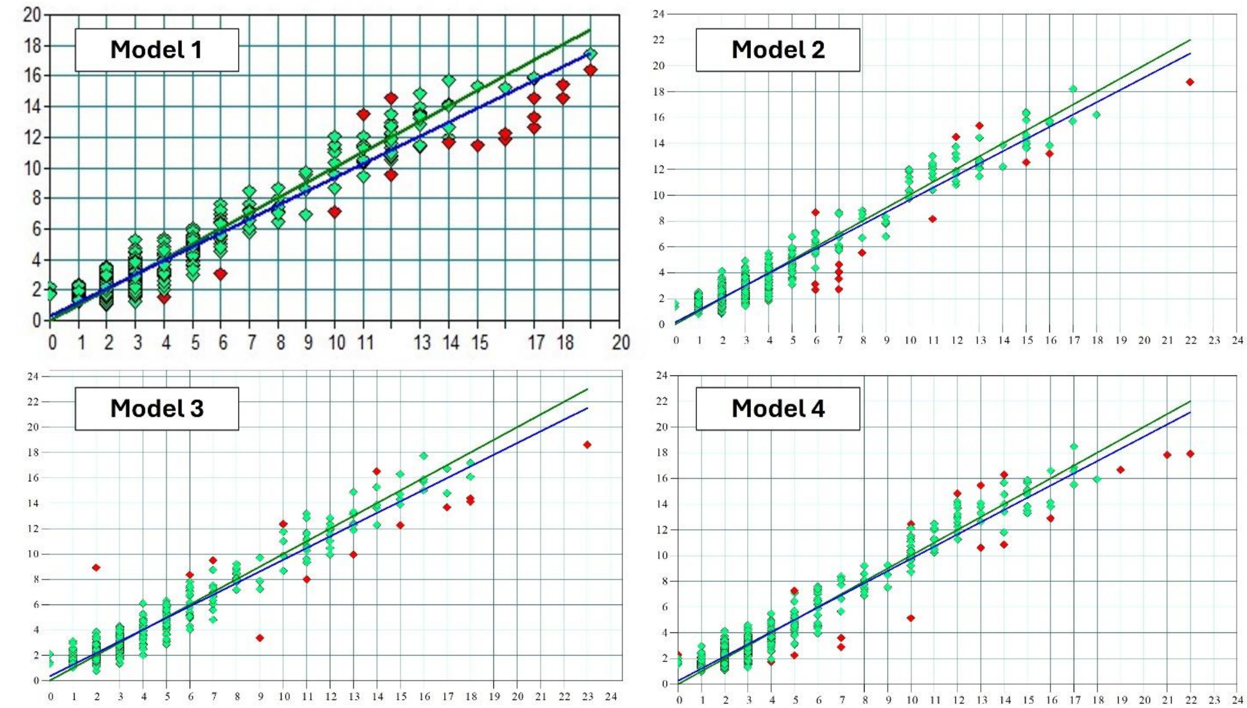

Figure 2. Comparing FT-NIRS generated age estimates from the testing sets with traditional age estimates for Red Snapper using the four calibration models. Increment count is on the x-axis. Model predicted age is on the y-axis. Green and red dots represent FT-NIRS estimated ages for individual fish. Red dots constitute outliers. The Green line is the one-to-one line and the blue line is the fitted line to the predicted ages.

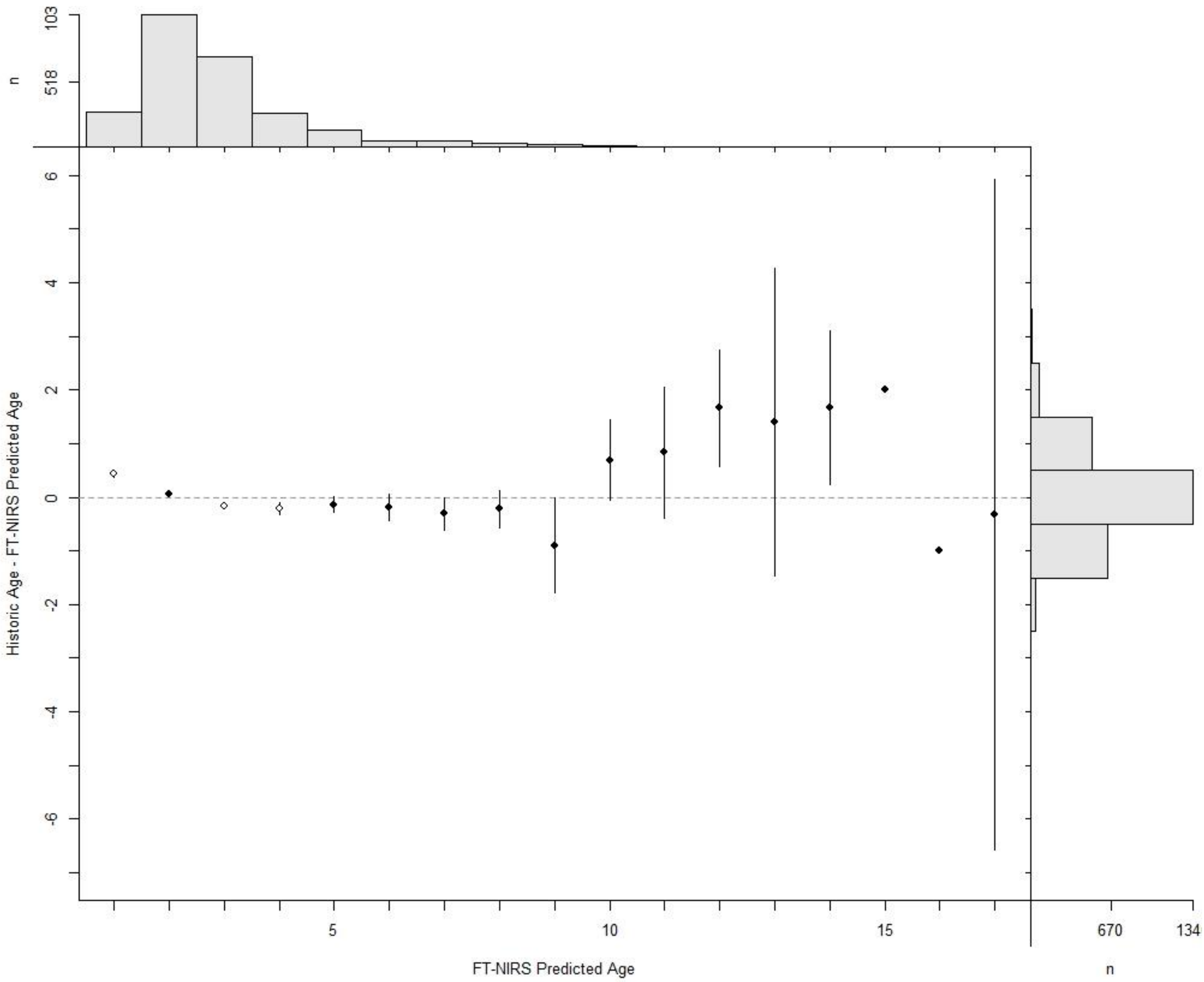

Figure 3. Distribution of error around traditionally derived ages when using FT-NIRS based predictions. Bar plots indicate sample sizes and points represent the mean deviation, with the whiskers being the 95% confidence intervals. Open circles indicate significant differences.

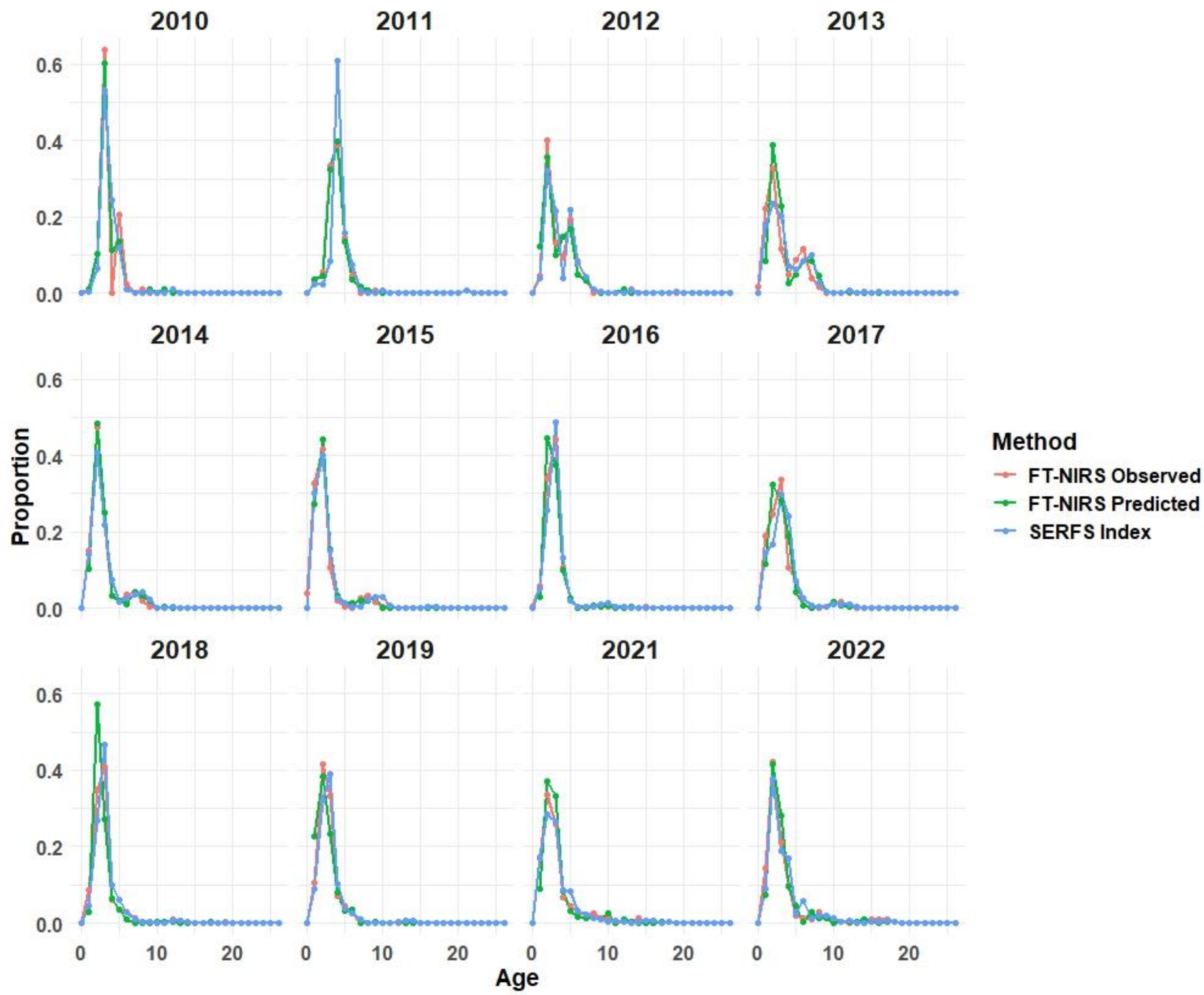

Figure 4. Annual proportion by age for the age composition used for the SERFS chevron trap index (SERFS Index), traditionally derived age estimates of the subset of otoliths used to create the calibration curve (FT-NIRS Observed), and the FT-NIRS predicted ages (FT-NIRS Predicted) from the subset of otoliths used to create the calibration curve.