Revised Estimates of Natural Mortality for Atlantic Menhaden

Sydney Alhale, Jeff Brust, Caitlyn Craig, Katie Drew, Brooke Lowman, Amy Schueller, Alexei Sharov

SEDAR102-RW-01

25 July 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:

Alhale, Sydney, Jeff Brust, Caitlyn Craig, Katie Drew, Brooke Lowman, Amy Schueller and Alexei Sharov. 2025. Revised Estimates of Natural Mortality for Atlantic Menhaden. SEDAR102-RW-01. SEDAR, North Charleston, SC. 49 pp.

SEDAR 102 WP-01: Revised Estimates of Natural Mortality for Atlantic Menhaden
M Working Group
March 2025

Executive Summary
The 2019 benchmark single-species assessment for Atlantic menhaden used an estimate of M based on the work of Liljestrand et al. (2019a). Liljestrand et al. (2019a) used a Bayesian mark-recovery model to estimate M and migration rates from an extensive tagging study conducted from 1966-1969 by the National Marine Fisheries Service. This study tagged over a million menhaden with internal ferromagnetic tags that were recovered by magnets installed in the reduction plants that processed harvested menhaden. Studies on tag shedding rates, tagging mortality rates, and magnet recovery rates were conducted as part of the tagging study. Liljestrand et al. (2019a) estimated an M of 1.17 for the fish tagged in that study. During the 2025 benchmark assessment process, Ault et al. (2023) submitted a working paper to the Atlantic menhaden Stock Assessment Subcommittee (SAS) and the Ecological Reference Points Work Group (ERP WG) that re-analyzed the historical tagging data and produced an estimate of M (0.56) that was significantly lower than Liljestrand et al. (2019a). However, Ault et al. (2023) had used a different subset of the data and a different approach to handling key parameters, which made direct comparisons with Liljestrand et al. (2019a) difficult. The SAS formed a working group (M WG) to consult with the authors and review the data and methods of both papers to understand the differences in the results and determine the best estimate of M.

The M WG determined that Liljestrand et al. (2019a) had overestimated the magnet efficiency rates for the Coston (1971) dataset. The M WG developed a revised time-series of magnet efficiency for use in the tagging model to correct this issue. In addition, the M WG found that the original time-series of confidential effort provided to Liljestrand et al. (2019a) could not be recreated by the Southeast Fisheries Science Center, and that the landings data for 1967 had been replaced with 1970 data in the original request. A reproducible time-series of confidential effort was developed and the landings data were corrected.

The M WG concluded the stepwise approach to modifying the magnet efficiencies used by Ault et al. (2023) was not appropriate, given how important this parameter (equivalent to the reporting rate in other tagging models) is to the tagging model estimates of M. While Ault et al. (2023) were able to fit the observed tag recoveries more closely using the adjusted magnet efficiencies, they did it at the expense of the observed magnet efficiency data. The adjusted magnet efficiency estimates varied widely from month-to-month but overall were lower than what was observed in the plant test data. This lower efficiency rate is what is driving the lower estimate of M from the Ault et al. (2023) analysis.

The M WG recommended using the estimate of M from the Bayesian mark-recovery tagging model using the Coston (1971) data with the revised magnet efficiency estimates and updated confidential effort and landings data as the base run. The lower estimate of M from the Ault et al. (2023) stepwise magnet efficiencies and the confidential effort data will be used as a

sensitivity run. For both runs, a vector of M-at-age for the assessment model was developed by scaling the Lorenzen (1996) curve so that M at age-1.5 was equal to the point estimate of M from the tagging model, based on the size of tagged fish.

Introduction

NOAA Fisheries Menhaden Tagging Study, 1966-1969

NOAA Fisheries Southeast Fisheries Science Center (SEFSC) initiated a large-scale tagging study of Atlantic menhaden in July 1966 (Dryfoos et al. 1973). Atlantic menhaden were captured by commercial purse seines and pound nets as well as fishery independent sampling (Ahrenholz et al. 1991) and tagged with internal ferromagnetic tags injected using hand-held tagging guns (Carlson and Reintjes 1972; Kroger and Dryfoos 1972). Tagged fish were released only off the coast of North Carolina in 1966, but from 1967-1969, releases occurred over the full range of the Atlantic menhaden fishery from Massachusetts to northern Florida. A total of 1,066,378 fish were tagged over those 3.5 years, making it one of the largest mark-recapture studies ever conducted (Liljestrand et al. 2019a).

Tags were recovered on magnets installed in the reduction plants where menhaden landings were processed. Magnets were installed in 17 of the 18 plants that were in operation at this time. The original researchers conducted complementary studies to estimate tag shedding and tagging mortality rates, as well as studies to estimate the magnet efficiency at each plant (Kroger and Dryfoos 1972; Dryfoos et al. 1973).

Coston (1971) summarized the releases and recoveries on primary magnets by month and region for 1966-1969, but the original digital records were lost during a data transfer in the 1990s. The surviving paper copies of the records were re-digitized in the early 2010s, referred to here as the NMFS re-digitized dataset to distinguish it from the summarized Coston (1971) dataset.

A review of the historical literature on this study is included in Appendix 1 to this report.

Liljestrand et al. (2019a) Bayesian Mark-Recapture Model

Liljestrand et al. (2019a) developed an instantaneous rate version of the Brownie dead recovery model that used Bayesian techniques to estimate fishing mortality, natural mortality, and the probability of movement between regions from the NOAA Fisheries' tagging data. They considered both the Coston (1971) summarized dataset and the NMFS re-digitized dataset, and chose to use the Coston (1971) dataset since the NMFS re-digitized dataset was incomplete and missing records were not distributed equally over time and space (see "Comparison of Release and Recapture Data" below and Table 1 for more details).

The model incorporated empirical estimates of tag shedding and tagging mortality from the studies conducted contemporaneously by Kroger and Dryfoos (1972). Estimates of magnet efficiency (equivalent to the reporting rate in other tagging models) were recalculated from the raw plant efficiency study data that were re-digitized along with the tag-recapture data.

Confidential fishing effort data at the regional, monthly level were used to inform the model estimates of fishing mortality.

The model operated on a monthly time-step, although October-May was treated as one period for the movement rate estimation, since fishing effort was low or non-existent during these months and therefore, so were recaptures. Liljestrand et al. (2019a) also condensed two of the geographic areas reported in Coston (1971). Coston (1971) broke the data out into five areas: Area 1 (MA-NY), Area 2 (NJ-DE), Area 3 (MD-VA), Area 4 (NC-SC), and Area 5 (FL-GA). Due to lower sample sizes in the northern two areas, Liljestrand et al. (2019a) combined the data from Coston's (1971) Area 1 and Area 2 into a single region. The regions described in this report follow the naming conventions of Liljestrand et al. (2019a) (Figure 1):

- Region 1 (MA-DE)
- Region 2 (MD-VA)
- Region 3 (NC-SC)
- Region 4 (FL-GA)

Liljestrand et al. (2019b) simulation-tested the model framework and found the model was able to estimate natural mortality accurately and precisely with very little bias; movement rates were also estimated accurately, but accuracy was improved in scenarios with more spatially and temporally uniform fishing effort and higher numbers of releases.

Liljestrand et al. (2019a) estimated an M of 1.17 yr^{-1}, which was significantly higher than previous estimates which ranged from 0.50-0.53 yr^{-1} (e.g., Dryfoos et al., 1973; Reish et al., 1985). The average age of tagged fish was 1.35 years, so the 2020 benchmark assessment scaled the estimates of M-at-age calculated using Lorenzen (1996) so that the estimate of M at age 1.5 was equal to Liljestrand et al.'s (2019a) estimate of 1.17.

<u>Ault et al. (2023) Re-Analysis of the Tagging Data</u>

During the 2025 benchmark assessment process, Ault et al. (2023) submitted a working paper to the Atlantic Menhaden SAS and the ERP WG that re-analyzed the historical tagging data and produced an estimate of M (0.56) that was significantly lower than Liljestrand et al. (2019a). Ault et al. (2023) argued that this lower value is more biologically realistic because it is in line with historical estimates of M for Atlantic menhaden, including estimates based on life history meta-analyses. The methods and conclusions of Ault et al. (2023) are described in detail in a working paper provided by Ault et al. (SEDAR 102-RW-02 Ault et al. Re-analysis of Tagging Data).

The SAS and the ERP WG reviewed the paper and found there were a number of differences in the input data and methods that made it difficult to evaluate what was causing the discrepancy in the estimate of M.

The SAS identified the following differences in the input and methods:

- **Release and recapture datasets**: Liljestrand et al. (2019a) used the Coston (1971) data, while Ault et al. (2023) used the NMFS re-digitized dataset.
- **Effort data:** Liljestrand et al. (2019a) used a confidential time-series of effort data in vessel-weeks at the region and month level, while Ault et al. (2023) were not granted access to the confidential data and had to derive a time-series of effort at the region and month level from coastwide, non-confidential data.
- **Magnet efficiency estimates:** Liljestrand et al. (2019a) calculated plant-level estimates of magnet efficiency from the plant-test dataset and then developed monthly, regional estimates of magnet efficiency by taking the average of the plant-specific estimates for each region, weighted by the proportion of landings from each plant within the region. Ault et al. (2023) used a "stepwise" approach where the estimates of magnet efficiency from the plant-test data were adjusted outside the model to improve the model fit to the recapture data and the model was re-run with the adjusted magnet efficiencies until a minimum likelihood was reached.

Both analyses used the same tagging model, as Liljestrand et al. provided the source code and input file for their model to Ault et al., as well as to ASMFC.

The SAS formed a working group to review the data and methods of both papers to understand what was driving the differences in the results, in consultation with the authors, and determine the best estimate of M.

Comparison of Release and Recapture Datasets

The Coston (1971) and NMFS re-digitized represent slightly different subsets of the complete tagging dataset from this experiment. Specifically, the Coston (1971) dataset, which is summarized to total number of releases by month and region, is presented at a coarser resolution than the NMFS re-digitized dataset which is at the individual tag level. In addition, the Coston (1971) dataset only includes information on tag releases and recoveries through 1969, while the NMFS re-digitized dataset has information on releases through February 1970 and recoveries through February of 1971.

However, the NMFS re-digitized dataset is not complete and is missing releases that are present in Coston (1971). From 1966-1969, Coston (1971) recorded 1,066,357 tag releases, while the NMFS re-digitized dataset only includes 768,977 releases, meaning the NMFS re-digitized dataset is missing approximately 28% of the records included in the Coston (1971) dataset. Furthermore, the missing records are not distributed evenly across time or space (Table 1). For example, from 1966-1969, the NMFS re-digitized dataset included 92% of the Coston (1971) releases in the MD-VA area, with data in 22 of the 23 months that Coston (1971) reported, but only 2% of the Coston (1971) releases in the NJ-DE area, all from one month, while Coston (1971) had releases in 12 months (Table 1).

About 50% of recaptures occurred within 80-90 days after release, at which point recovery slowed down considerably, and about 80% of recaptures occurred within one year of release, and nearly all recaptures occurred within two years (Figure 2).

Ault et al. (2023) noted discrepancies between the Coston (1971) data and the data presented by Liljestrand et al. (2019a). After consultation with Liljestrand et al. and a review of the Coston (1971) data and the input and output files from their model runs, the M WG determined that although the supplementary tables presented in the Liljestrand et al. (2019a) paper were incorrect, Liljestrand et al. (2019a) had used the correct Coston (1971) data in the actual model runs that produced their results.

M WG Decision

The M WG recommended using the Coston (1971) dataset for the base run of the tagging model. The Coston (1971) dataset is more spatially complete and will provide better information on movement rates of Atlantic menhaden as well as M. Spatially explicit assessment and ecosystem models are a high priority for management, and it is important that estimates of movement and M are consistent for future models.

Landings and Effort Data

The raw commercial reduction fishery landings and effort data files for 1966-1971, housed at the SEFSC, were converted from DTA files to CSV and uploaded to SQL for the following analysis. The raw data file was compared to the files provided to Liljestrand et al. and Ault et al., as well as the SEDAR 69 data. The files were filtered to species = 0 to exclude Atlantic thread herring.

Landings

The landings in the raw files were compared on a yearly basis to Liljestrand et al. (2019a), Ault et al. (2023), and SEDAR 69. The annual totals corresponded well, except the 1967 total in the Liljestrand et al. (2019a) file. It was later discovered that the file provided 1970 data as 1967 data, which explained the discrepancy. Other years had a 0.01 to 1 percent difference, which was not considered significant for this analysis.

Effort

The effort data were represented in vessel weeks, which was calculated by counting the number of weeks a vessel had at least one effort (i.e., fishing time on the water during the week, regardless of catch). The fishing week (fweek) was calculated by using the WEEKNUM function in Excel to determine the week of the year the set date fell into, with weeks starting on Monday. The fishing week was then used to determine the effort each vessel had per week and was then summed across vessels by week to determine vessel weeks.

The vessel weeks reported in SEDAR 69, as well as the Liljestrand file, were significantly less than the vessel weeks calculated from the raw file. Unfortunately, the calculation method used to generate the time-series for SEDAR 69 and Liljestrand et al. (2019a) was not well documented, so the methods could not be compared, and the SEDAR 69/Liljestrand et al. (2019a) time-series could not be reproduced. The method used for this analysis is the current method used in the stock assessment and menhaden reports and is reproducible.

The monthly landings and effort were shared with Ault et al., without area or region information, to preserve data confidentiality. The number of trips (days with landings) were also shared. The Ault et al. (2023) analysis used the nominal effort by area by year reported in Liljestrand et al. (2019a) and converted vessel weeks to portions of effort by area by year as a proxy to calculate a ratio of vessel weeks per trips. The ratio was then used to convert the originally provided monthly data (Ault et al. 2023).

The number of trips were recalculated for the M WG using the raw data files and matched relatively well, with a 1-2 percent difference.

<u>Differences between Liljestrand et al. (2019a) and Ault et al. (2023) datasets</u>
The major difference observed between the analyses was the file structure of the data provided. Liljestrand et al. were provided effort and landings by area by plant by month. Ault et al. were provided with monthly effort and landings because of confidentiality constraints. Ault et al. used the Liljestrand et al. vessel weeks to create a proxy for effort by area by year using the number of trips provided to them.

<u>M WG Decision</u>
The M WG recommended using the confidential time-series of landings and effort developed for this assessment. Using the finer scale confidential data eliminates a source of uncertainty in the back-calculations from non-confidential data, and the new estimates are reproducible for future assessments and correct previous errors.

Magnet Efficiency Estimates
Tags were recovered on magnets installed in the reduction plants where menhaden landings were processed. Understanding how effective these magnets were at recovering tags was critical, as that was essentially the reporting rate for this study: if a tagged fish was recaptured, how likely was it that the tag would be found and reported? Therefore, the original study included an extensive plant test study where tagged menhaden were seeded into the catch at active plants and recoveries of these known tagged fish were recorded (Kroger and Dryfoos 1972).

The plant test data were re-digitized along with the release and recovery data. A total of 964 batches of tags were released into 19 plants from 1966-1971 (Table 2), representing nearly 96,000 fish across all four model regions. Batches were comprised of 100 tagged fish; in some cases, the tagging gun malfunctioned and not all one hundred tags in the batch were used, but the number of tags successfully inserted was noted for each batch. Not all plants were tested in all years, but all plants had multiple tests conducted. There was only one active plant on the coast that reported both landings and tag returns but was not tested.

<u>Primary vs. secondary magnets</u>
Reduction plants often had magnets already installed to remove scrap metal from the catches, and additional magnets were installed in support of this project. Magnets were categorized as “primary” magnets and “secondary” magnets. The tags passed through the primary magnets

immediately after landing or as the processed meal was transferred from the reduction plant to the scrap shed for drying. Secondary magnets detected tags when the final product was transferred to trucks for transport to the buyer. Kroger and Dryfoos (1972) concluded that time of capture could be accurately determined for tags recovered on primary magnets, but secondary magnets recovered the tags too late in the reduction process to accurately determine the time of capture. As a result of this finding, Coston (1971) only reported recoveries by month on primary magnets.

The re-digitized datasets included information on which magnet each tag was recovered on. However, it was not clear from the historical reports which magnets were considered primary and which were secondary. The M WG consulted with J. Smith, a retired SEFSC employee who was more familiar with the final stages of the tagging program in the late 1970s, and compared the number of recaptures reported in Coston (1971) and the NMFS re-digitized dataset for 1966, a year where the number of releases matched exactly in both datasets in order to identify the primary magnets. From this, the M WG determined that magnets labeled “1” and “2” in the re-digitized datasets were considered primary magnets, and all other magnets were considered secondary.

A comparison of the timing of recovery on magnet types agreed with Kroger and Dryfoos’s (1972) conclusions, with the median time between the release and recovery of the plant test tags was 2 days for primary magnets and 13 days for secondary magnets. The 95th percentile of time in the plant was 26 days for primary magnets (i.e., less than a month) and 148 days for secondary magnets, meaning non-negligible numbers of tags could be recovered on secondary magnets weeks or even months after the fish was captured.

Based on the review of the literature and the examination of the re-digitized data, the M WG determined that Liljestrand et al. (2019a) had calculated the magnet efficiency estimates incorrectly for the Coston (1971) dataset. Liljestrand et al. (2019a) had calculated the magnet efficiencies by plant using the recoveries from all magnets in the plant test dataset, but the release-recapture matrix from Coston (1971) only included recoveries on primary magnets. Therefore, using all magnets in the magnet efficiency estimates would overestimate the efficiency of each plant.

<u>Revised magnet efficiency estimates</u>

The M WG recalculated the plant-specific magnet efficiency estimates to align correctly with the Coston (1971) dataset by using only recoveries on primary magnets (i.e., magnets 1 and 2).

The M WG decided to pool the plant test data across years and months for each plant, to maximize the overall sample size for each plant. The recovery rate for primary magnets was variable, ranging from as low as 0% to as high as 100% for individual batches, but most plants had some kind of central tendency, not a uniform distribution (Figure 3). The M WG estimated the mean magnet efficiency rate for each plant using a negative binomial distribution, as Liljestrand et al. (2019a) did. That is, for each batch, a, and plant, p, the likelihood of recovering x individuals from a batch of n releases was predicted using a negative binomial distribution and the magnet efficiency parameter for that plant, ε_p, which was estimated by minimizing the negative log likelihood of observed and predicted data. The magnet efficiency rate varied from plant to plant, even within regions, so in order to develop estimates of regional magnet efficiency at the year and month time-step of the model, the M WG followed the approach of Liljestrand et al. (2019a) and calculated an average regional, monthly magnet efficiency from the magnet efficiency rates of each plant, weighted by the proportion of the menhaden landings from each plant for that year, month, and region. This reflects the fact that not all plants within a region were active in all months, or processed the same volume of landings in every month over the time-series, and prevents the regional average from being skewed by plants that had a high number of samples in the plant test data, but passed fewer menhaden through their magnets overall.

From 1966-1969, there were only three plants that reported landing menhaden but did not have magnet efficiency data. Two plants did not have magnets, so the magnet efficiency rate was set to zero for those plants when calculating the regional average. The third plant did recover tags, but based on the NMFS re-digitized dataset, those recaptures all occurred on a secondary magnet, and so would not have been reported in Coston (1971), making the magnet efficiency rate for that plant zero as well. Some plants in the plant test data were not active from 1966-1969 and reported no landings, so those rates were effectively dropped from the regional means. When landings did not occur for a month in a region, or no attempts were made to clean the magnets and recover the tags (as occurred in some regions at the start of the project), the magnet efficiency rates were set to zero.

Unlike Coston (1971), the NMFS re-digitized dataset included recoveries on all plant magnets, so the M WG also developed plant-specific magnet efficiency estimates to use with the NMFS re-digitized dataset that included recoveries on all magnets. The distribution of magnet efficiency rates on all magnets were generally similar to the rates on the primary magnets alone, with more of a central tendency towards higher efficiency rates (Figure 4). The same negative binomial approach was used to estimate plant-specific magnet efficiency rates, and the same landings-weighted approach was used to develop annual, monthly, regional magnet efficiency rates. The revised magnet efficiency rates for the Coston (1971) dataset developed by the M WG were lower than what Liljestrand et al. (2019a) used (Figure 5).

Ault et al. (2023) magnet efficiency estimates

Ault et al. (2023) did not have access to the confidential landings by plant and month, so they used the plant test data to estimate regional magnet efficiencies by year and month. They then took a "stepwise" approach to adjust the annual, monthly magnet efficiency estimates outside the model so the tagging model could more closely match the observed recaptures. The adjusted magnet efficiencies were constrained to be between 0.1 and 0.9, but no other prior distribution was applied. They also explored allowing the model to fit magnet efficiency directly and obtained lower estimates of M than through their stepwise approach or the fixed efficiency approach, which they did not consider plausible.

The adjusted estimates of magnet efficiency were more variable from month to month than the estimates developed by the M WG (Figure 6), but overall, the distribution of their adjusted efficiencies skewed lower than the distributions of the observed plant efficiency data at the regional level (Figure 7).

M WG Decision

The M WG recommended using the revised magnet efficiencies developed from observed plant test data using the corrected methods of Liljestrand et al. (2019a). There is not enough information in the tag-recapture data to estimate both M and the magnet efficiency rate, as these parameters are too highly correlated. Ault et al. (2023) were able to improve the fit of the tagging model to the observed recaptures, but at the expense of the observed magnet efficiency data: the median of their adjusted estimates of magnet efficiency were lower for all regions than what was observed, their distributions were skewed low compared to the observed data, and their adjusted values were frequently at the lower bound of the uniform distribution they imposed.

Ault et al. (2023) argue that the adjusted magnet efficiencies better capture the variability in the plant test data than the M WG's method, but if that were the case, the distribution of their adjusted estimates should still match the distribution of the plant test data. Both a visual inspection (Figure 7) and the Kolmogorov-Smirnov test indicated that the adjusted magnet efficiency estimates come from a significantly different distribution than the plant test data ($p < 0.001$ for the pooled dataset and $p<0.05$ for all regions).

Tagging Model Results

Model specifications and base run

The base run of the tagging model included the following data components: 1) releases and recoveries from the Coston technical report, 2) confidential effort data, and 3) magnet efficiencies based on primary magnets and weighted by confidential landings. The base run configuration was the same as the configuration in Liljestrand et al. (2019a), only the data input components differed. The model contained four regions, as described above, and 42 monthly time steps to reflect the data.

The parameters estimated from the model were the same as those from Liljestrand et al. (2019a) and included the log of natural mortality on the monthly time step, log catchability (q) for each region, theta for each region and monthly time step that had effort or recoveries, and the migration rate parameters for each time step between regions. The annual value estimated for natural mortality was obtained by taking the exponential of the estimate and multiplying that by 12 months.

The uncertainty surrounding natural mortality was determined using an MCMC analysis for the base run and with data inputs at the upper and lower confidence intervals for the maximum likelihood estimate. Each MCMC was run 4,000,000 times with chains saved every 1,000 runs and the first half of the chain removed as the burn in period.

Sensitivity runs

Sensitivity runs were used in order to explore multiple avenues of investigation. The first avenue was to provide a continuity of the model used to estimate M from the last assessment to the updated data decisions used to estimate M for this update assessment. This was done in a stepwise fashion with three total intermediary sensitivity analyses. The first sensitivity analysis used the updated confidential effort data. The second sensitivity analysis used the updated magnet efficiencies using only primary magnets and weighted by the confidential landings data. The final sensitivity analysis in this series used the updated Coston release and recovery information, which differed by two tags in the recovery data.

The second set of sensitivity runs assessed the differences that would occur if the NFMS re-digitized data were used, as opposed to the Coston data. The analysis helped to delineate the consequences of the spatial coverage differences in the data, in particular the dearth of data for the northernmost region. Two sensitivity runs were completed with the NMFS re-digitized data, one with the time series the same as the Coston data set including 1966-1969 and one with the full time series of data for the NMFS re-digitized data including 1966-1971. The use of the NMFS re-digitized data required a change in the number of theta parameters for the model given the differences in recoveries in the dataset and increase of 14 months of data including 1970 and January and February of 1971. These sensitivity runs used magnet efficiency rates based on all magnets, rather than only the primary magnets.

Finally, two additional sensitivity runs were completed using the stepwise method proposed by Ault et al. (2023). The first run was provided by Ault et al. and used the Coston dataset with the non-confidential effort time-series. The magnet efficiencies were the stepwise-adjusted values that minimized the log-likelihood with the non-confidential effort data. The second run was developed by the SAS by replacing the non-confidential effort data with the confidential time-series in the Ault et al. input file and applying the stepwise process until the log-likelihood was minimized or the model failed to converge. These sensitivity runs evaluate the impact of using the fixed, empirically estimated magnet efficiencies in the base run vs. using the stepwise-adjusted approach to estimate magnet efficiency.

While other sensitivity runs were completed during the exploration of these data options, those differences were resolved by finalizing data; thus, the full outputs of those runs are not included in this writeup.

Results – Base run

The base tagging model converged and the Hessian inverted. The overall fits to the observed recovery data were good and balanced high and low residuals across months and regions (Figure 8). The annual instantaneous fishing mortality was estimated as the highest for Region 3 (NC-SC), followed by Region 1 (MA-NY) (Figure 9), while the Chesapeake Bay region (Region 2 MD-VA) and the most southern region (Region 4 GA-FL) had lower estimated annual fishing mortality rates. The estimated monthly instantaneous fishing mortality rates were highest

during the summer months for all regions (Figure 10), while Region 3 (NC-SC) also had a high fishing mortality rate for the end of the year, or winter months including November and December. These estimates in the latter months of the year match the fishery dynamics during this time period whereby there was a large "fall fishery" operating off the coast of North Carolina, as has been described in past stock assessment documents (SEDAR 2020).
The migration dynamics estimated from the base run model suggest that most of the fish remain in the same area during May through October (Figure 11), which is consistent with migration theories historically proposed. The exception to that is from May to July, there is movement of fish from Region 3 (NC-SC), into Region 2 (MD-VA). Otherwise, most of the movement dynamics occur during October to May, over the winter months (Figure 11). In the northernmost region, approximately one-third of the fish remain resident in the same region, while two-thirds migrate south into Region 2 (MD-VA) and Region 3 (NC-SC) in approximately equal proportions. In Region 2 (MD-VA), the majority, approximately 75% of the fish, remain in the region with smaller amounts moving north and south. In Region 3 (NC-SC), approximately 75% of the fish remain in the region with about 25% of the fish moving into Region 2 (MD-VA). Finally, in Region 4 (GA & FL), the majority of fish move from this southernmost region into Region 3 (NC-SC), the region directly north.

The estimated natural mortality value from the base run of the tagging model was 0.92 (95% CI: 0.88, 0.97) (Table 3, Figure 12). The estimated natural mortality parameters were normally distributed. In addition, the median value of the MCMC uncertainty analysis was the same as the estimate from the maximum likelihood estimators.

Results – Sensitivity runs
The first set of sensitivity runs was used to determine which data set or sets had the largest influence on the estimate of natural mortality (Table 3). Stepping through the data inputs one by one from the Liljestrand et al. (2019a) base run used in the 2020 assessment, where M was 1.16 (95% CI: 1.13, 1.19) using the maximum likelihood estimator, to the data configuration used for this current base run assessment allowed for determining which data set had the largest influence. Specifically, the sensitivity runs that updated the confidential effort data (M = 1.19; 95% CI: 1.17, 1.21) and changed the recoveries by two tags (M = 1.16; 95% CI: 1.13, 1.19) had no or very little influence on the estimation of natural mortality. The sensitivity run that adjusted the magnet efficiencies based on the primary magnets had the largest impact (M = 0.94, 95% CI: 0.90, 0.98).

The second set of runs compared the use of the Coston data in the base run to the use of the NMFS re-digitized data (Table 3). Using the same time series of data available for Coston, 1966-1969, the estimate of M from the NMFS re-digitized data sensitivity analysis was lower at 0.75 (95% CI: 0.67, 0.82). Using the complete release and recovery time series (1966-February 1971, an additional 14 months of data), the estimate of M from the sensitivity analysis was higher at 0.96 (95% CI: 0.91, 1.00), which overlapped with the uncertainty bounds from the base run using the Coston dataset. The reduced data in the northernmost region led to differences in the estimates of the fishing mortality, specifically a reduced fishing mortality rate was estimated as the recoveries in the region were substantially reduced (Full time series; Figure

13). This lack of recoveries suggested that the fishing mortality was low, which is contrary to what the fuller data set provides. In addition, the migration rates estimated were also impacted with the clearest divergence occurring during the winter months in the northernmost region (Full time series; Figure 14). Specifically, higher retention rates were estimated in the northernmost region, with those that did migrate, about 25%, moving further south to Region 3 (NC-SC). Finally, the observed and expected recoveries are well fit with a balance in residuals over space and time (Figure 15). Of note, is the reduced recoveries in the northernmost region, Region 1, for this time series.

The two "Stepwise" approach runs (Table 3) were provided to explore the model robustness to this proposed method (Ault et al. 2023). The sensitivity run using the "Stepwise" method with non-confidential data as provided by Ault et al. estimated a M = 0.51 (95% CI: 0.43, 0.60), while the sensitivity run with the confidential effort included estimated a M = 0.47 (95% CI: 0.37, 0.57). The annual instantaneous fishing mortality rate is higher in the northernmost region using the "Stepwise" approach, nearing levels in Region 3 (NC-SC) (Figure 16). That pattern is less pronounced with the confidential effort included; however, the stepwise approach with the confidential effort included estimates a much higher fishing mortality rate in the southernmost region (Region 4 GA-FL). The monthly instantaneous fishing mortality estimated using the "Stepwise" approach is highest in the northernmost region during the fall fishery (Figure 17), which only operated off of NC and SC (Region 3). This result doesn't match with the knowledge about the fishery during this time period, and it was unlikely that the fishery was exerting the largest fishing mortality in the northernmost region during the fall fishery time period. When including the confidential effort data, the monthly instantaneous fishing mortality estimated using the "Stepwise" approach is highest in Region 4 (GA-FL), which is also unexpected given NC and SC were the center of the fishery during this time period (Figure 17). The migration rate estimates for the "Stepwise" approach are substantially different for the May to June and October to May time periods (Figure 18). Specifically, the residency rates in the northernmost region are approximately zero, which does not match with survey data collected during that time period that captures menhaden. The movement rates with inclusion of the confidential data are more in line with the base run.

Application to the Stock Assessment Model

Based on the size of the tagged fish, the SAS assumed that the point estimate of M from the tagging model represented M on age-1.5 fish. To develop an age-varying vector of M for use in the Beaufort Assessment Model (BAM), the statistical catch-at-age model used in the single-species assessment, the SAS used the same approach as the 2020 benchmark assessment in scaling the Lorenzen curve to the tagging estimate. The Lorenzen (1996) estimates of M-at-age, based on a time-constant mid-year weight-at-age for Atlantic menhaden, were scaled so that the M at age-1.5 was equal to the point estimate from the tagging model. As a result, M on age-0 and age-1 fish was higher than the tagging estimate, while M on ages 2-6+ was lower.

Life History Approaches

As part of the 2020 benchmark assessment, the SAS explored a number of life-history based approaches to calculate M, both age-constant and age-varying. Age-constant methods used information on maximum age (t_{max}), von Bertalanffy growth parameters (K, L_{∞}), and average water temperature (T°C); the updated equations from Then et al. (2014) were used for these methods, and the calculations were done with growth parameters derived from both the full time-series and from the most recent five years of the benchmark data (2011-2015) (Table 4).

The SAS also explored several approaches that have been used to provide age-varying estimates of *M* for menhaden in the past. This included Peterson and Wroblewski (1984), Boudreau and Dickie (1989), Lorenzen (1996), and Charnov et al. (2013). All of the approaches use an inverse relationship between length or weight and *M*. To apply these methods, weight-at-age was calculated for the middle of the model year (September 1).

The method of Lorenzen (1996) has been used extensively in recent years to provide age-varying estimates of *M* (Table 5) and it has become common to scale the Lorenzen (1996) estimates of age-varying *M* such that cumulative survival from age-1 through the maximum age is equal to a specified percent, usually 1.5% based on Hewitt and Hoenig (2005). The updated work of Then et al. (2014) suggested a lower survival at maximum age (0.3% - 0.6%). When similarly scaled, the resulting *M* from Peterson and Wroblewski (1984), Boudreau and Dickie (1989), and Lorenzen (1996) provide very similar results (Table 6).

The age-constant methods provided estimates that were similar to the earlier tag-based estimates of Dryfoos et al. (1973) and Reish et al. (1985) and lower than the estimate of M from the revised tagging model (Table 4), a major point of criticism from Ault et al. (2023). However, these estimators are the product of regressions of observed values of *M* and maximum age or growth parameters and include process error in the form of variability in the relationship between *M* and maximum age or *K* from species to species and observation error in the measurement of *M,* maximum age, and *K* in the underlying data. Hamel and Cope (2022) in their re-estimation of the Then et al. (2014) inverse maximum age approach found that the natural log-space standard deviation around the relationship between *M* and maximum age was 0.44, and, assuming half the observed variance is due to errors in estimates of *M* in the underlying data, recommended a 95% prediction interval for *M* of the point estimate divided by 1.8 to the point estimate multiplied by 1.8. For Atlantic menhaden, the point estimate using the Hamel and Cope (2022) estimator was 0.54, so the prediction interval would be 0.3-0.97. The prediction interval around the estimate of *M* using *K* would be even wider, as there is more variability in that relationship (Hamel and Cope 2022). For the assessment, the SAS is using an age-varying estimate of M scaled to the tagging estimate, which means that M on older ages is lower, reaching an asymptote at M=0.57 for age-6+ fish. The age-constant M that would leave the same proportion of fish alive at the oldest age class is M=0.75, which is well within Hammel and Cope's (2022) prediction interval for the inverse maximum age estimator.

Under the age-varying estimate of M from scaling the Lorenzen (1996) curve to the tagging estimate, approximately 0.06% of a cohort would be left alive at the oldest age. While this is low compared the historical understanding that 1.5% should be alive from Hewitt and Hoenig (2005), Hamel and Cope's (2022) inverse maximum age approach would result in 0.45% left alive, Then et al.'s (2014) revised non-linear fit would leave 0.26% alive, the unscaled Lorenzen

(1996) would leave 0.16% left alive, and the upper bound of the prediction interval for Hamel and Cope's (2022) inverse maximum age estimator would leave 0.01% alive. Thus, the age-varying tag-based estimator is not out of line with other estimators in terms of the proportion of a cohort that survives to the maximum age (Figure 19). Atlantic menhaden are batch spawners and their larvae are found in coastal waters nearly year-round (Simpson et al 2017); it is not unreasonable to suppose that they could produce enough recruits to be encountered in sampling even if such a low proportion survive to the maximum observed age, particularly since those age-10 fish were collected at a time when fishing effort was higher and distributed more widely than it is currently.

Estimates of M for Other Forage Species
Ault et al. (2023) criticized the estimates of M from Liljestrand et al. (2019a) as unrealistically high. Direct estimates of M for other forage species are rare, but the M of 0.92 for age-1.5 fish is comparable to estimates for other species. Wilhelm (2023) used a tagging model to estimate M for age 1-2 Gulf menhaden (*Brevoortia patronus*) from the Gulf equivalent to the NOAA Fisheries Atlantic menhaden tagging study and found an annual M of 1.08 (95% confidence interval 1.04 - 1.12). Uriarte et al. (2016) used a catch curve to estimate M for anchovy (*Engraulis encrasicolus*) from survey data during a fishery closure in the Bay of Biscay and found an M of 0.7 for age-1 fish, with age-2+ fish experiencing an M of twice that, consistent with this species' shorter lifespan and previous hypotheses that anchovies experience senescence at older ages.

Discussion
The M WG carefully vetted the raw data and the methods underlying the Liljestrand et al. (2019a) and Ault et al. (2023) analyses, in consultation with both sets of authors, to understand what was causing the differences in results between the two studies and to determine the best approach to estimating M from the available tagging data. The M WG recommended the use of the summarized Coston (1971) dataset over the NMFS re-digitized dataset due to its more comprehensive spatial coverage, but found that both datasets produced similar estimates of M. The M WG concluded that the major source of the discrepancy between the two estimates was the handling of the magnet efficiency rates (i.e., the reporting rate), with the issue of confidential effort data contributing to a lower degree.

The M WG agreed with Ault et al.'s (2023) conclusion that there was an error in how Liljestrand et al. (2019a) had calculated the magnet efficiencies for use with the Coston (1971) dataset and therefore developed revised estimates of magnet efficiencies using the correct subset of the plant test data.

However, the M WG did not agree with Ault et al.'s (2023) stepwise approach to modify the observed magnet efficiency estimates to reduce the overall log-likelihood of the model. While Ault et al. (2023) is able to improve the fit to the observed recaptures, they do so at the expense of the observed plant efficiency data: the annual, monthly, regional estimates they produce are biased low compared to the distribution of the observed data at the regional level. It is this lower magnet efficiency that results in their lower estimates of M.

The tagging model framework was tested via simulation analyses to estimate migration and M based on the input data, including the magnet efficiencies. Formulating the model to estimate magnet efficiencies along with M and movement rates, either internally or via this external, "stepwise" approach, has not been simulation tested, and the SAS believes this is likely not a robust use of the model.

Ault et al. (2023) argue that the higher estimate of M from the Liljestrand et al. (2019a) method is unrealistic because it does not agree with earlier estimates using the same dataset or with life history estimators. While the revised M from the Liljestrand et al. (2019a) method is higher than Dryfoos et al. (1973) and Reish et al. (1985), the Brownie model is recognized as an improvement over the catch-curve-based (Dryfoos et al. 1973) or VPA-based (Reish et al. 1985) approaches that produced those lower estimates. In addition, the Liljestrand et al. (2019a) version of the Brownie model takes into account the movement rates of Atlantic menhaden, as well as the fact that tag releases and effort were not distributed evenly over space and time.

As part of this argument, Ault et al. (2023) presented a figure showing the estimate of M for Atlantic menhaden as an outlier when compared to other assessed species (Figure 3.2 in SEDAR 102-RW-02). This figure is somewhat misleading, as the majority of species on that figure use age-varying natural mortality, but the average or age-constant equivalent is presented, and in some cases is out of date. When the age-1 estimate of M from those assessments is plotted with the revised age-1 estimate of M for Atlantic menhaden, Atlantic menhaden is no longer an outlier (Figure 20).

More significantly, the Atlantic menhaden M is one of only two values derived from empirical data. The other species use life-history based estimators, most commonly Hoenig (1983), either as an age-constant value or to scale the age-varying values (Table 5). It is not surprising that they all seem to follow a similar curve, as they are all using the same relationship between maximum age and M, which was derived from a meta-analysis of multiple different species across a number of different trophic guilds. They do not represent independent estimates of M. In addition, this figure does not take into account the uncertainty around the estimators themselves. The revised tag M is within the bounds of uncertainty for the most recent estimation of the inverse maximum age and one-*K* approaches (Hamel and Cope 2022).

The SEDAR (2020) benchmark for Atlantic menhaden used the tagging estimate of M to scale the Lorenzen (1996) curve to produce estimates of M-at-age because the tagging estimate represented an empirical estimate derived from a robust, well-designed study: over 1 million tagged fish were released across three and a half years, with complementary studies to estimate tag loss and tag mortality rates, as well as a comprehensive effort to measure reporting rate in the form of 964 magnet efficiency tests conducted at nearly all active plants over the study time period. The final base run of the tagging model using the M WG findings provides biologically plausible results for fishing mortality and migration that comport with other available fishery dependent and fishery independent data sources and past knowledge about Atlantic menhaden. The M WG agreed with the SAS's decision during the 2020 benchmark assessment to use the tagging estimate instead of the life-history based approaches, consistent with recommendations in the literature (e.g., Maunder et al. 2022) that empirical approaches should be preferred over life history estimators wherever possible.

Overall, the tagging model estimate of M as developed by Liljestrand et al. (2019a) and revised by the M WG as described here represents the best available estimate of natural mortality for Atlantic menhaden and is consistent with other empirical estimates of M for forage species as well as life history estimators, particularly when the uncertainty in those estimators is considered.

Literature Cited

Ahrenholz, D. W. 1991. Population biology and life history of the North American menhadens, Brevoortia spp. Marine Fisheries Review 53(4): 3-19.

Ault, J.S., L. Luo, and C.E. Porch. 2023. Understanding Atlantic Menhaden Population Demographics: Re-evaluation of the 1960's NMFS Tagging Data. Miami, FL.

Boudreau, P.R., and L.M. Dickie. 1989. Biological model of fisheries production based on physiological and ecological scalings of body size. Canadian Journal of Fisheries and Aquatic Sciences 46:614-623.

Carlson, F.T., and J.W. Reintjes. 1972. Suitability of internal tags for Atlantic Menhaden. Fishery Bulletin of the National Oceanic and Atmospheric Administration 70:514-517.

Charnov, E., H. Gislason, and J.G. Pope. 2013. Evolutionary assembly rules for fish life histories. Fish and Fisheries. 14, 213–224.

Coston, L.C., 1971. Summary of Tags Released and Recovered for the Atlantic Menhaden, 1966-69. U.S. Dep. Commer. Nat. Mar. Fish. Serv. Data Rep. 66, 117pp.

Dryfoos, R. L., R. P. Cheek, and R. L. Kroger. 1973. Preliminary analyses of Atlantic menhaden, *Brevoortia tyrannus*, migrations, population structure, survival and exploitation rates, and availability as indicated from tag returns. Fishery Bulletin 71(3):719-734.

Hamel, O. S., and J.M. Cope. 2022. Development and considerations for application of a longevity-based prior for the natural mortality rate. Fisheries Research 256: 106477.

Hewitt, D.A., and J.M. Hoenig. 2005. Comparison of two approaches for estimating natural mortality based on longevity. Fishery Bulletin 103: 433-437.

Hoenig, J.M. 1983. Empirical use of longevity data to estimate mortality rates. Fishery Bulletin 82: 898-903.

Kroger, R.L. and Dryfoos, R.L., 1972. Tagging and tag-recovery experiments with Atlantic menhaden, *Brevoortia tyrannus.* NOAA Tech. Rep. U.S. Dept. Commer. NMFS SSRF664, 11pp.

Liljestrand, E.M., M.J. Wilberg, and A.M. Schueller. 2019a. Estimation of movement and mortality of Atlantic menhaden during 1966-1969 using a Bayesian multi-state mark recapture model. Fisheries Research 210: 204-213.

_______, Wilberg, M.J., Schueller, and A.M., 2019b. Multi-state dead recovery mark-recovery model performance for estimating movement and mortality rates. Fisheries Research 210: 214-233.

Lorenzen, K. 1996. The relationship between body weight and natural mortality in juvenile and adult fish: A comparison of natural ecosystems and aquaculture. Journal of Fish Biology 49: 627-647.

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

Peterson, I., and J.S. Wroblewski. 1984. Mortality rate of fishes in the pelagic ecosystem. Canadian Journal of Fisheries and Aquatic Sciences 41: 1117-1120.

Reish, R.L., R.B. Deriso, D. Ruppert, and R.J. Carroll. 1985. An investigation of the population dynamics of Atlantic menhaden (*Brevoortia tyrannus*). Canadian Journal of Fisheries and Aquatic Sciences 42 (Suppl 1): 147-157.

Southeast Data, Assessment, and Review (SEDAR). 2020. SEDAR 69 - Atlantic Menhaden Single-Species Benchmark Stock Assessment and Peer Review Report. SEDAR, North Charleston, SC. 691 pp.

Simpson, C.A., H. Bi, D. Liang, M.J. Wilberg, A.M. Schueller, G.M. Nesslage, and H.J. Walsh. 2017. Spawning locations and larval dispersal of Atlantic menhaden during 1977-2013. ICES Journal of Marine Science 74: 1574-1586.

Smith, W.E., Schueller, A.M., and J.W. Smith. 2014. Application of a spatial tag-return model to historic Atlantic menhaden data. Final report to Atlantic States Marine Fisheries Commission. 22p. Available at: https://sedarweb.org/documents/sedar-40-stock-assessment-report-atlantic-menhaden/

Then, A. Y., J. M. Hoenig, N. G. Hall, and D. A. Hewitt. 2014. Evaluating the predictive performance of empirical estimators of natural mortality rate using information on over 200 fish species. ICES Journal of Marine Science 72(1): 82-92.

Uriarte, A., Ibaibarriaga, L., Pawlowski, L., Masse, J., Petitgas, P., Santos, M. and Skagen, D., 2016. Assessing natural mortality of Bay of Biscay anchovy from survey population and biomass estimates. Canadian Journal of Fisheries and Aquatic Sciences 73(2): 216-234.

Wilhelm, C. 2023. Estimation of mortality rates for the gulf menhaden stock and reference point implementation for the fishery. Graduate Thesis, The University of Southern Mississippi.

Tables

Table 1. Spatial breakdown of tag release data from Coston and the re-digitized NMFS records.

	Area 1 MA-NY	Area 2 NJ-DE	Area 3 MD-VA	Area 4 NC-SC	Area 5 GA-FL
Coston: number of months with Release Data 1966-69	8	12	23	36	23
Re-digitized NMFS: number of months with Release Data 1966-69	4	1	22	20	19
Coston: Releases 1966-69	12,931	36,149	308,305	386,171	322,801
Re-digitized NMFS: Releases 1966-69	8,268	700	283,155	199,631	274,417
Re-digitized NMFS: Percent of Coston Releases 1966-69 (by area)	64%	2%	92%	52%	85%
Re-digitized NMFS: Percent of Coston Releases 1966-69	72%				
Coston: number of months with Release Data 1970	0	0	0	0	0
Re-digitized NMFS: number of months with Release Data 1970	0	0	2	2	0
Coston: Releases 1970					
Re-digitized NMFS: Releases 1970			35,789	585	

Table 2. Batches of tags released in plant-test study by model region, plant, and year. Batches were generally comprised of 100 fish.

Model Region	Plant	1966	1967	1968	1969	1970	1971	TOTAL
1	01			15	14			29
	02		18	24	18	2	7	69
	04		7	5	13			25
	23*						2	2
	25*						4	4
2	07		33	34	23	6	24	120
	09		21					21
	10		35	37	25	25	29	151
	11		30	3		19		52
	29		5	12	1			18
3	12	8	7	1	5	7	3	31
	13	11	17	9	12	10	16	75
	14	5	6	1	7	8	4	31
	15	5	5	2	4			16
	16	5	5	1	7	3	1	22
	17	5	3	7	18	13	18	64
	28	8		11	13	2	15	49
4	19	3	21	28				52
	20	10	15	23	49	18	18	133

*: Plants 23 and 25 were not active from 1966-1969.

Table 3. M estimates and 95% credible intervals for the base run of the tagging model and sensitivity runs.

Run	M (95% Credible Interval)
Base run (Coston dataset, updated confidential effort, recovery data, & magnet efficiencies)	0.92 (0.88, 0.97)
Liljestrand et al. (2019a) (SEDAR 69)	1.16 (1.13, 1.19)
Liljestrand et al. (2019a) with updated effort data	1.19 (1.17, 1.21)
Liljestrand et al. (2019a) with updated recovery data	1.16 (1.13, 1.19)
Liljestrand et al. (2019a) with revised magnet efficiency estimates	0.94 (0.90, 0.98)
NMFS re-digitized data (1966-1969)	0.75 (0.67, 0.82)
NMFS re-digitized data (1966-1971)	0.96 (0.91, 1.00)
Stepwise magnet efficiencies + non-confidential effort data	0.51 (0.43, 0.60)
Stepwise magnet efficiencies + confidential effort data	0.47 (0.37, 0.57)

Table 4. Constant M from life history approaches from SEDAR (2020) using K & L_∞ bias-corrected averaged across annual values, either full period of 1955-2015 (full) or recent period of 2011-2015 (recent). Equations were updated using Then et al. (2015) (Hoenig and Pauly methods) and Hamel and Cope (2022) (inverse maximum age and one-parameter *K*).

Parameters
t_{max} = 10 years
K = 0.5 (recent); K = 0.301 (full)
L_∞ = 29.8 cm (recent); L_∞ = 42.2 cm (full)

Method	Equations	*M* Estimate
$Hoenig_{nls}$	$4.899 t_{max}^{-0.916}$	0.59
$Pauly_{nls\text{-}T}$	$4.118 K^{0.73} L_\infty^{-0.33}$	0.80 (recent); 0.49 (full)
Inverse maximum age	$5.4/t_{max}$	0.54
One parameter *K*	$1.55K$	0.78 (recent); 0.46 (full)

Table 5. Summary of M estimates used to develop species-comparison

Species	M Average (range for all ages, youngest to plus group)	Type	Reported Max Age	Method Type	Specific Method(s)	Reference
Atlantic Menhaden	**- (1.39-0.57)**	**Age-varying**	**10**	**Empirical estimate**	**Tagging model**	**This assessment**
Atlantic Croaker	0.35 (0.461-0.214)	Age-varying	17	M estimator	Lorenzen (2005) scaled to target M (Hoenig (1983))	SEDAR 20
Atlantic Herring	0.35 (-)	Age-constant	12	M Estimator	Hoenig (1983)	NEFSC 2022 update
Atlantic Sardine (Bay of Biscay)	0.44 (1.07-0.4)	Age-varying	15*	M estimator	Gislason (2010), scaled to minimize AIC of assessment model	ICES. 2024. Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA).
Atlantic Sardine (Atlantic Iberian waters)	- (0.98-0.32)	Age-varying	15*	M estimator	Gislason (2010), scaled to minimize AIC of assessment model	ICES. 2024. Working Group on Southern Horse Mackerel, Anchovy and Sardine (WGHANSA).
Black Grouper	0.2 (0.579-0.168)	Age-varying	33	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 48
Black Sea Bass (South Atlantic)	0.375 (1.20-0.311)	Age-varying	11	M estimator	Lorenzen (2022) scaled to target M (Averaged estimates between Hammel and Cope (2022) and assessment model estimates)	SEDAR 76
Bluefish (Atlantic)	- (0.85-0.27)	Age-varying	14	M estimator	Lorenzen (1996)	2022 NEFSC Research Track Assessment
Goliath Grouper	0.18 (0.54-0.13)	Age-varying	37	M estimator	Lorenzen (1996) scaled to target M (Then et al. 2014 (Hoenig-nls))	SEDAR 47
Gray Snapper (Gulf)	0.15 (0.50-0.13)	Age-varying	28	M estimator	Lorenzen (2005) scaled to target M (Hoenig (1983))	SEDAR 51, SEDAR 75
Gray Triggerfish (South Atlantic)	0.38 (0.61-0.39)	Age-varying	16	M estimator	Lorenzen (1996) scaled to target M (Then et al. 2014 (Hoenig-nls))	SEDAR 82

Greater Amberjack (South Atlantic)	- (0.82-0.29)	Age-varying	17	M estimator	Charnov et al. (2013)	SEDAR 59
Greater Amberjack (Gulf)	0.28 (0.74-0.26)	Age-varying	15	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 70
Hogfish	0.179 (0.597-0.109)	Age-varying	25	M estimator	Lorenzen (2005) scaled to target M (Hoenig (1983))	SEDAR 37 (Update 2017)
King Mackerel (Atlantic)	0.16 (0.657-0.157)	Age-varying	26	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 38 (Update 2019)
Mutton Snapper	0.129 (0.3-0.118)	Age-varying	42	M estimator	Lorenzen (2005) scaled to target M (Hammel and Cope (2022))	SEDAR 79
Pacific Hake	0.235	Age-constant	25	Estimated by assessment model	Estimated in assessment with a prior with mean: Hoenig (1983)	2024 Hake Assessment (NWFSC, Fisheries and Oceans Canada)
Pacific Mackerel	0.851 (0.7-1.2)	Age-varying	8	M estimator	Lorenzen (2005) scaled to target M (Hammel and Cope (2022))	NOAA-TM-NMFS-SWFSC-688
Pacific Sardine	0.55 (-)	Age-varying	8	Estimated by assessment model	Prior: Hamel and Cope (2022); Age-varying: Lorenzen (1996)	NOAA-TM-NMFS-SWFSC-698
Red Grouper (South Atlantic)	0.14 (0.3-0.12)	Age-varying	26	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 53
Red Grouper (Gulf)	0.14 (0.5579-0.1209)	Age-varying	29	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 61
Red Porgy	0.22 (0.455-0.209)	Age-varying	25	M estimator	Charnov et al. (2013) scaled to target M (Average of Hoenig (1983) and Then et al. (2014))	SEDAR 60
Snowy Grouper (South Atlantic)	0.08 (0.328-0.071)	Age-varying	56	M estimator	Charnov et al. (2013) scaled to target M (Hoenig (1983))	SEDAR 36, 2020 Update
Spanish Mackerel (South Atlantic)	0.35 (0.68-0.34)	Age-varying	12	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 78
Spanish Mackerel (Gulf)	0.38 (1.26-0.35)	Age-varying	11	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 81
Striped Bass	- (1.13-0.15)	Age-varying	30	Empirical estimate	Estimated M for ages: 1-3 with tagging study (Jiang et al. 2007), age 7+ Hoenig (1983) estimator, ages4-6 scaled between using Lorenzen (1996)	NEFSC 66th SAW

Tilefish (South Atlantic Tilefish)	0.135 (0.39-0.13)	Age-varying	40	M estimator	Lorenzen (2022) scaled to target M (Hammel and Cope (2022))	SEDAR 89
Vermillion Snapper	0.22 (0.341-0.194)	Age-varying	19	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 55
Yellowtail Snapper	0.223 (0.558-0.198)	Age-varying	20	M estimator	Lorenzen (1996) scaled to target M (Hoenig (1983))	SEDAR 96

- : No age-constant estimate of M was provided

(-): No age-varying estimates of M were provided

*: Max age from FishBase, not assessment report

Table 6. Summaries of various age-specific estimates of M including those as inverse function of size-at-age. Scaled Lorenzen estimates have been scaled so that M at age-1.5 is equal to the M from the tagging studies (Smith et al. 2014 for SEDAR 40, Liljestrand et al. 2019a for SEDAR 69, and the M WG's revised tagging M).

Age	Petersen & Wroblewski 1984	Boudreau & Dickie 1989	Lorenzen 1996	Charnov et al. 2013	Lorenzen 1996 Scaled (SEDAR 40)	Lorenzen 1996 Scaled (SEDAR 69)	Lorenzen 1996 Scaled (M WG)
0	1.13	1.1	1.18	1.41	1.12	1.76	1.39
1	0.89	0.79	0.88	0.84	0.82	1.31	1.03
2	0.73	0.62	0.7	0.57	0.65	1.03	0.82
3	0.65	0.53	0.6	0.45	0.57	0.90	0.71
4	0.6	0.48	0.55	0.38	0.52	0.81	0.64
5	0.57	0.44	0.51	0.34	0.5	0.76	0.60
6+	0.54	0.42	0.49	0.31	0.48	0.72	0.57

Figures

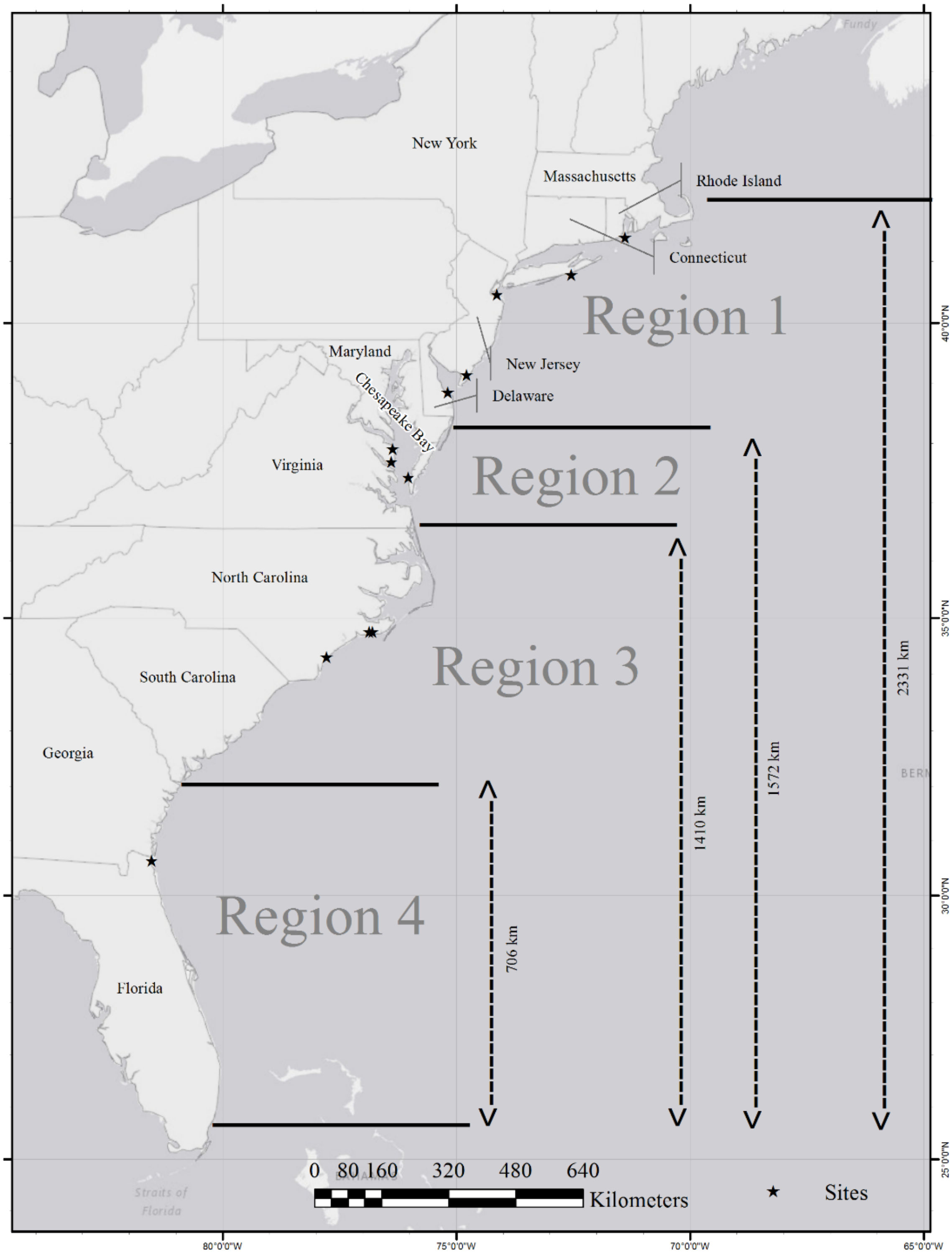


Figure 1. Map of regions used in the tagging model. Stars indicate the locations of reduction plants operating at the time of the tagging study, with some stars representing multiple plants. From Liljestrand et al. (2019a).

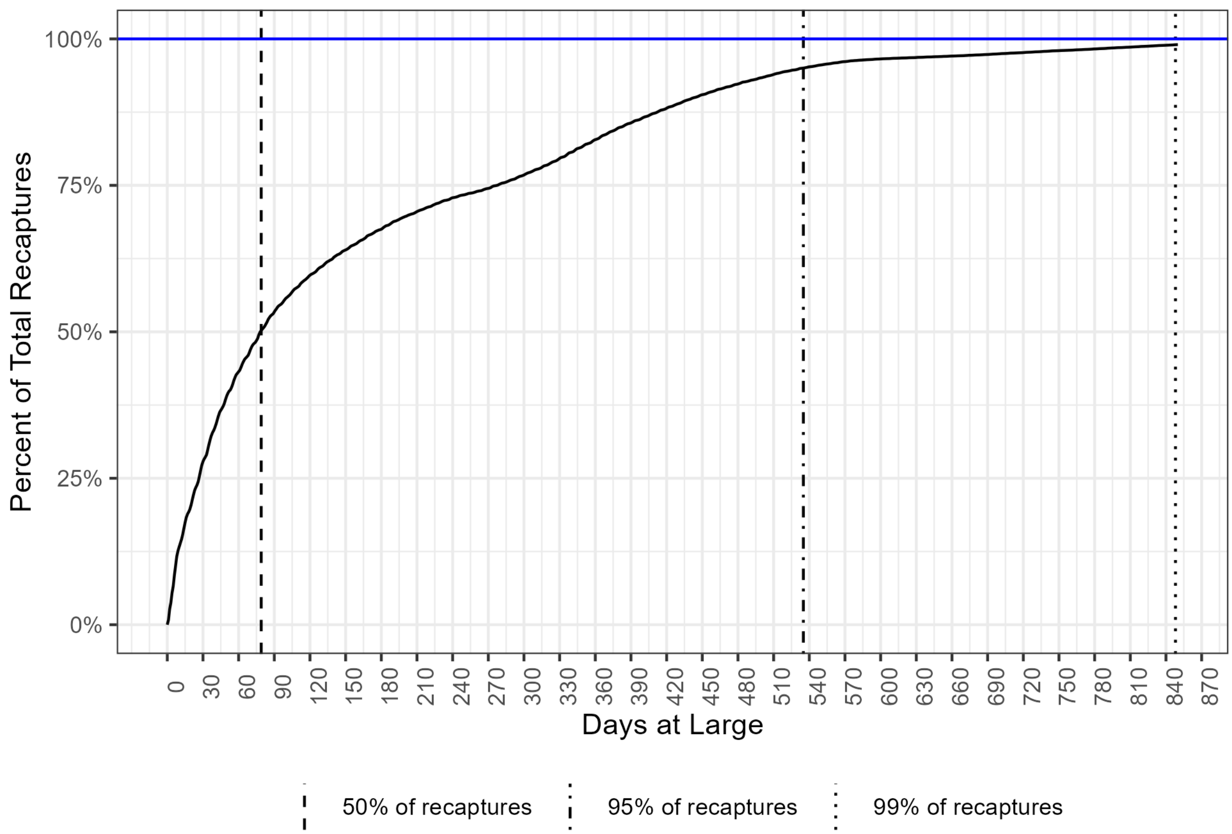


Figure 2. Percent of recaptures on all magnets as a function of time-at-large. Recaptures slowed considerably after the first 80-90 days, and less than 1% of recaptures occurred after 2.5 years.

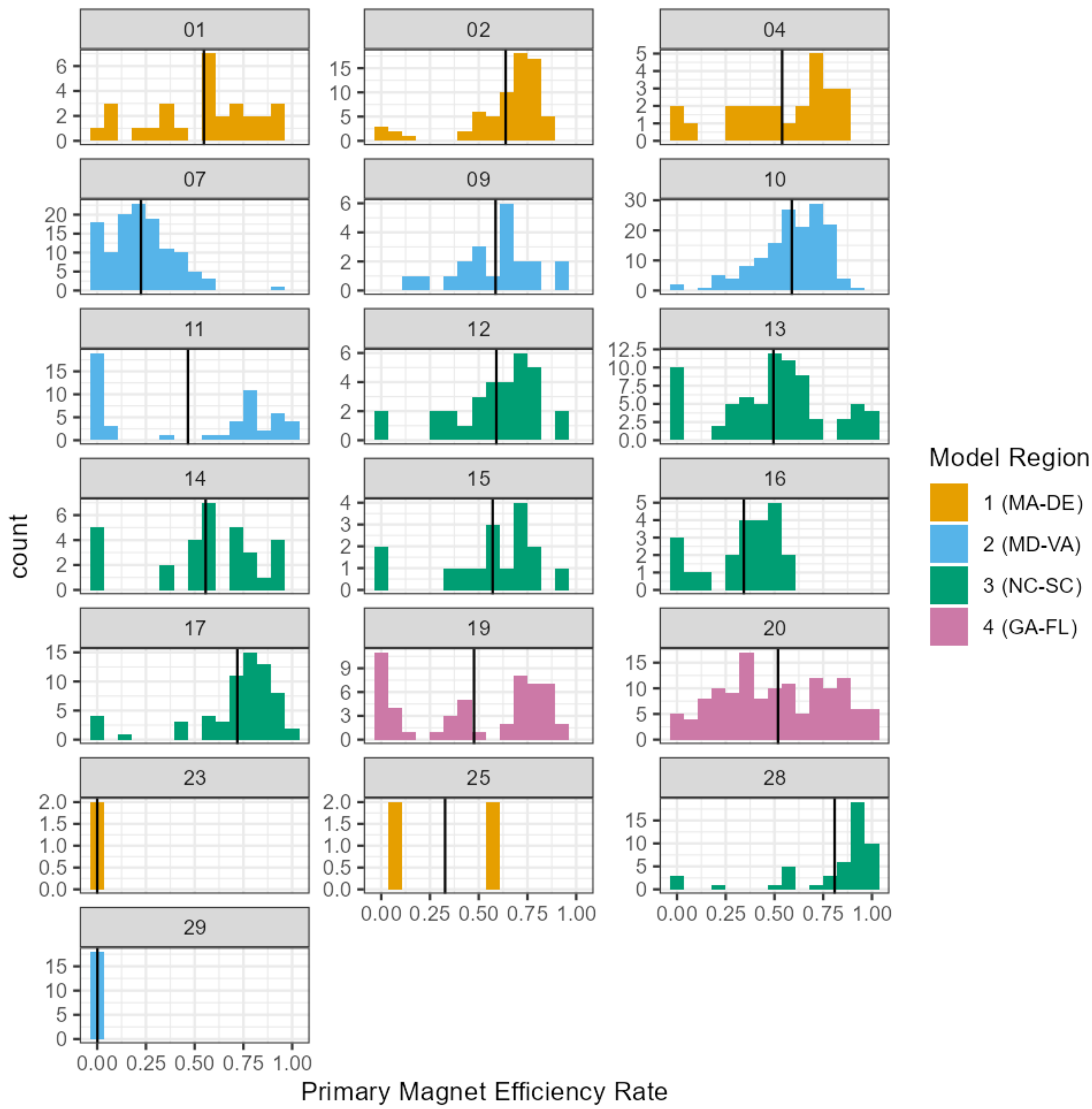


Figure 3. Distribution of primary magnet efficiency rates by plant. Vertical lines indicate the overall magnet efficiency rate for the plant, calculated by minimizing a negative binomial log likelihood. Note that plants 23 and 25 were not active from 1966-1969, and so were excluded from the calculations of magnet efficiency for Region 1 (MA-DE) for the Coston (1971) dataset.

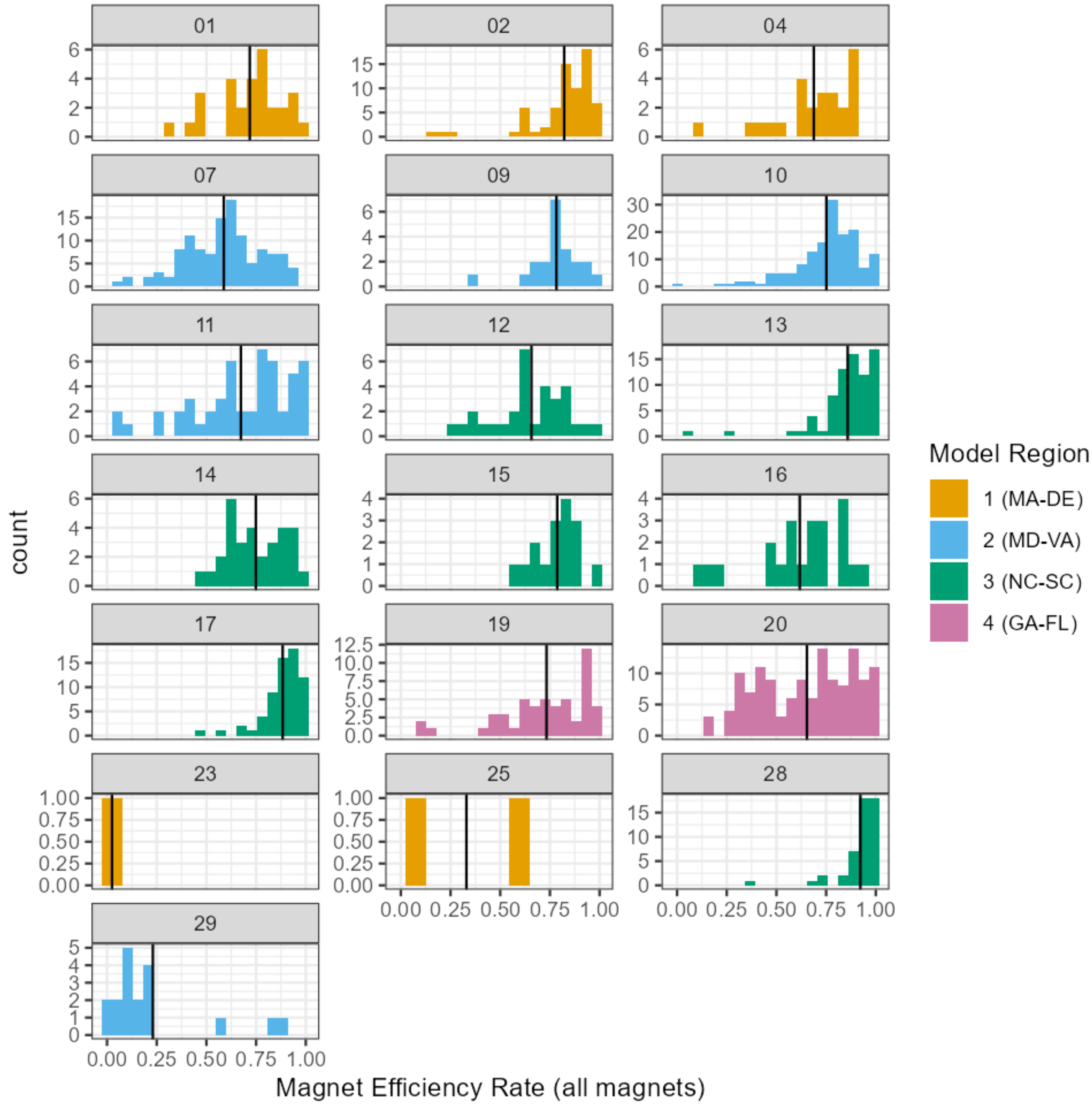


Figure 4. Distribution of magnet efficiency rates by plant using recoveries from all magnets. Vertical lines indicate the overall magnet efficiency rate for the plant, calculated by minimizing a negative binomial log likelihood for each plant.

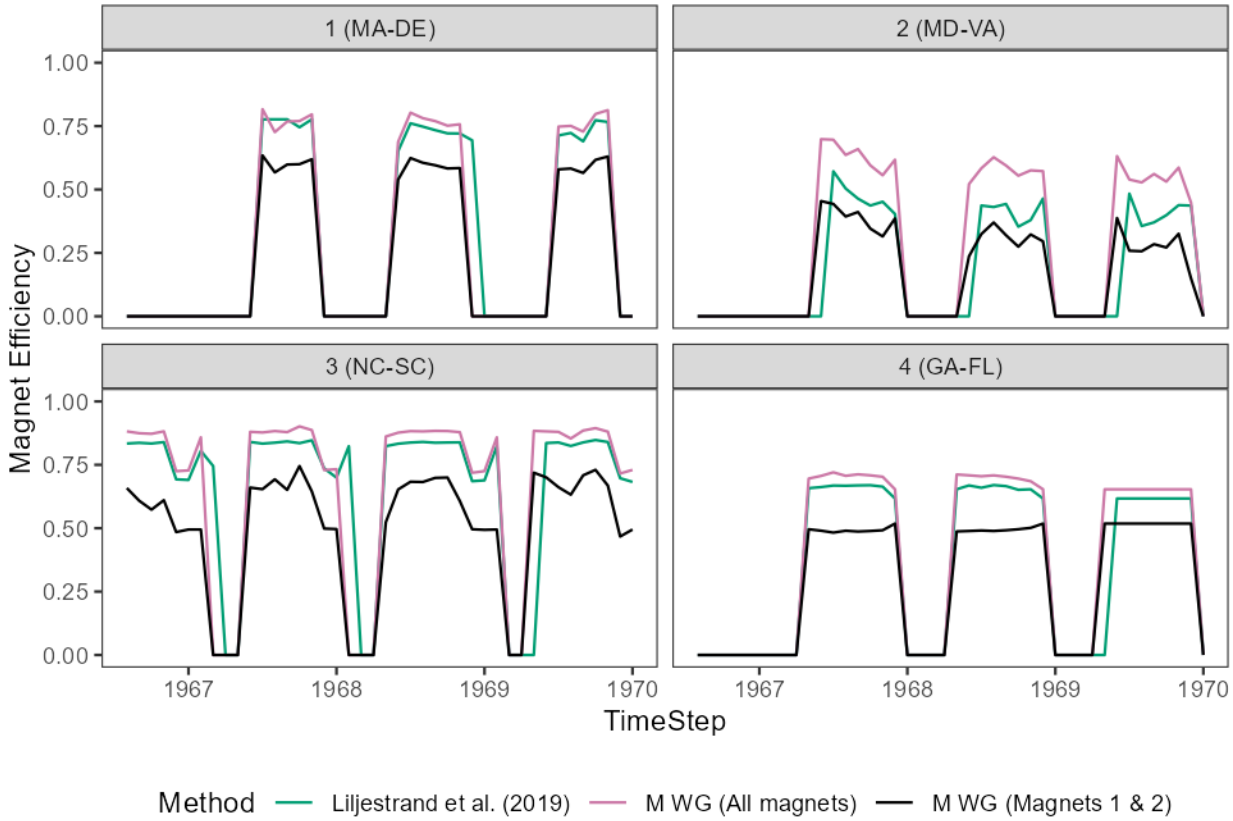


Figure 5. Revised regional magnet efficiencies by year and month compared to Liljestrand et al. (2019a) values.

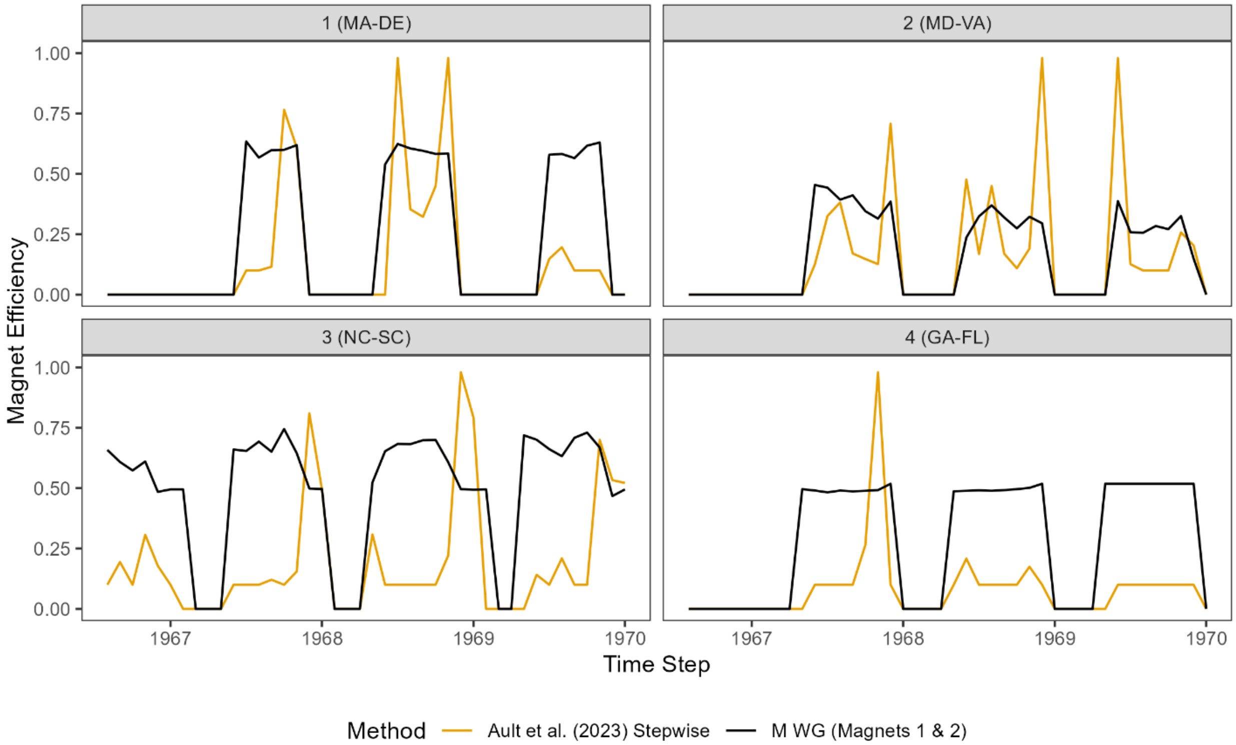


Figure 6. Revised regional magnet efficiencies by month and year as calculated by the M WG compared to those calculated from the Ault et al. (2023) stepwise method with the confidential effort data.

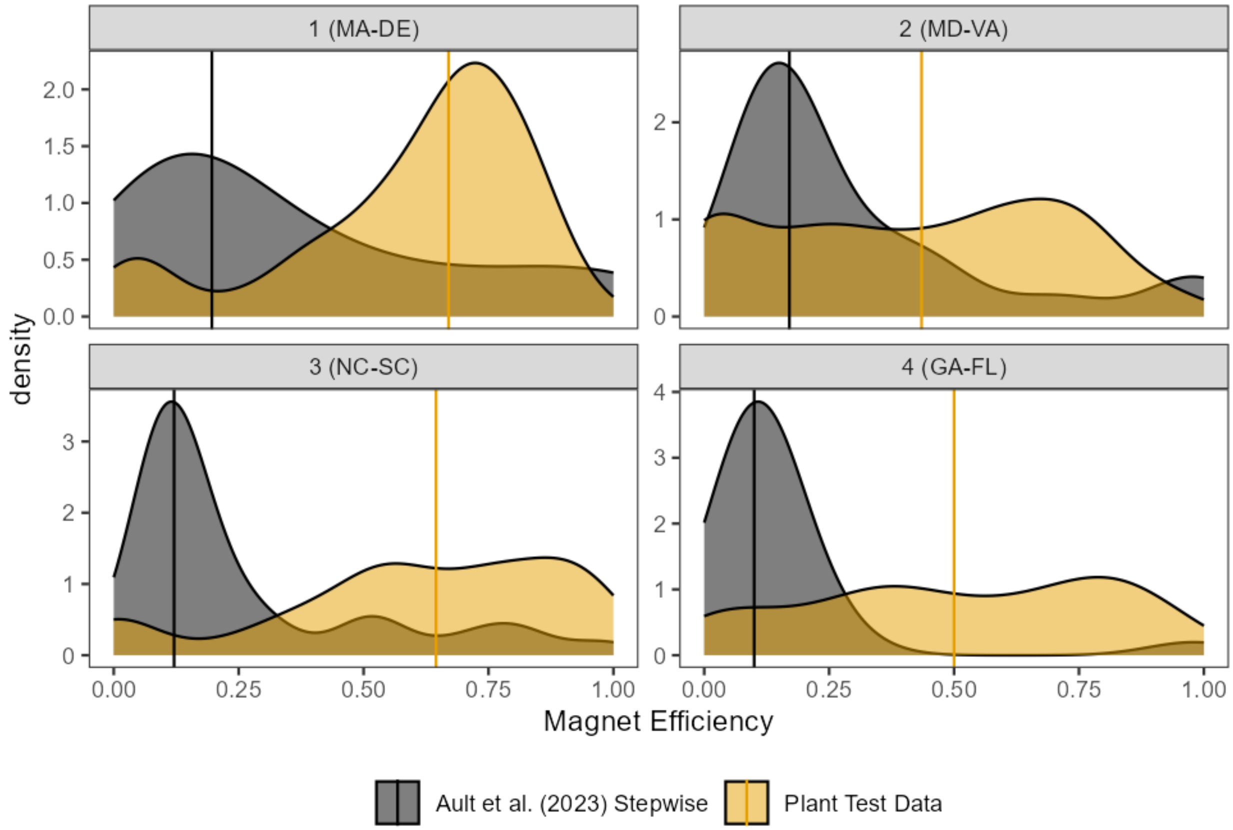


Figure 7. Distribution of adjusted magnet efficiency estimates by region calculated from the Ault et al. (2023) stepwise method using the confidential effort data compared to distributions of observed magnet efficiencies from the plant test data. Vertical lines indicate the median of each distribution.

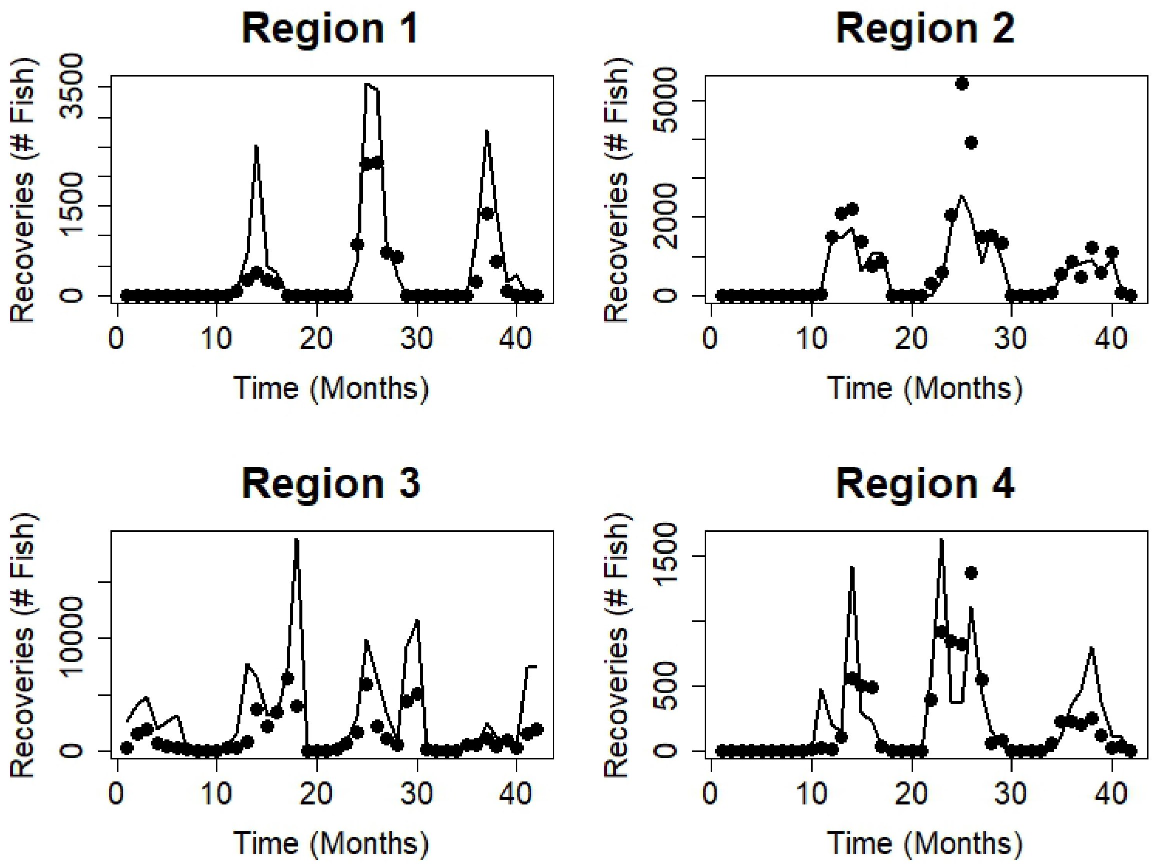


Figure 8. Observed recoveries (black circles) and predicted recoveries (black line) from the base run of the tagging model in numbers of fish for each region from July 1966 until December of 1969.

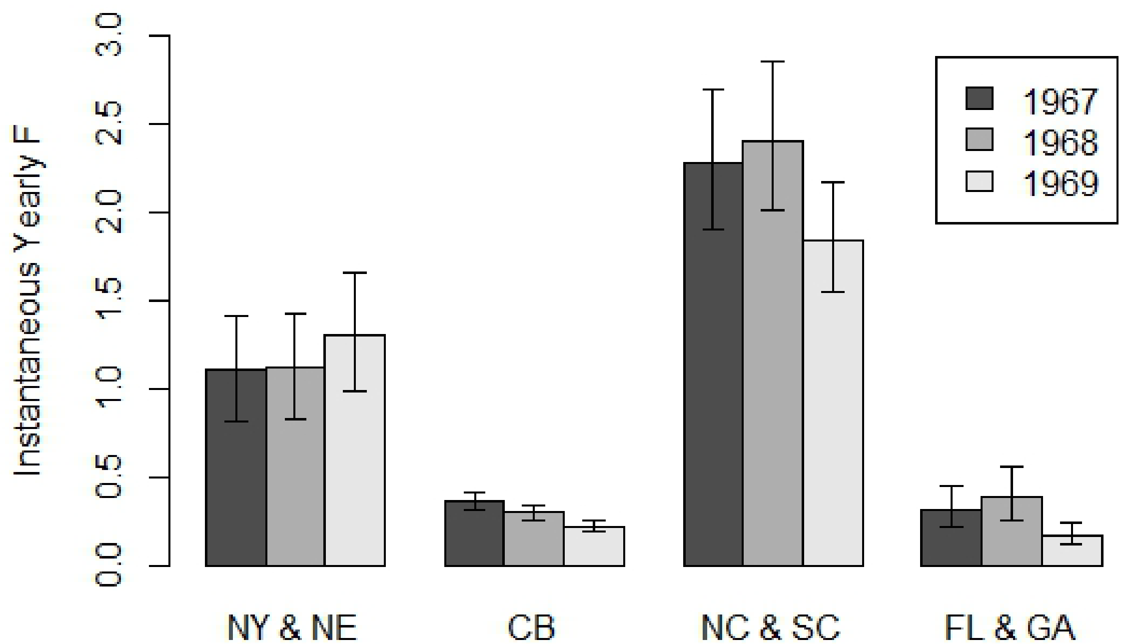


Figure 9. Yearly instantaneous fishing mortality rates for each region from the base run of the tagging model.

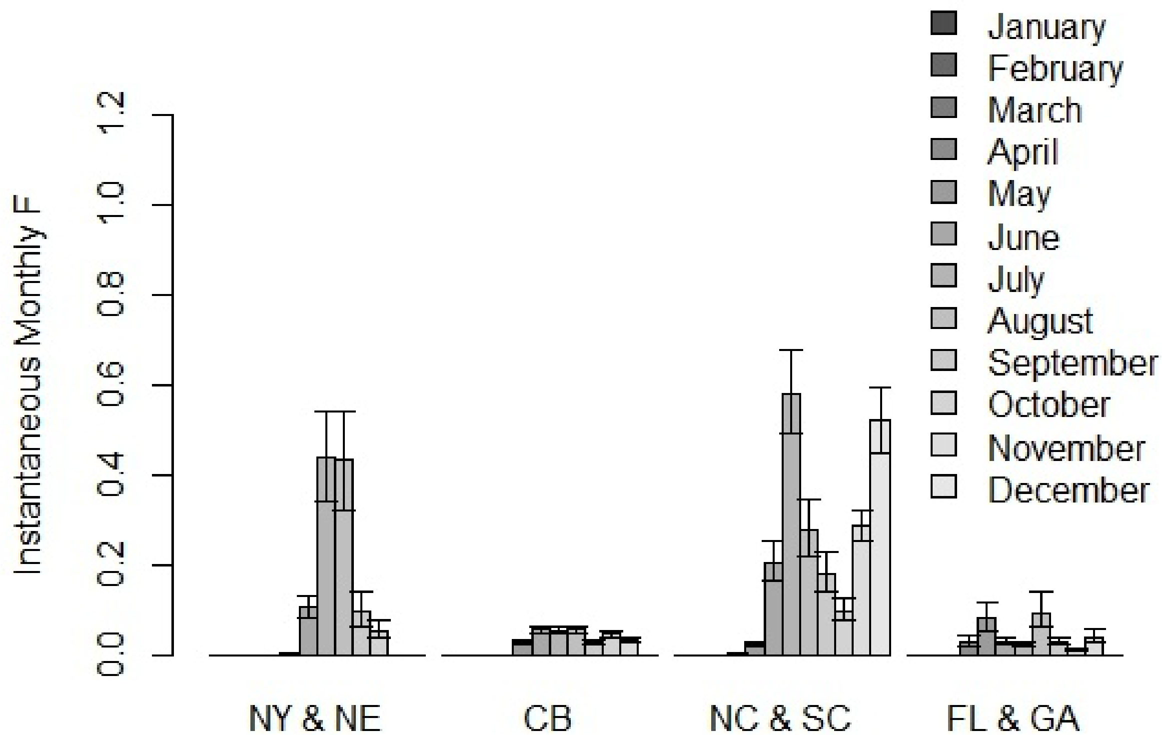


Figure 10. Monthly instantaneous fishing mortality rates for each region from the base run of the tagging model.

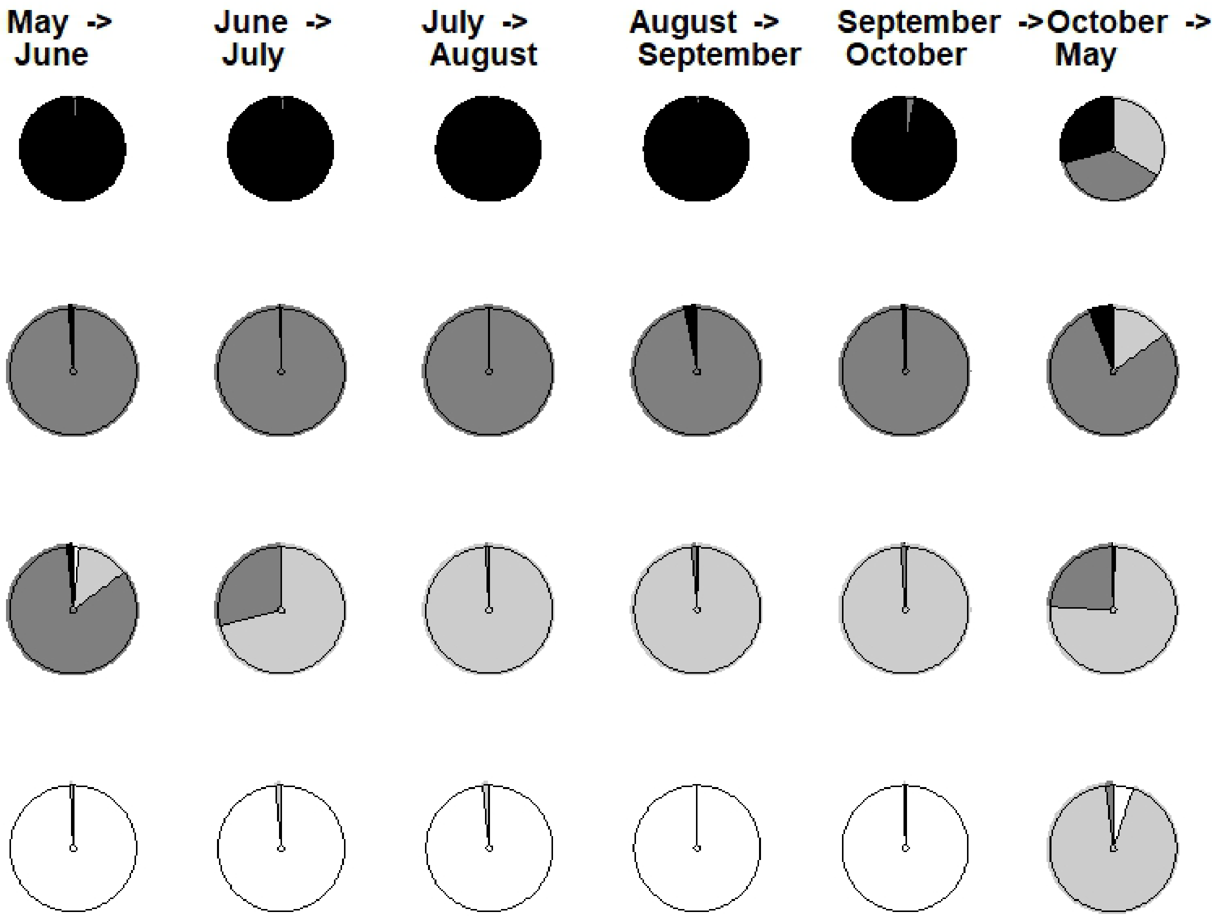


Figure 11. Migration proportions from each region to each region by time step from the base run of the tagging model, as indicated by the monthly time step from May to September and a six-month time step from October to May at the top of the figure. The rows indicate the region from north to south with region 1 (NE & NY), region 2 (NJ & DE), region 3 (NC & SC), and region 4 (GA & FL). The colors also indicate the region with region 1 (NE & NY) being black, region 2 (NJ & DE) being charcoal gray, region 3 (NC & SC) being light gray, and region 4 (GA & FL) being white.

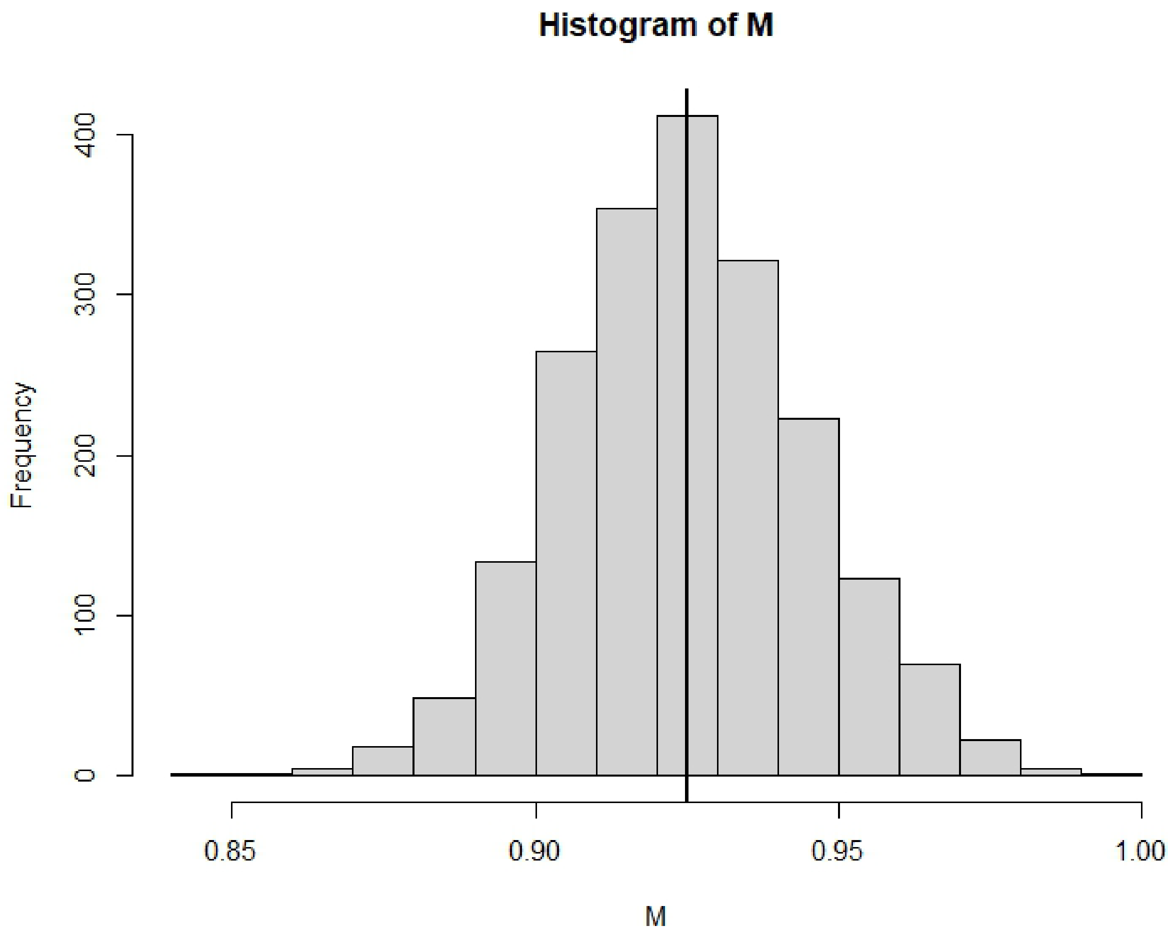


Figure 12. Distribution of natural mortality (M) from the MCMC analysis for the base run of the tagging model. The central tendency or median was 0.92 while the 95th percentiles are [0.88, 0.97].

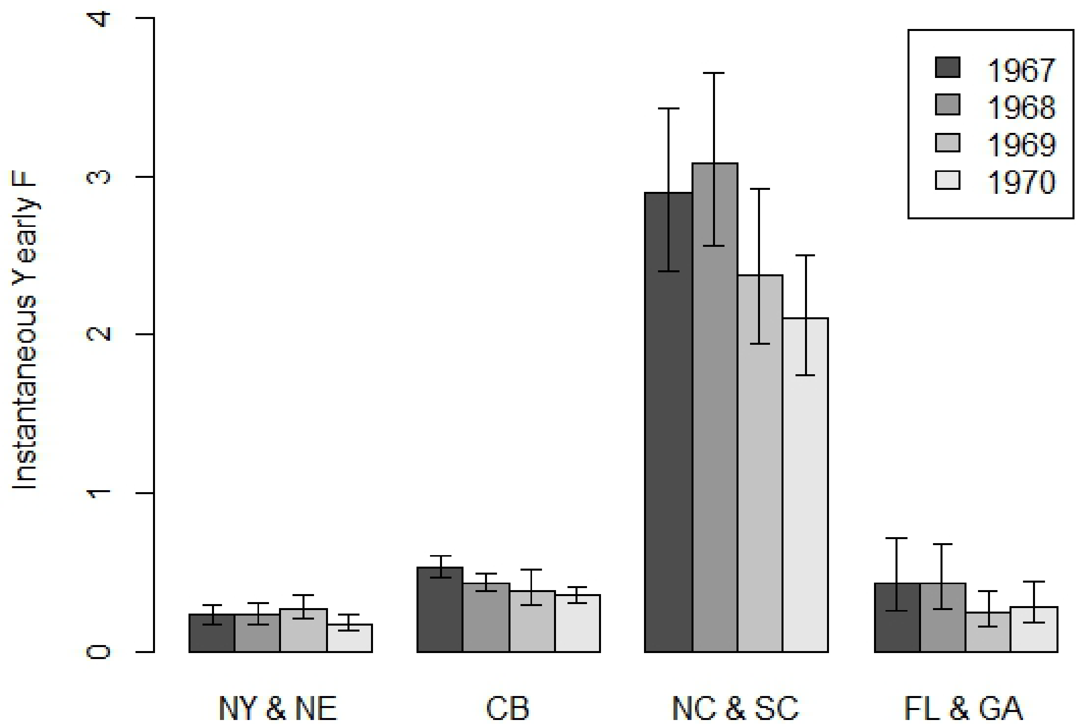


Figure 13. Yearly instantaneous fishing mortality rates for each region from the sensitivity run using the NMFS re-digitized time series from 1966-1971.

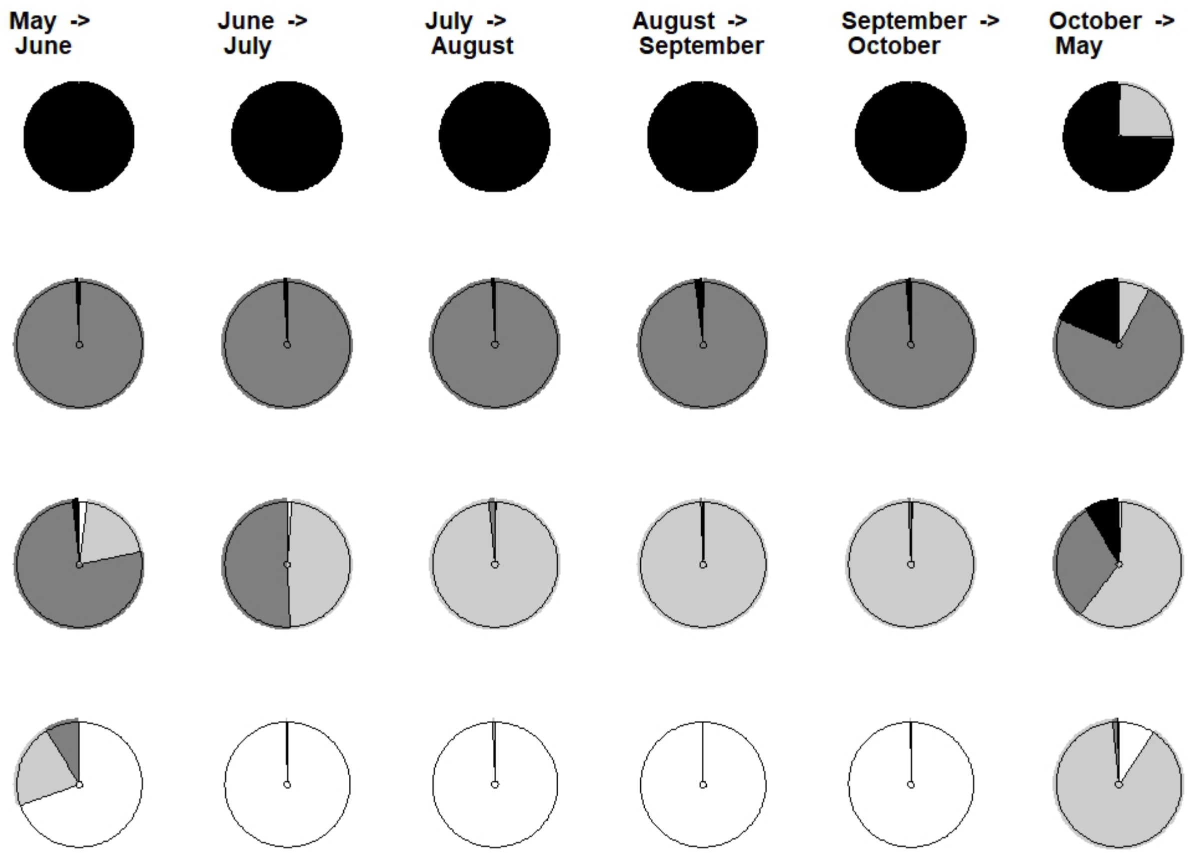


Figure 14. Migration proportions for the sensitivity run based on the NMFS re-digitized time series from 1966-1971. The monthly time step from May to September and a six-month time step from October to May are indicated at the top of the figure. The rows indicate the region from north to south with region 1 (NE & NY), region 2 (NJ & DE), region 3 (NC & SC), and region 4 (GA & FL). The colors also indicate the region with region 1 (NE & NY) being black, region 2 (NJ & DE) being charcoal gray, region 3 (NC & SC) being light gray, and region 4 (GA & FL) being white.

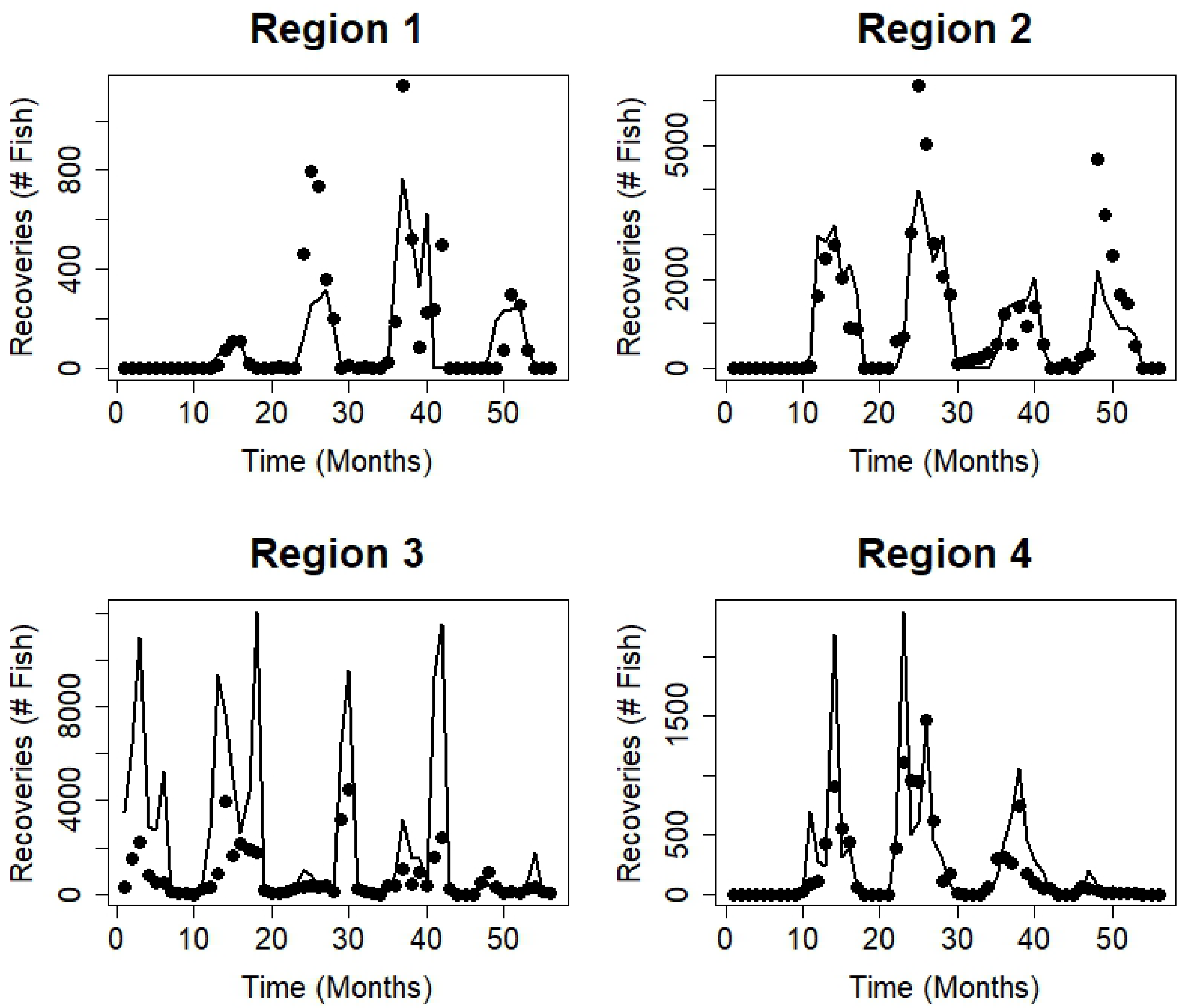
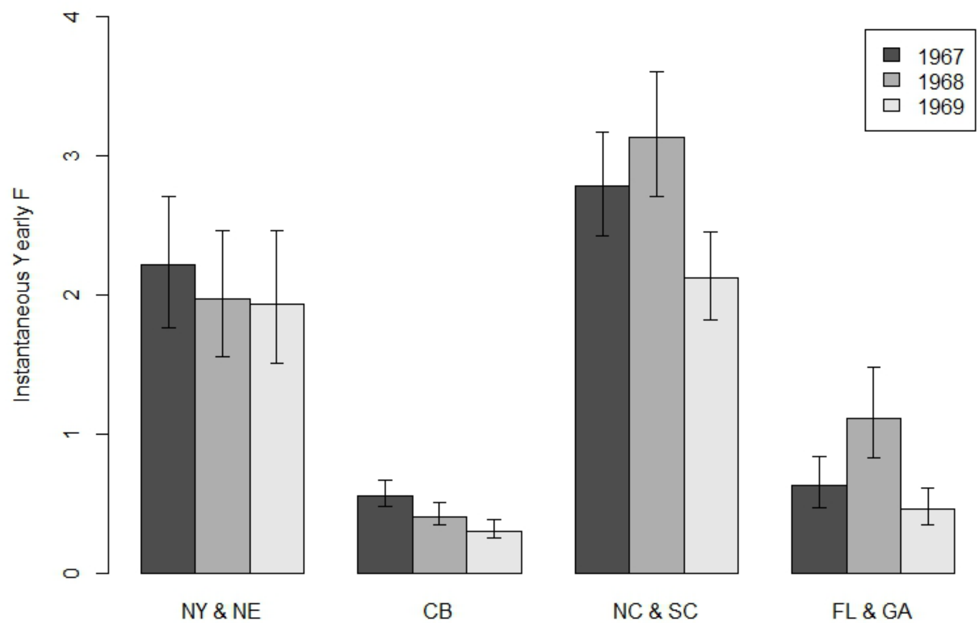


Figure 15. Observed recoveries (black dots) and predicted recoveries (black line) as estimated from the sensitivity run of the tagging model using the NMFS re-digitized data set for the full time series (1966 – February 1971).



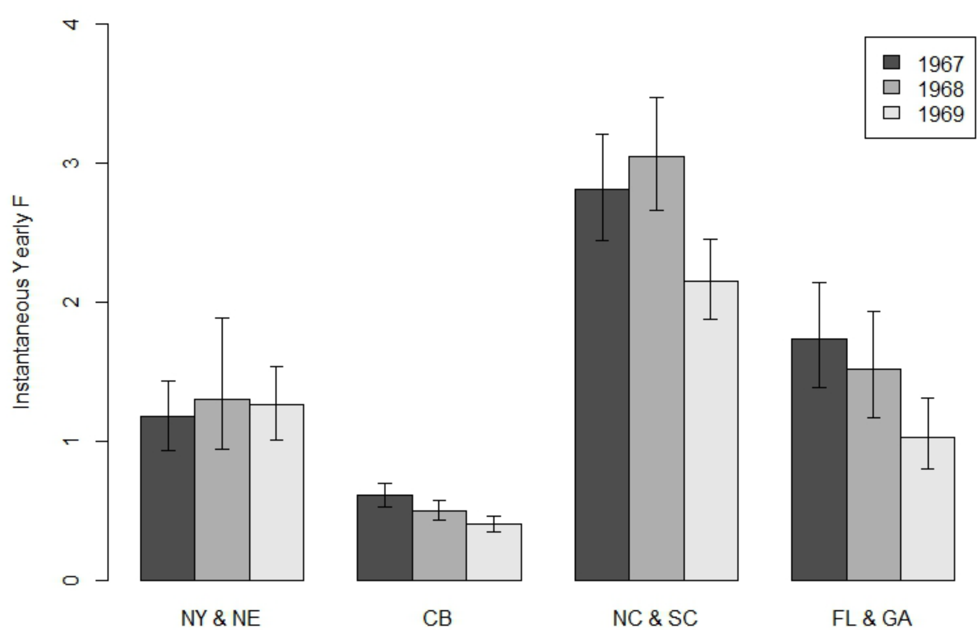


Figure 16. Yearly instantaneous fishing mortality rates for each region from the sensitivity runs using the provided "Stepwise" approach (top) and with the confidential effort data, step 6 (bottom).

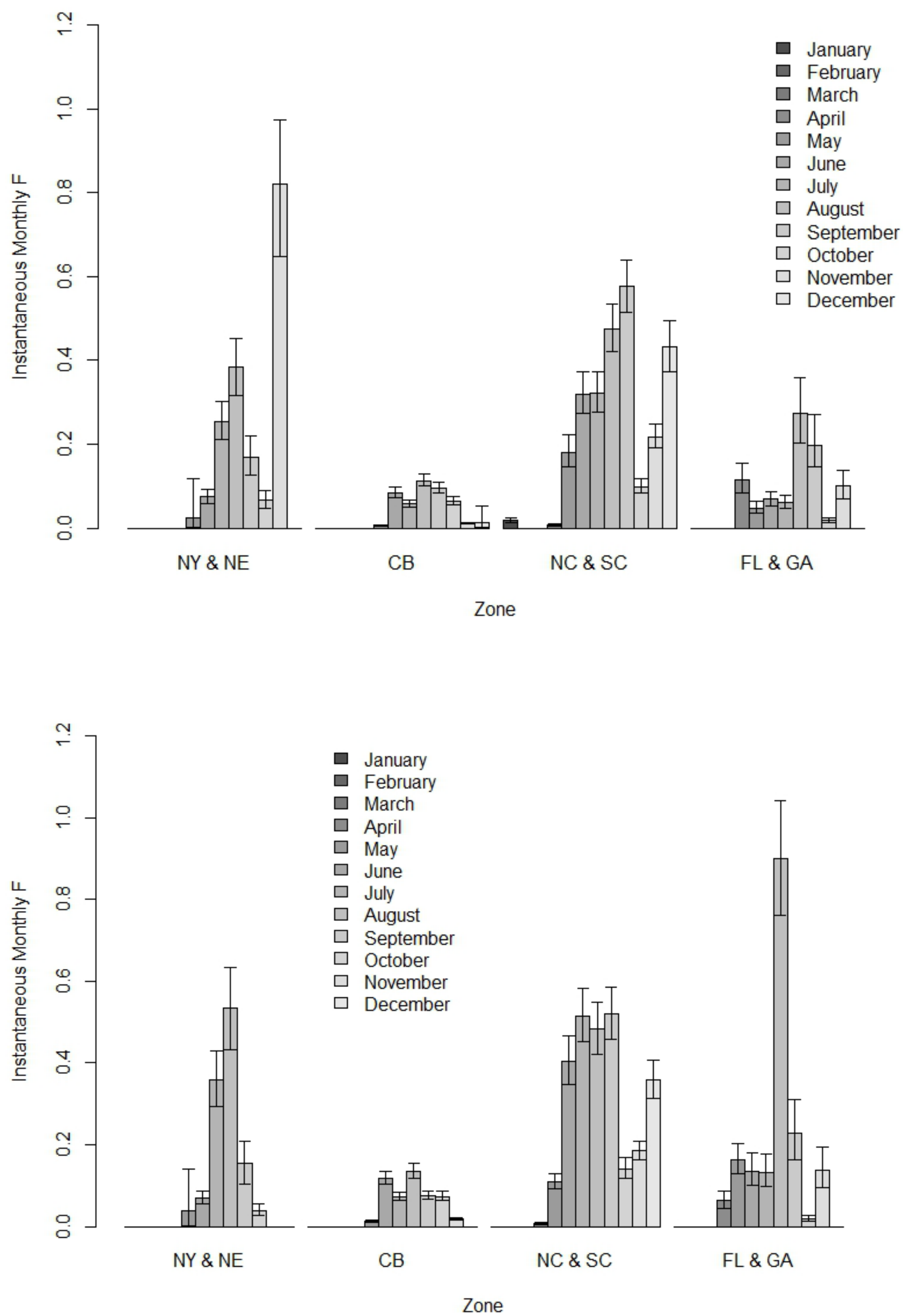


Figure 17. Monthly instantaneous fishing mortality rates for each region from the sensitivity runs using the "Stepwise" approach with non-confidential effort data (top) and with the confidential effort data (bottom).

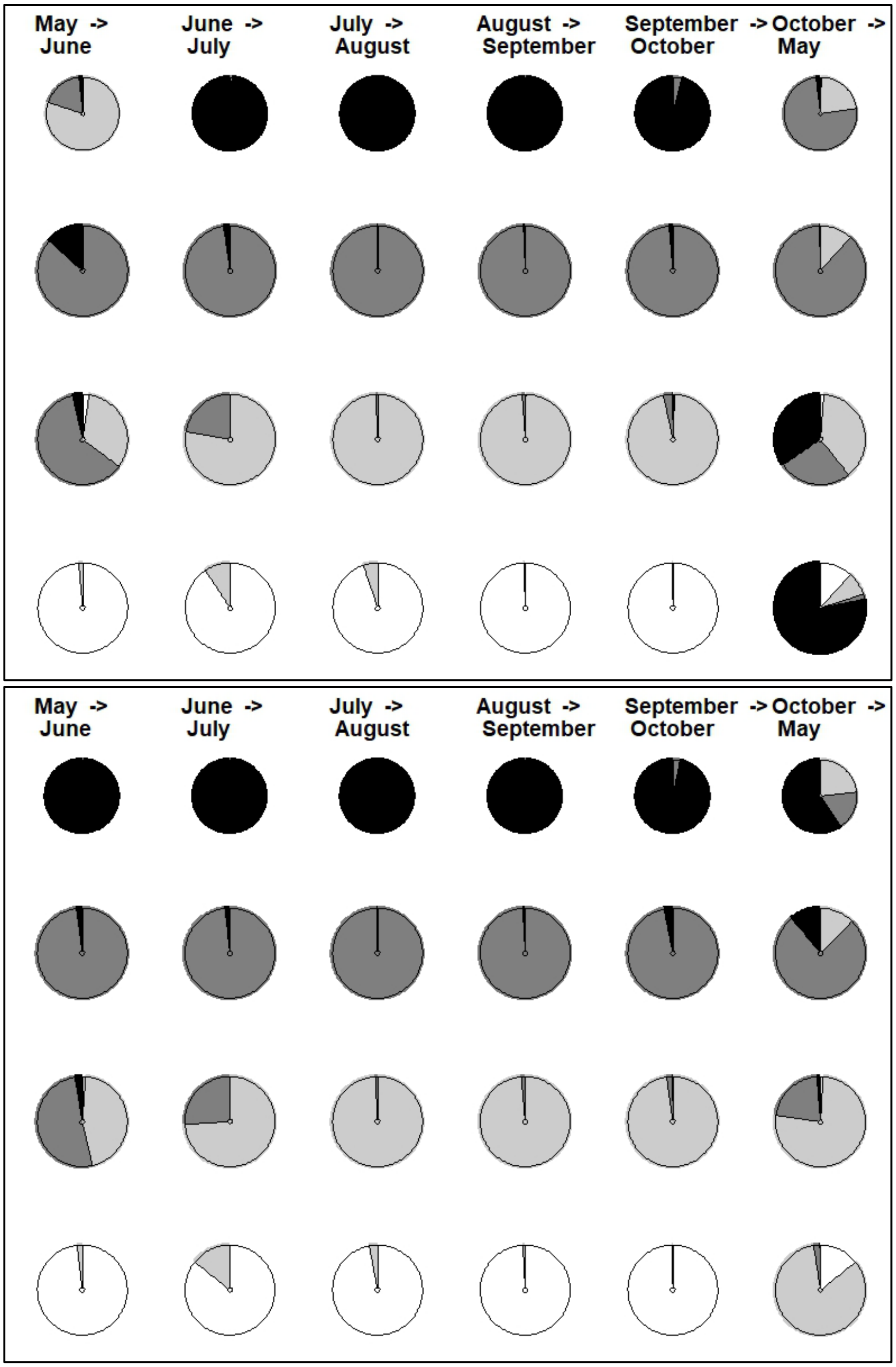


Figure 18. Migration proportions for the sensitivity runs based on the "Stepwise" approach with non-confidential effort data (top) and with the confidential effort data (bottom). The monthly time step from May to September and a six-month time step from October to May at the top of the figure. The rows indicate the region from north to south with region 1 (NE & NY), region 2 (NJ & DE), region 3 (NC & SC), and region 4 (GA & FL). The colors also indicate the region with region 1 (NE & NY) being black, region 2 (NJ & DE) being charcoal gray, region 3 (NC & SC) being light gray, and region 4 (GA & FL) being white.

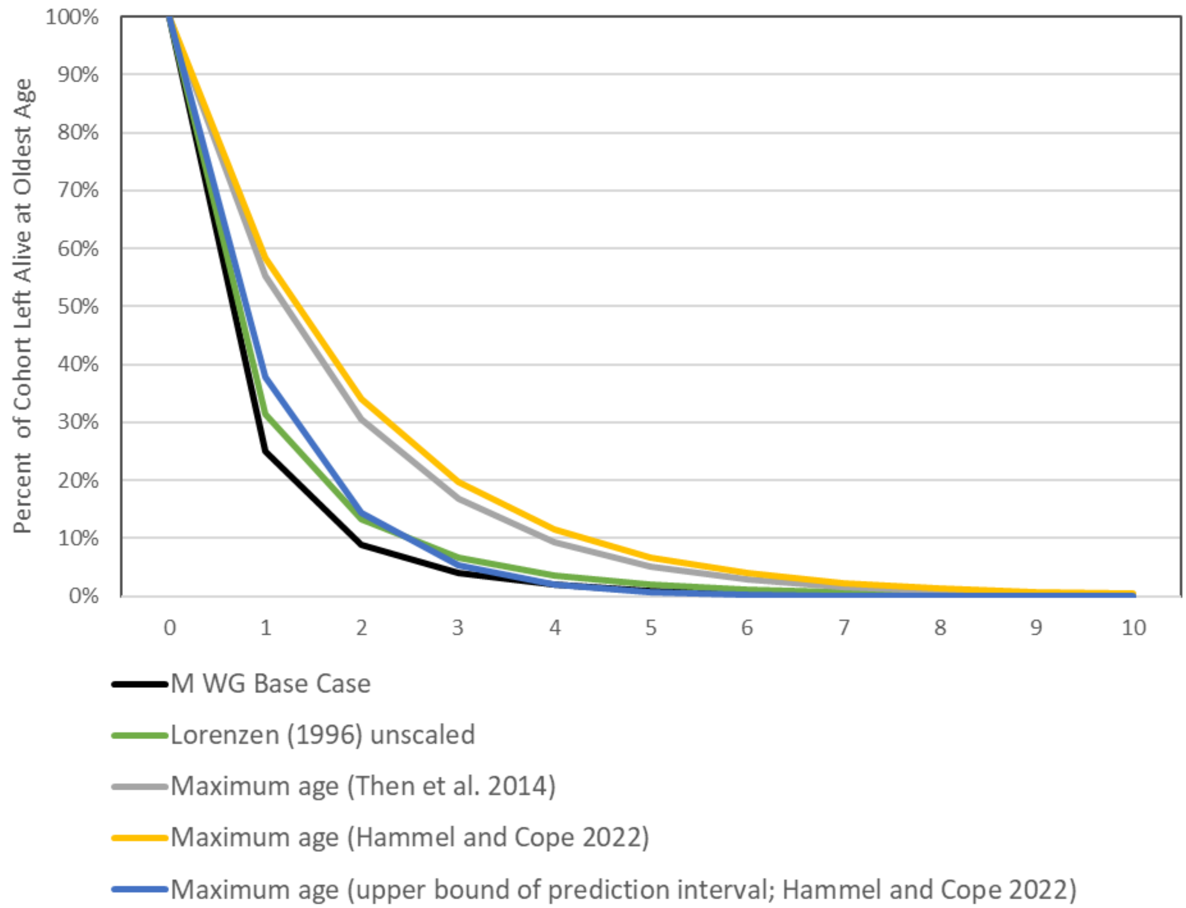


Figure 19. Percent of the cohort left alive at each age under different M estimators.

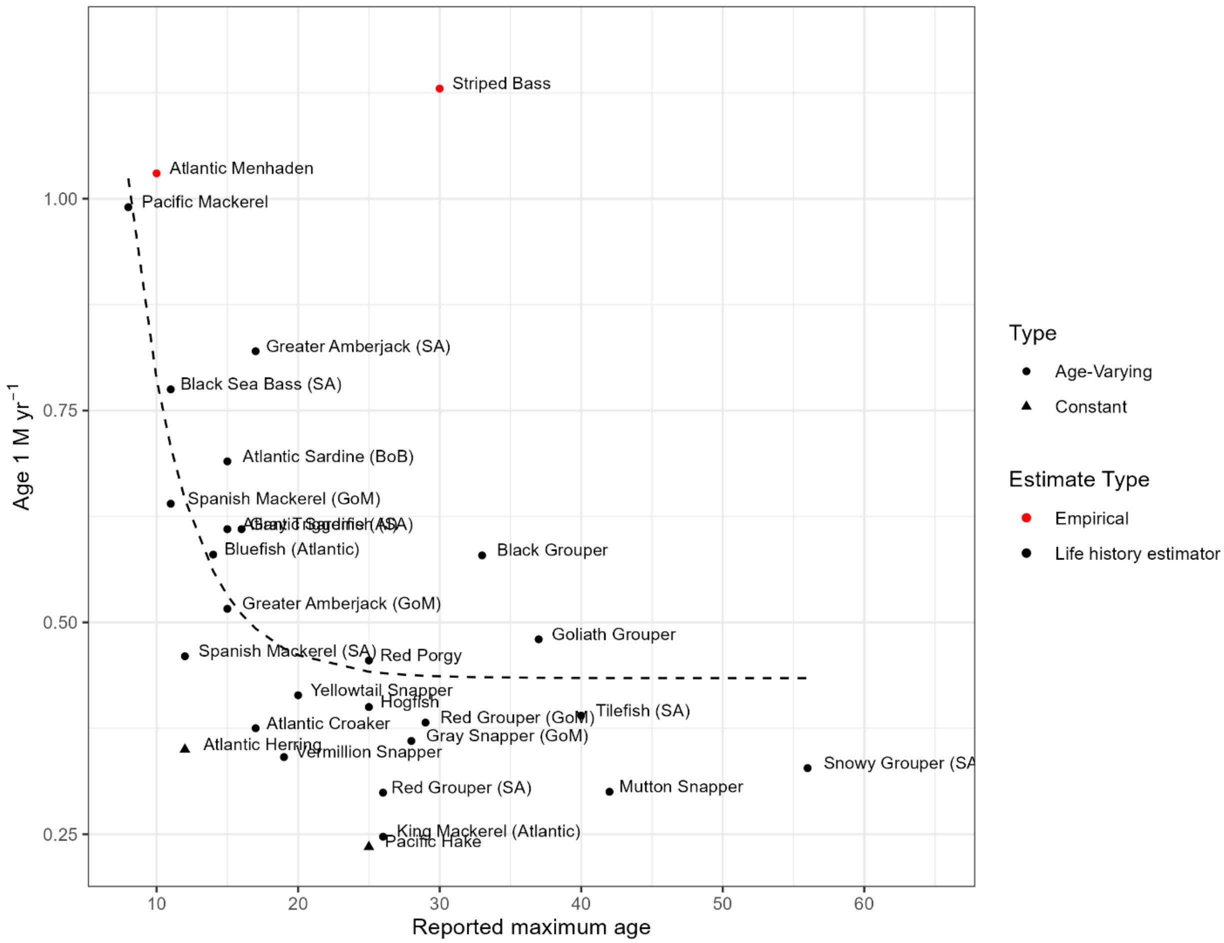


Figure 20. Estimates of age-1 M used in stock assessments plotted against reported maximum age for species selected by Ault et al. (2023). The Atlantic menhaden value is from the revised base run from the M WG. Letters in parentheses indicate stock for species with multiple stocks: (SA) = South Atlantic; (GoM) = Gulf of Mexico; (BoB) = Bay of Biscay; (AI) = Atlantic Iberian

Appendix 1: Review of the Historical Literature on Menhaden Tagging

A literature search was conducted to seek additional sources that would inform decisions that were made in the working group's analysis as to how to best estimate natural mortality (M) of Atlantic menhaden for use in stock assessments. Overall, there was not any additional literature (historic, or more recent) to help inform how to best assess the tagging recovery data from menhaden reduction plants. Besides what was already available to the group, no additional published information was found regarding schematics of reduction plants and where magnets were located within the plants. Limited published work is available to compare natural mortality estimates of other forage fish species and it may not be useful for this specific study. Work done of Gulf menhaden was also researched for comparison. Summary of the reviewed literature is below.

Ahrenholz, D.W. 1991. Population biology and life history of the North American menhadens, Brevoortia spp. Marine Fisheries Review, 53(4): 3-19.

> This paper reviews a variety of life history characteristics of different species of menhaden including: geographic range, the different species physical descriptions, migratory and spawning behavior, and general life cycle.
>
> This paper states the mean estimated natural mortality (M) for Atlantic menhaden is 0.45, and 1.1 for Gulf menhaden. The other species of menhaden (yellowfin or finescale) did not have available M estimates.
>
> These were the values used at the current time for stock assessments. These rates factored in predation, disease and bycatch.
>
> The mean estimate of 0.45 for Atlantic menhaden consisted of estimates by Reish et al (M=0.50, based on tag returns of 2- and 3-year-old fish), Dryfoos et al. (M=0.52, based on adult tagged fish), and Schaaf and Huntsman (M=0.37, based on catch statistics).
>
> The mean estimate of 1.1 for Gulf menhaden came from the mean of six estimates of M ranging from 0.69 to 1.61 which was determined using mark-recapture data.

Ahrenholz, D.W., 1981. Recruitment and exploitation of gulf menhaden, Brevoortia patronus. Fishery Bulletin, 79(2):325-335.

> This author estimates a natural mortality rate (M) for Gulf menhaden of 1.093. Tags retained in plants for more than a year were found to have minimal effect on M so they were ignored and it was noted that delayed recoveries appeared to be low in number.

Ahrenholz, D.W., Dudley, D.L., and Levi, E.J. 1991. Overview of Mark-recovery studies on adult and juvenile Atlantic menhaden, Brevoortia tyrannus, and gulf menhaden, B. patronus. Marine Fisheries Review, 53(4): 20-27.

> This paper summarizes and examines the methodologies of the Atlantic menhaden tagging study and also looks at adjustments that were needed for analysis of the data. Details on magnet setup within the plants are described including a schematic as to how these plants operated, and how tags were recovered.
>
> Some key details that were described include: description of three different types of magnets that can be used in plants (plate, grate (drawer) tubular, and rotating grate), as well as the observation that magnets adjacent to meal dryer and along conveyor system recovered magnets within days of the fish being landed as opposed to the magnets in scrap storage shed (secondary magnets). These magnets could recover tags from fish that were landed weeks or months earlier.
>
> Adjustments to the numbers of fish released had to be made to account for tag-induced mortality or tag shedding and a tag efficiency rate was calculated. Another adjustment had to be made for tags that could get lodged in the plant machinery resulting in late recoveries and potential overestimation of recovered tags for that year.

Similar to other papers, the descriptions of plant set up and magnet efficiencies is informative for our study when determining which magnets to use in the new natural mortality analysis.

Ahrenholz, D.W., Nelson, W.R. and Epperly, S. 1987. Population and fishery characteristics of Atlantic menhaden, Brevoortia tyrannus. Fishery Bulletin, 85(3): 569-600.

This paper states the annual instantaneous rate of natural mortality (M) to be 0.45, similar to other papers published by this author. This paper's main goal was to examine the stock status of Atlantic menhaden through 1981 while looking at historic events that may have led to the conditions of the stock at that time. This paper doesn't add any additional estimators of natural morality and is limited on describing methodology for calculating M's. It does not add any additional information to help inform the current task of reassessing the current natural mortality estimate.

Coston, L.C., 1971. Summary of Tags Released and Recovered for the Atlantic Menhaden, 1966-69. U.S. Dep. Commer. Nat. Mar. Fish. Serv. Data Rep. 66, 117pp.

This paper summarized the number of tags released and recaptured from 1966 through 1969 as part of a tagging study for Atlantic menhaden. The goal of this tagging study was to help determine population structure, movements, growth, and survival of Atlantic menhaden. It is important to note that in this particular report, recaptures were only counted for tags that were recovered on 'primary magnets' within the fish processing plants. The date the tags were recaptured served as a reasonable proxy for the date that the fish were caught as the primary magnets were checked routinely.

This paper lacked fine scale detail on any of the plant operations or additional information on magnet position within the plants that could help further inform this study.

Dryfoos, R. L., Cheek, R. P., and Kroger, R. L. 1973. Preliminary analyses of Atlantic menhaden, Brevoortia tyrannus, migrations, population structure, survival and exploitation rates, and availability as indicated from tag returns. Fishery Bulletin, 71(3), 719–734.

This paper examined the results from the Atlantic menhaden tagging study that was conducted from 1966 through 1969 and was useful in informing many of the decisions of this study. Ultimately, an estimate of instantaneous natural mortality (M) is derived and is estimated at 0.52 and an instantaneous fishing mortality (F) is estimated at 0.95, leading to the conclusion that 50% of the fish are being caught and 24% are dying of natural causes.

Further details and key findings of this study are:

1) 99% of recovered tags were retrieved from the magnets with most being detected within 2 weeks of the fish entering the processing plants
2) Secondary magnets might not encounter tags until a month after fish are landed.
3) Whole tagged fish that were detected by electronic recovery systems were ejected from the system, prior to coming into contact with magnet system.
4) Primary magnets were found between fish scrap drier and scrap storage area
    - 75% of tags pass over magnets same day tagged fish are processed
    - Following day an additional 10% of fish pass by the primary magnet
    - >95% of tags pass primary magnet within a 2 week time frame
    - Primary magnets usually indicated approximate date of capture

Additionally, this paper examines migrations of Atlantic menhaden. Adult menhaden migrate northward in early spring and early summer and southward in fall. Atlantic menhaden migrate farther northward as they grow older and juveniles migrate southward in fall

Kroger, R.L. and Dryfoos, R.L., 1972. Tagging and tag-recovery experiments with Atlantic menhaden, Brevoortia tyrannus. NOAA Tech. Rep. U.S. Dept. Commer. NMFS SSRF664, 11pp.

This paper highlights the desirability of mark-recapture for estimating life history parameters and summarizes many different experiments in helping to decide the best tagging methodology.

One key finding was that primary magnets and electronic detectors are valuable in recovering the ferromagnetic tags, but secondary magnets do not show the same utility as they can't help to inform the time of capture. An additional key finding was that there was a lower mortality rate for treated tags but this was offset by higher shedding of tags that were treated with antibiotics. Antibiotics result in higher shedding because of longer healing time due to antibiotics. These findings led to the conclusion that clean, dry untreated tags provided the lowest rate of tag loss. It was determined through the variety of experiments that both adult and juvenile tags can be efficiently and economically recovered on magnets from dry fish meal at reduction plants.

A recovery rate of adult tags was calculated and ranged from 55 to 90% depending on the type and location of magnets as well as the individual plant conveyor systems. Additionally, it was noted that the smaller juvenile tags had about a 70% recovery efficiency rate when compared to the recovery rate of adult tags.

Nicholson, W.R., 1978. Movements and population structure of Atlantic menhaden indicated by tag returns. Estuaries 1, 141–150.

This paper discusses population movement and age structure within the population while focusing on different areas of Atlantic menhaden up and down the coast. It does produce an estimate of M, which was the primary focus of this work group. It contains a brief summary of tagging and recovery methods used in the large-scale menhaden tagging program, including a description of primary magnet locations within the plants. There was no novel information presented in this paper regarding magnet locations or more details on plant operations.

Parker, Jr., R.O., 1973. Menhaden Tagging and Recovery: Part II- Recovery of Internal Ferromagnetic Tags Used to Mark Menhaden Genus Brevoortia. NOAA Technical Report.

This paper examined the efficiency rate of recovery of internal ferromagnetic tags in menhaden reduction plants. A brief summary and description of the processing procedure and magnet types (both found originally within the plants, and additionally added one for this study) are described within this work.

Efficiency rates for tag recoveries were estimated and described in the abstract. Adult menhaden tagged with large tags (14.0 x 3.0 x 0.5 mm) had an average efficiency rate of 64% with a range from 14-97%. Juvenile menhaden tags had an average 39% recovery rate, with a range of 5 to 86%. Juvenile tags were smaller and measured 7.0 x 2.5 x 0.4 mm. Additionally, it was noted in the abstract that 'magnet installations do not interfere with the reduction operation, but facilitate it by removing scrap metal.'

No additional information was found to help inform the working group on decisions pertaining to magnet set up and tag recaptures and efficiency rates.

Parker, R. O., Jr. 1972. An electronic detector system for recovering internally tagged menhaden genus Brevoortia. NOAA technical report NMFS SSRF 654.

This paper examines the use of an electronic detector-recovery system for fish with internal ferromagnetic tags. Menhaden on the Atlantic coast were generally recovered the day of capture (or close to), so tag recoveries made it easier to assume area and date of capture. Gulf coast vessels tended to stay out for longer making it difficult to assume where the fish were caught, and the day they were caught. This detector-recovery system was intended to help improve recovery of tags with the Gulf fishery and could potentially be used onboard a vessel. A single detector had a 36% efficiency, and a double detector had a 60% efficiency. 64% of the tags were missed using this system but eventually they were recovered by the magnets. It was found that this system was effective for recovering tags but ultimately the system did not need to be used as it was costly for menhaden studies and the magnets were recovering tags adequately. Additional information on best location for injected tags was discussed in this paper.

Pristas, P.J. and Willis, T.D., 1973. Menhaden tagging and recovery: part 1- Field methods for tagging menhaden, genus Brevoortia. Mar. Fish. Rev. 35, 31–35.

This paper described the tagging methodology for the tagging program of Atlantic menhaden, which ran from 1966 through 1971. In 1968 the program was expanded to include Gulf menhaden in addition to Atlantic.

It was determined from previous studies that internal ferromagnetic stainless steel tags measuring 14.0 x 3.0 x 0.5 mm were best suited for menhaden greater than 100 mm. These tags had a unique six-digit code. The smaller tag, used for menhaden less than 100 mm measured 7.0 x 2.5 x 0.4 mm and had a three-digit identifying code. It was noted that 100 smaller tags have the same identifying three-digit code.

This paper further details the methods for acquiring and retaining the fish, as well as the exact methodology for tagging the fish. There is a brief summary of additional studies examining survival rates of fish released in large batches vs released in groups, with it being noted that those fish released individually had better survival rates.

Reish, R.L., Deriso, R.B., Ruppert, D., Carroll, R.J., 1985. An investigation of the population dynamics of Atlantic menhaden (Brevoortia tyrannus). Can. J. Fish. Aquat. Sci. 42, 147–158.

Abstract discusses that the age-specific natural mortality estimate seems biologically reasonable, except age-1 seems low. There is no mention of the value of the natural mortality estimate used, but this paper is referenced in Ahrenholz 1991 and mentions a 0.50 estimate of M.

Vaughan. 1996. Population Characteristics of Gulf Menhaden, Brevoortia patronus. Fishery Bulletin. NOAA Technical Report NMFS 125.

This paper notes that Gulf menhaden exhibit high natural mortality and short lifespans. Thus, in comparison to Atlantic menhaden, they have higher natural mortality rates and are shorter lived. Additionally, it notes that the Gulf menhaden population is highly responsive to environmental and fishing pressures. The annual instantaneous mortality of 1.1 from Ahrenholz (1981) was discussed as long as the proposed range of 0.7 to 1.6 from tagging studies. Additionally, there was discussion about estimates from Hoenig (1983) that ranged from 0.7 for age 6 and 1.1 for age 4 fish, and lastly estimate from Pauly (1979) that presents a range for M of 0.9 to 1.1. Ultimately an M of 1.1 was used. Further, the relationship between M and fishing mortality (F) is discussed noting that when lower estimates of M are used, estimates of F are increased.

This author estimates a constant instantaneous natural mortality rate (M) for Gulf menhaden of 1.08 with a 95% confidence interval of 1.04 to 1.12 which validates Ahrenholz's 1981 estimate and is only slightly lower than the Ahrenholz recommended rate. Under the base model, the author consistently estimated a monthly M of 0.09 and an annual rate of 1.08. This estimated M likely represents fish aged 1 or 2 due to the selectivity of the fishery on those ages. Tagging mortality was based on Atlantic menhaden tagging mortality due to the lack of adult Gulf menhaden tagging mortality data. Magnet efficiency was also examined and estimated for each year of the study.