Comparing scale and otolith age estimates to understand the potential bias of long-term American shad age data in the Penobscot River

Camille Wilson
Chesapeake, Virginia

A Thesis presented to the
Faculty of the University of Virginia
in Fulfillment of the
Distinguished Majors Program

Department of Environmental Sciences
University of Virginia April 2022

Max C. N. Castorani,
Thesis Advisor

Joe Zydlewski,
Thesis Co-Advisor

Thomas A. Smith,
Director of Distinguished Major Program

ABSTRACT

American shad (*Alosa sapidissima*) are indigenous to the Atlantic Ocean, but despite their long history in this range, they currently represent only a fraction of their historical abundance. Since they are an ecologically and commercially important species that has experienced an extensive decline, it is important to understand the demographics of the populations to best manage stocks and promote sustainable harvest. One key aspect of demography is population age structure. Fish age can be estimated by reading the annuli found on scales or otoliths. For roughly a decade, scales have been collected from shad within the Penobscot River to estimate the age structure of this population over time. While studies have found that otoliths provide a more accurate estimate of shad age, it is impossible to retrieve otoliths while keeping the fish alive, so this study has relied on scales alone for past age estimates. In 2020 and 2021, 185 dead shad were retrieved from the Milford Dam fish lift for this study, and both scales and otoliths were taken from each individual fish. For each fish, I compared the otolith age estimate to the scale age estimate, as well as the precision of age determinations among readers for each sample. I concluded the otoliths were the easiest to read with the most precision among both readers, and the scales will underestimate the age of older fish and overestimate the age of younger fish. This information on the exact extent of these differences might be useful in understanding the bias of previously recorded scale data in order to get a more accurate depiction of the age structure of American shad in the Penobscot river over time.

ACKNOWLEDGMENTS

I would like to thank the NOAA Hollings Scholarship and the University of Maine for their generous funding and providing resources to conduct this research. I also thank Erin Peterson and Joe Zydlewski for their helpful instruction in working with American shad age structures as well as David Hugo, another intern without whom we would not have finished processing the samples in time. I am also extremely grateful to my faculty advisor Max Castorani who introduced me to my first research experience as a lab technician and has helped with anything I have needed in completing this thesis.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

INTRODUCTION

American shad (*Alosa sapidissima*) are native to the western Atlantic Ocean with a historic range from rivers in central Florida to southeastern Canada. American shad are anadromous, so they migrate upstream from the ocean to the river habitat to spawn (Limburg, 2003). Shad spend the majority of their life in the ocean, but when they make these migrations up and downstream, they often have to pass through dams that fragment the rivers. These dams can have negative impacts on shad populations by both blocking access to historic upstream habitats and by directly reducing shad survival due to death or injury during dam passage (Hall, 2011).

American shad were once among the top commercially fished species, but overfishing and waterway modification have caused widespread declines. This has led to a drop in harvest levels that has persisted into the 21st century (Limburg, 2003). Even with this decline, female shad are a commercially and recreationally sought-after fish for their roe. While their commercial importance may have diminished, they are still an ecologically important species. American shad mainly feed on small animals like zooplankton and insects while many other fish species and larger predators rely on shad as a food source, thus making shad an important forage base in many ecosystems and providing an important link in the food web from lower to higher trophic levels (Limburg, 2003; Elzey, 2015).

Due to their ecological importance and conservation status, shad are regarded as a species that should be restored and managed. Shad are currently protected under the Anadromous Fish Conservation Act and managed in the United States under the Interstate Fishery Management Plan for Shad and River Herring. This plan requires annual biological data to be collected for American shad from Florida to Maine. This includes age data, which can be particularly useful in estimating mortality, determining age of recruitment into spawning populations, and quantifying age structure by sex (Elzey, 2015). All of these metrics are extremely important when determining how to best manage a species. However, until a study conducted in 2010-2011, there was little information on the demographics of shad in the Penobscot River, Maine, USA (Grote, 2014). This study marks the beginning of a long-term collection of American shad age data in this river.

The Penobscot River is the second largest river in New England, and it is important for American shad management for both cultural and biological reasons. Before their decline, American shad were important food sources for Native American tribes in Maine, and they are still culturally significant to these tribes today (Government of Maine, 2016). Additionally, at high latitudes, American shad populations exhibit iteroparity. Iteroparous fish reproduce more than once time in a lifetime, and these repeat spawners are very important to the survival of populations. Repeat spawners tend to be larger with a higher fecundity, making them extremely important to the reproductive success of populations, and studies suggest the Penobscot River should have a very high frequency of iteroparity (Limburg, 2003; Grote, 2014). Early findings of this long-term study in the Penobscot River support this with 95% of the sampled shad in 2010 being repeat spawners and 75.4% being repeat spawners in 2011 (Grote, 2014). While other American shad populations in the Northeast are stable, American shad populations in Maine are still declining (Government of Maine, 2016). This makes it even more important to have the best data possible in order to make the right management decisions.

Additionally, due to the long-term nature of this study and the study location being the Penobscot River, there is the unique ability to assess the demographics of American Shad before and after the removal of two dams, the Great Works Dam in 2012 and the Veazie Dam in 2013, with the Penobscot River Restoration Project (Grote, 2014). The assessment of a population after habitat restoration can have large conservation implications, so it is important that the data is as accurate as it can be.

Now over a decade after this study began, data are abundant for understanding the changes in age structure in this population as the river has changed over time. However, the study’s methods may have led to bias. There are several structures in shad that can be used to estimate the age of the fish, but not all of them are created equal. This past data has been collected by retrieving scales from individual fish and reading the annuli to determine age (Grote, 2014), but numerous studies have demonstrated that otoliths (ear bones in fish) are much easier to read, and their annuli are much better indicators of age for American Shad (McBride, 2005; Elzey, 2015; Duffy, 2011). Otoliths cannot be retrieved without killing the fish, so while it is not possible to collect these structures to the extent needed to have sufficient data, it is possible to calibrate the existing scale data using otoliths if necessary.

In order to calibrate past data, it’s important to have reference data that was collected under as similar conditions as possible. While extensive research has been done regarding the validity of using scales and otoliths for aging in general, a comprehensive analysis that focuses on my population and study system of interest (American Shad in the Penobscot River) is missing. Past studies might not have the data I am looking for, but they provide concrete expectations and well-defined methods that I can use in my own experiment (McBride, 2005; Elzey, 2015; Marcy, 1961; Duffy, 2011). In 2020 and 2021, 185 dead American Shad were retrieved from the Milford Dam fish lift, and both otoliths and scales were taken from each individual fish. The goal of this study is to provide information to understand the potential biases of past age estimates by comparing the otolith and scale age determinations of these fish. Of particular interest are the frequency of disagreements, the magnitude of the disagreements, whether or not scales tend to underestimate or overestimate age, and how these variables change with fish age class.

It is well known that otolith annuli are a better indicator of age for American Shad and various other fish, so I predicted that otolith age determinations would be more precise among readers. I also expected frequent disagreements between the otolith and scale age determinations. Other studies also suggest that American Shad scales tend to overestimate the age for younger fish and underestimate the age for older fish (Elzey, 2015; Duffy, 2011). I also hypothesized that the age determinations would generally fall within two years of each other based on past studies (Elzey, 2015). While these precedents are useful, it’s important that we have data from the particular population and study system of interest in order to best calibrate the previous data.

METHODS

Existing Data

There is an existing dataset that includes sex, tail length, age estimates, and number of previous spawnings for American Shad in the Penobscot river system. This dataset includes roughly a decade of raw data that has been collected since 2010. I added the age estimates and spawning information for the new scales and otoliths to this dataset. This study is focused on analyzing the age estimates of the newly acquired scales and otoliths as outlined below, but the results of this analysis might provide insight in order to calibrate the age estimates of this dataset.

Study Site

As anadromous fish, American shad swim upstream from the ocean to river habitats to spawn. They pass through several dams during their river passage in the Penobscot river, and the first dam they encounter swimming upstream is the Milford Dam. The Milford Dam fish lift is a new feature to help make fish passage easier for different diadromous fish species as well as provide opportunities to study these populations. Instead of passing directly through the dam, the fish are ushered into the fishway entrance and into the lift hopper, which will essentially act as an elevator to transport the fish up and out the dam flume (Figure 1a). While the lift is meant to lower the mortality rates of the fish that pass through the dam, there are still fish that die within it or before it. These mortalities pass through the flume just as living, swimming fish, but they are taken out at the trap facility (Figure 1b). Both otoliths and scales were sourced from these mortalities to conduct this study.

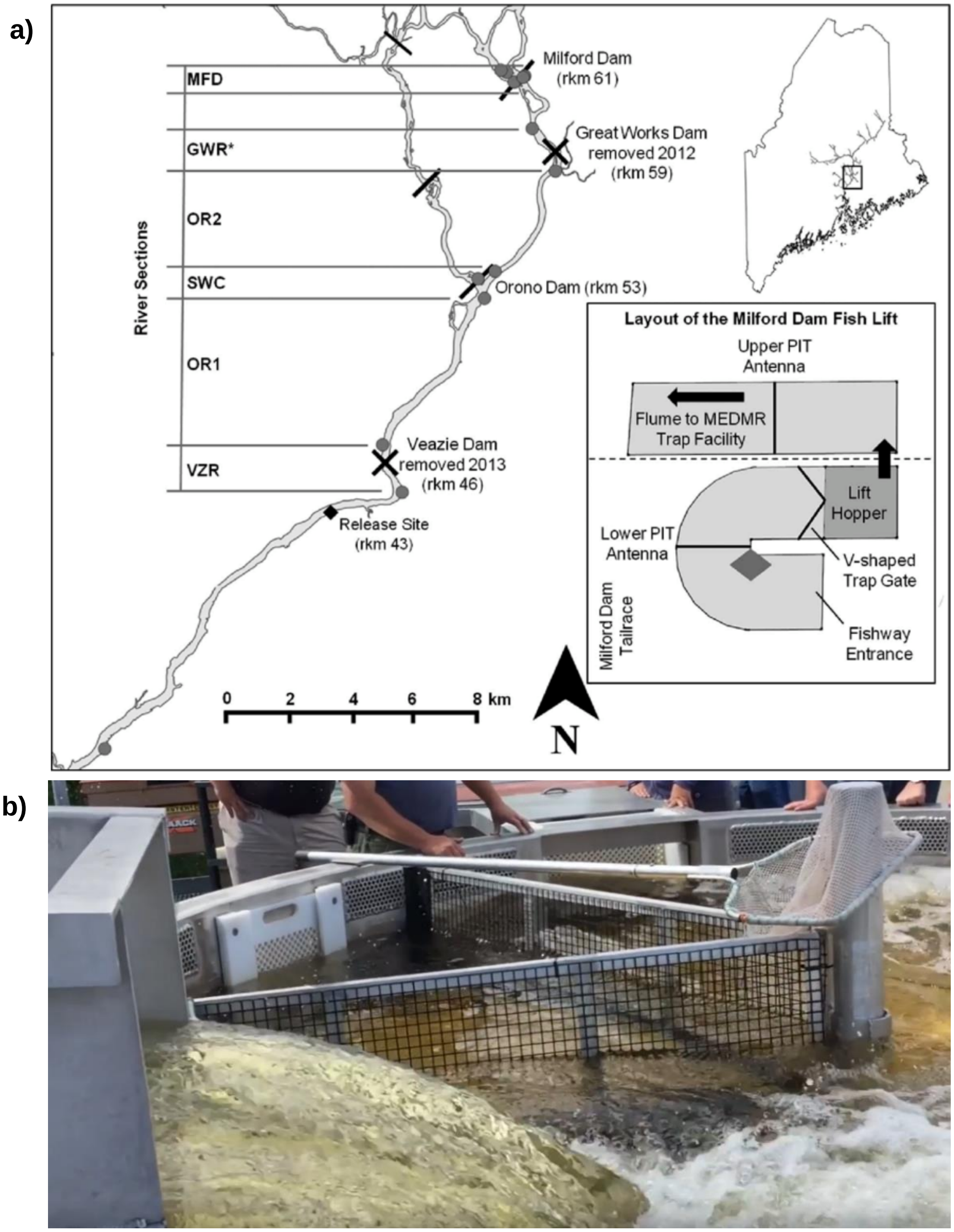

Figure 1. Milford Dam fish lift along the Penobscot River. (a) A map of the study site including the dams along the Penobscot River as well as a diagram of the Milford Dam Fish Lift. Photo credit: Izzo et al, 2016 (b) A picture of the flume fish travel through after the Milford Dam Fish Lift and where the mortalities were collected. Photo credit: NOAA Fisheries

Laboratory Methods

Both scales and otoliths were retrieved from 175 American shad mortalities within the Milford Dam fish lift in 2021 and from an additional 10 American Shad mortalities that were retrieved from the lift in 2020. Scales were taken from the left side of each fish, just below the dorsal fin, using locked hemostats. The scales from each fish were placed in labeled dry envelopes to be processed later. I needed to clean each scale in distilled water before we could read them. To clean them, I rubbed each scale between two fingers until all the residue was gone. Once cleaned and dried until clear, I placed up to 8 scales from an individual fish in between two glass microscope slides, taped the two slides together, and labeled them with the year and sample number. Two people individually looked at each scale sample to determine the age estimate, and they will be referred to as readers. Once the scales were mounted, each reader viewed them under a microfiche reader that projected a magnified image of the scale onto a screen. Both readers individually determined age estimates based on Cating's method. This entails counting the total number of spawning marks and annuli and adding one year to the age determination for the edge of the scale (Cating, 1953). Readers recorded fish age, whether or not it had spawned, and the number of previous spawnings. We repeated this process for all 185 scale samples. Of the 185 samples, we were unable to read 10 due to poor annuli visibility or tears in the scale. Once all scales were read, I compiled it into the existing raw dataset.

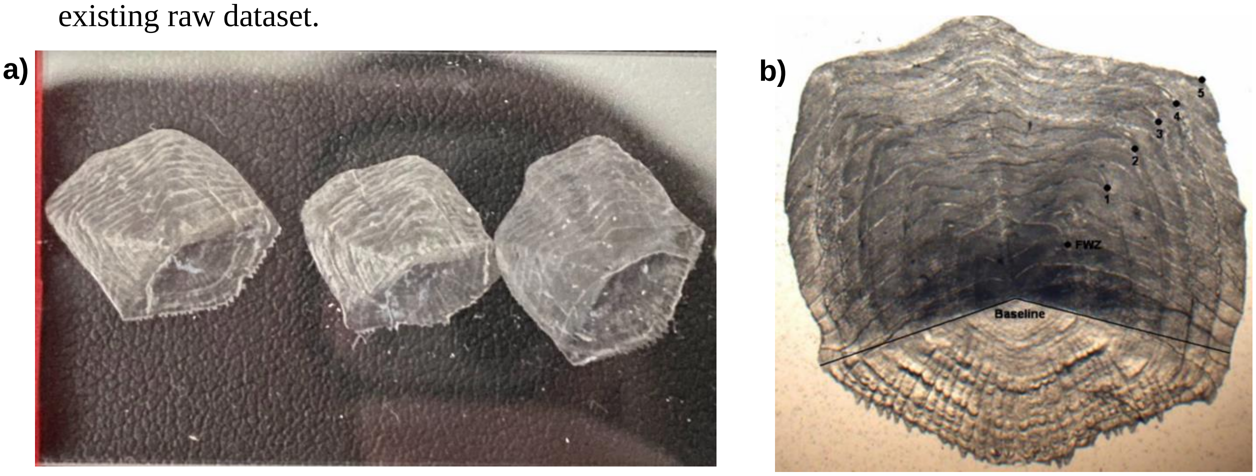

Figure 2. American shad scales. (a) A picture of three shad scales mounted between two microscope slides to be read. (b) A shad scale as seen on a microfiche reader with each dot indicating an annulus used to estimate the age. Photo credit: Elzey, 2015

In order to prepare the otoliths, I cleaned the samples from each fish by placing them in distilled water and removing the residue with small brushes and picks. Each sample had two otoliths from a single fish, and if one was already broken or broken during the cleaning process, then I set it aside and did not mount it onto a slide. Of the 185, there were 12 total samples that had no viable otoliths for reading. I attached the unbroken otoliths to the top of a single microscope slide using clear nail polish. It was important to place each otolith sulcus (groove) side down and flat side up to ensure that the annuli were as clearly visible as possible. For comparison, I placed both left and right otoliths of an individual fish on a single slide. I then labeled the slide with the year and sample number. Then, the same two readers used a dissecting microscope to count the presumed annuli, which we defined as the far edge of each hyaline zone. Both readers counted the annuli plus the edge of the otolith for the age determination (Elzey, 2015). Again, we repeated this process for all samples, and I added this data to the existing dataset as well.

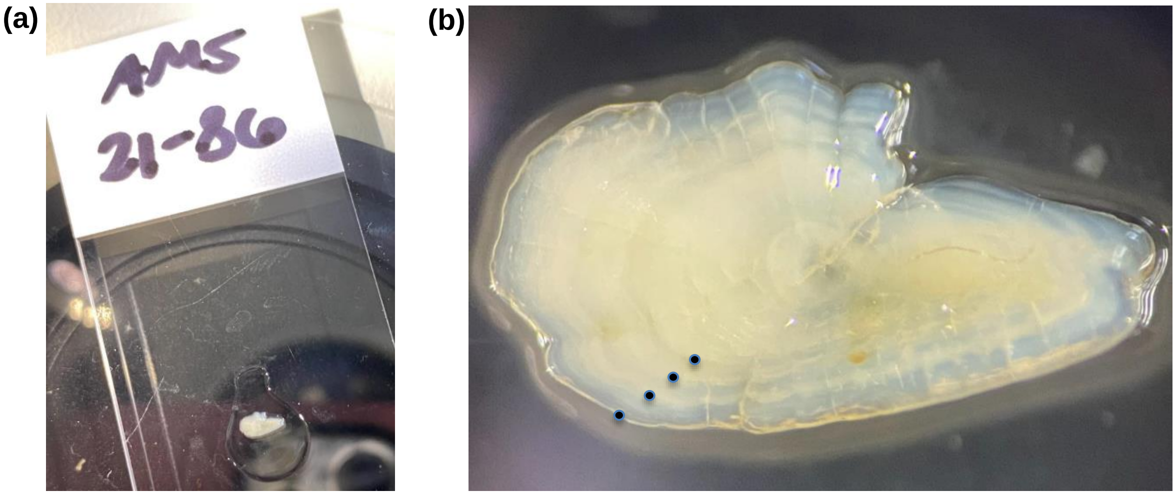

Figure 3. American shad otoliths. (a) A picture of a shad otolith mounted onto a labelled microscope slide to be read. (b) A picture of what the otolith looks like under a dissecting microscope where each dot indicates an annulus used to estimate the age.

Analysis

Once the dataset was ready, I used it to assess the precision of the age determinations between both readers for both scales and otoliths as well as the precision within each reader comparing the scale and otolith age estimate of each fish. I will calculate this by measuring percent agreement using the formula:

$$\frac{\#\, of\ times\ readers\ come\ to\ the\ same\ age\ determination}{Total\ number\ of\ readings} \times 100$$

where higher percent agreement will indicate more precision. With only two readers, their age estimates must match exactly in order to be in agreement. However, I included different percent agreements with precision within 1 year, 2 years, and 3 years for analysis as well (Elzey, 2015). I also calculated the coefficient of variation using:

$$CV_j = 100 \text{ x } \frac{\sqrt{\sum_{i=1}^{R} \frac{(X_{ij} - X_j)^2}{R-1}}}{X_j}$$

CVj = CV for the jth fish
Xij = ith age estimate of the jth fish
Xj = average age estimate of jth fish
R = number of times fish was read

where a smaller coefficient indicates more precision (Campana, 2001). This value can be averaged across all samples of the same type (i.e. all scales, all otoliths, reader 1 estimates, or reader 2 estimates). This value standardizes the variation in age estimates relative to the mean of the age estimates. Finally, I also used root mean square error estimates using:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{N}(O-E)^2}{N}}$$

N = number of samples that had estimates for both groups
E = expected correct age estimate; the age estimate for the group treated as the standard (otoliths and reader 2)
O = actual observed age estimate for the group being compared to the standard (scales and reader 1)

to assess the precision between readers and between aging structures where a larger value means less precision. This value quantifies how far the observations stray from the variable group I held as the standard.

Next, I compared the age estimates to find the magnitude and frequency of disagreements between the scale and otolith determinations for each fish. I also assessed if the magnitude and frequency of the disagreements change with respect to the fish's age class. Finding the frequency at which this happens and the average and range of magnitudes is necessary to better gauge the accuracy of past estimates. Finally, I was interested if there were any significant trends of scales over/underestimating age if the estimates differ and whether or not this trend had a linear or non-linear relationship with respect to the presumed age of the fish.

RESULTS

In order to determine that otoliths are in fact the better age structure to use for American shad in this system, I compared the agreement in the age estimates between readers. This includes several statistical tests to assess the error including percent agreement (PA), coefficients of variation (CV), and root mean square error (RMSE). When I compared the age estimate of both scales and otoliths from reader1 to reader 2, scales had a lower percent agreement and a higher coefficient of variation and root mean square error. The higher percent agreement by over 25% among otoliths means that the readers agree on the age estimate of an otolith ~25% more of the time than they agree on the age estimate of a scale.

	PA	CV	RMSE
Between Readers			
Scales	58.86%	6.00	0.73
Otoliths	84.05%	3.19	0.51

Table 1. Comparison of the precision of scale and otolith age estimates by comparing the readings reader 1 and reader 2 had for the same scale/otolith.

The magnitude of the disagreement ranged from 0 (perfect agreement between readers) to 3 years (the difference between reader 1 and reader 2’s age determination was 3 years) for otoliths and 0 to 2 years for scales. When I looked more specifically at the magnitude of disagreements between the two readers, I found that most age estimates for both scale and otoliths fall within 1 year of the other reader’s age estimate. All scale age estimates for reader 1 were within 2 years of reader 2’s age estimate, while the magnitude of disagreement between the age estimates for one otolith sample was 3 years.

	PA +/- 0	PA +/- 1	PA +/- 2	PA +/- 3
Between Readers				
Scales	58.86%	96.57%	100%	100%
Otoliths	84.05%	99.42%	99.42%	100%

Table 2. Percent agreement between reader 1 and reader 2's age estimates across all scales and across all otoliths. This includes the percent of the time the readers' age estimates are in perfect agreement, within 1 year, 2 years, or three years of each other.

I was also interested in the effect of fish age on the agreement between readers for both scales and otoliths. For both scales and otoliths, the average age estimates for young and old shad tend to stray further from the 1:1 line representing perfect agreement. Additionally, on average, reader 1 overestimates the age of the youngest age class and underestimates the age for the oldest age class for both scales and otoliths when compared to reader 2. These younger and older age estimates make up a very small minority (reader 2 only determined 1 otolith to be 3 years old and 2 otoliths to be 7 years old), and by far the most common age estimates were 5 and 6 years for both scales and otoliths. When looking at just these two points, the otolith age estimates much more closely match the 1:1 line of perfect agreement compared to the scales, and their standard deviation is much smaller as well.

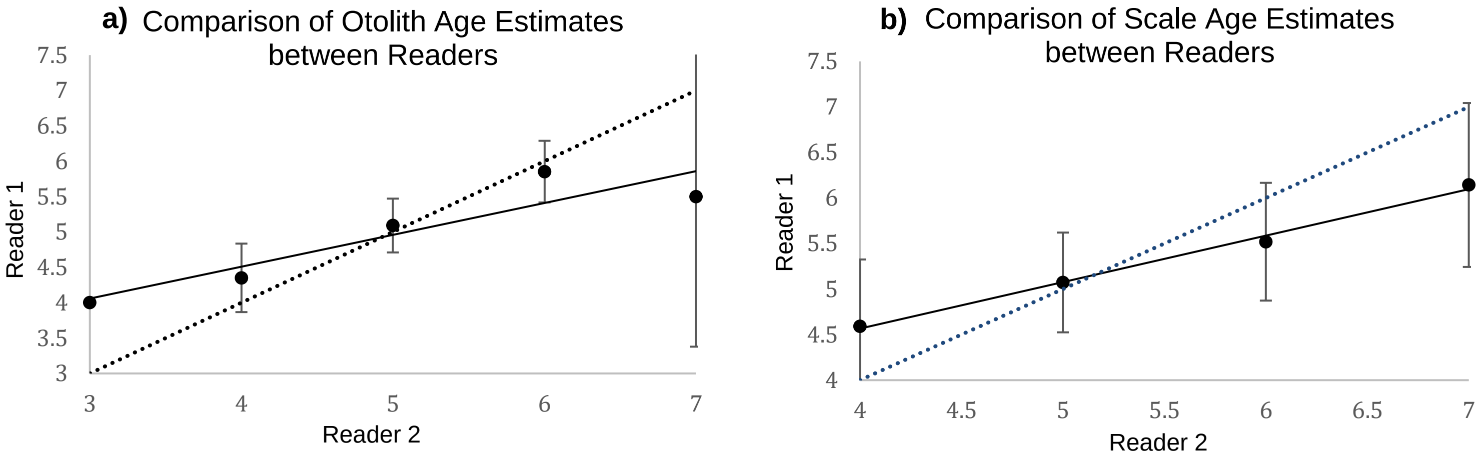

Figure 4: Comparing trends in the differences between the readers' age estimates for both scales and otoliths. (a) The points represent the average age estimate reader 1 determined from an otolith (y axis) whenever reader 2 estimated the age to be 3,4,5,6, or 7 respectively (x axis). (b) The points represent the average age estimate reader 1 determined from a scale (y axis) whenever reader 2 estimated the age to be 4,5,6, or 7 respectively (x axis). (a and b) The solid lines represent a trend line between the points, the dotted lines represent a 1:1 line that represent the relationship if all age estimates were in agreement, and the error bars represent one standard deviation above and below the mean.

To actually compare the differences in the age estimates determined from scales and otoliths, I used comparisons within readers using the age estimates a single reader determined from the otoliths and scales of the same fish. Again, I used percent agreement, coefficient of variation, and root mean square error to quantify these differences. For percent agreement for both readers, the value falls below 50%, so the scale age estimate matched the otolith age estimate less than half of the time. The coefficient of variation values are both very large (>> 1), indicating that each age estimate for a fish varies greatly from the average age estimate calculated from the scale and otolith age estimate of that fish. Therefore, many samples had differing otolith and scale age estimates. The RMSE values are also high, indicating that the observed scale age estimate differed from the otolith age estimate that is held as the standard very often.

	PA	CV	RMSE
Scales vs. Otoliths			
Reader 1	44.17%	8.79	0.91
Reader 2	39.88%	9.21	0.89

Table 3. Assessing the accuracy of scale age estimates by comparing the scale age estimate from each fish to the corresponding otolith age estimate for both reader 1 and reader 2.

I was also interested in the magnitude of the disagreement between the age estimates of scales and otoliths for the same American shad. For both readers, over 90% of scale age estimates were within one year of the fish's corresponding otolith age estimate. All but one scale age estimate was within 2 years of the otolith age estimate for both readers, and all scale age estimates were within 3 years of the otolith age estimate.

	PA +/- 0	PA +/- 1	PA +/-2	PA +/-3
Scales vs. Otoliths				
Within Reader 1	44.17%	92.02%	99.39%	100%
Within Reader 2	39.88%	94.48%	99.39%	100%

Table 4. The percent of the time the scale age estimate agreed with the corresponding otolith age estimate for both readers separated into perfect agreement, agreement within 1 year, agreement within 2 years, and agreement within 3 years.

I also wanted to determine if there were any trends in the differences between otolith and scale age estimates across age classes. I separated the otolith age estimates into age classes of one-year increments and found the percentage of perfect agreement the scales had with an otolith of that age class. For both readers, the highest percent agreement was when otoliths estimated the age to be 5, while the youngest and oldest age, age class 3 and 7 respectively, had the lowest agreement at 0%. Following age class 5, age class 6 and age class 4 had the next highest percent agreements respectively.

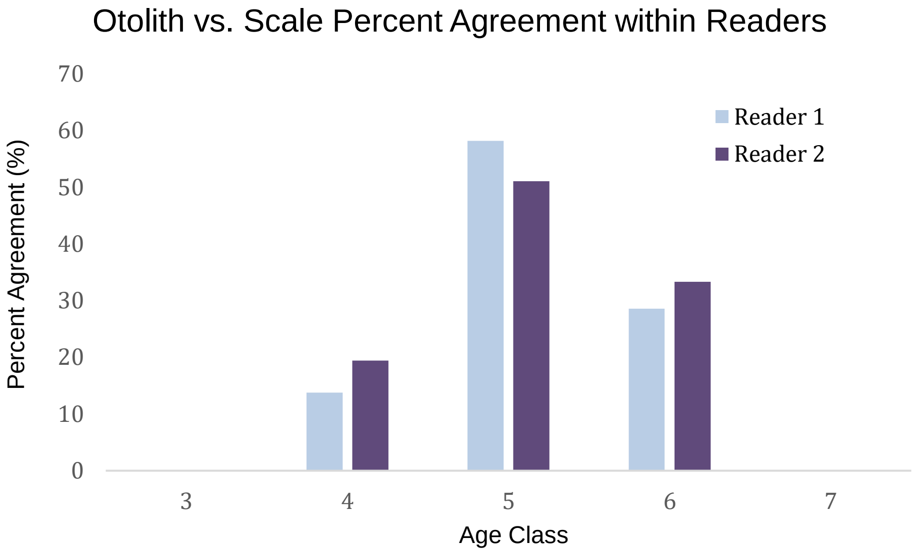

Figure 5. The percent of the time scale age estimates had perfect agreement with the corresponding otolith age estimates for each age class.

Of particular interest were the trends in the magnitude and direction of the differences. When reader 1 estimated the age of the otolith to be 4, 5, 6, or 7, the average age estimates for the corresponding scales was roughly 5. Reader 1 only estimated 2 otoliths to be 7 years old, and the reader estimated both of the corresponding scales to be 5 years old, so the average magnitude of the difference between the age estimates was two years. When reader 1 estimated the otolith to be 4 or 6, the average difference between the otolith and scale age estimates was 1 year. When reader 1 estimated the otolith to be 4, the scale was on average estimated to be one year older while the corresponding scale for 6-year-old otoliths was estimated to be one year younger on average. Similarly, reader 2 estimated the scale to be older than the otoliths for younger shad and younger than the otolith for older fish. On average, when reader 2 estimated the otolith to be 3 and 4, the scale estimate was one year older. When reader 2 estimated the otolith to be 5, the estimate for the corresponding scale was older on average, but it tended to be closer to 5. When reader 2 estimated the otolith to be 6 or 7, the corresponding scale was estimated to be younger by an average of roughly 0.5 and 1 year respectively.

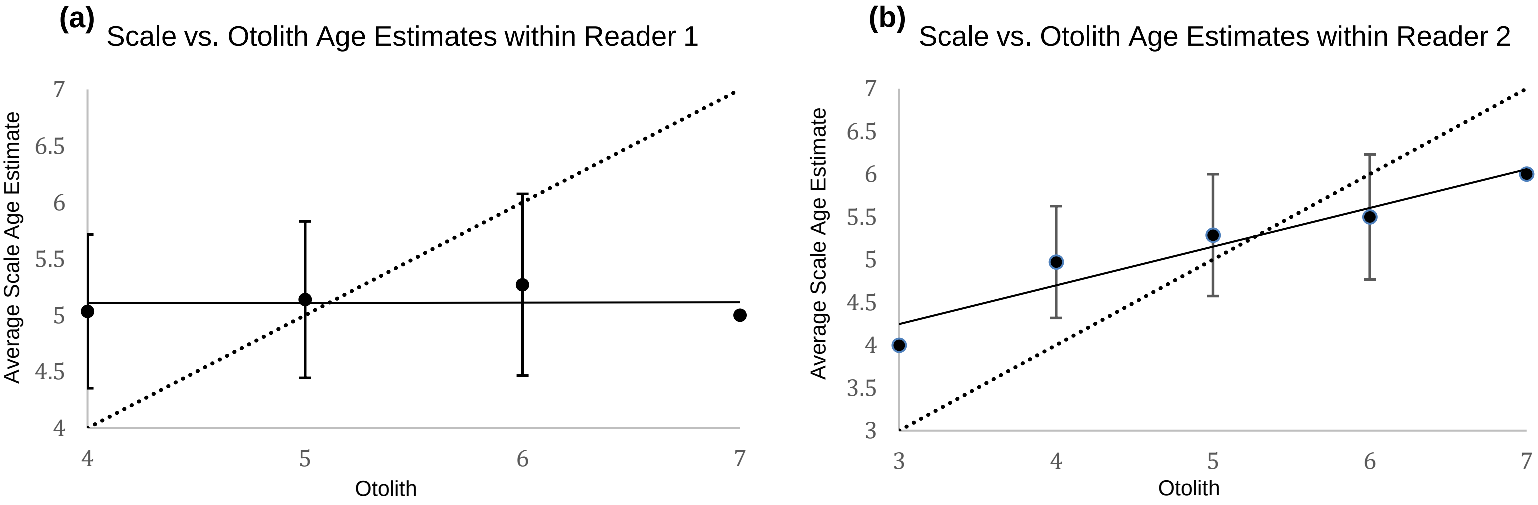

Figure 6. Comparing trends in the differences between the scale age estimate and the corresponding otolith age estimate for both readers. (a) The points represent the average age estimate reader 1 determined from a scale (y axis) whenever the corresponding otolith was estimated to be 4,5,6, or 7 respectively (x axis). (b) The points represent the average age estimate reader 2 determined from a scale (y axis) whenever the age of the corresponding otolith was estimated to be 3,4,5,6, or 7 respectively (x axis). (a and b) The solid lines represent a trend line between the points, the dotted lines represent a 1:1 line describing the relationship if all age estimates were in agreement, and the error bars represent one standard deviation above and below the mean.

Finally, the total count in each age class as determined by the scale and otolith age estimates is important to this study. For both readers, the 5-year-old scale and otolith age estimates have relatively similar frequencies at close to 100, which is over half of the total number of the samples. The frequencies of the other age classes stay relatively similar between scales and otoliths for reader 1, but they differ more for reader 2. Of the rest of the age classes, 6-year-old age estimates have the second highest frequency among scales, followed by 4, 7, and 3 for both readers. Reader 1's otolith age estimates also fit this pattern, but reader 2's year 4 otolith age estimate frequency is just slightly higher than the year 6 age estimate frequency.

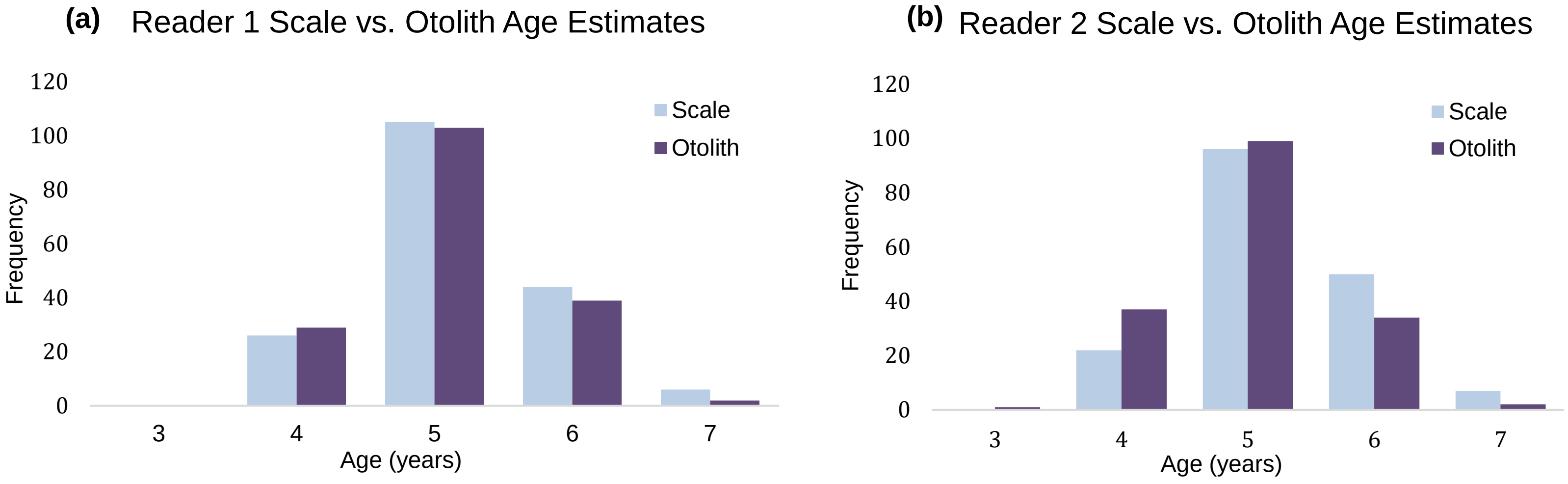

Figure 7. Comparison of the frequencies of fish in each age class as determined by each aging structure for both readers. (a) Reader 1. (b) Reader 2.

DISCUSSION

The significance of otoliths having higher percent agreement among readers and lower RMSE and CVs indicates that they are easier to read and interpret as the readers agreed on the age estimate more often. Due to this, I can conclude that with this study system and species, otoliths are indeed better structures for estimating the age of the fish compared to scales. Given this, I could use the otoliths as a standard age that is assumed to be correct and compare the corresponding scales to it. If scales had higher precision, there would be no need to do any further comparisons as the goal of this study was to be able to understand the potential bias of the past scale age estimates using otoliths, but if scales were better aging structures all along, this goal would be pointless.

I was also interested in how much the readers differed in their age estimate for a particular scale or otolith. For the most part, the reader's age estimates were within one year of the other for both scales and otoliths, so neither aging structure consistently had very different readings between readers. While it is good that neither structure is so hard to read that the readers consistently differ by several years in the age estimates, on the scale of a fish, a single year difference is a long time, and can mean very different properties and behaviors. This means it is important to choose an aging structure that is as precise as possible in order to have the best information to be able to understand and manage a population.

When observing trends of over- or underestimation and average magnitudes of disagreement among age classes, examination of regression plots can lead to different interpretations. If you look solely at the regression lines, the two aging structures seem to be very similar in how closely they follow the 1:1 relationship of perfect agreement between readers, thus implying both structures are equally as precise; however, examining the data points with more information provides a different picture. For both otoliths and scales, reader 1 overestimated the age of younger fish and underestimated the age of older fish when compared to reader 2. For otoliths, this created a trendline that deviated greatly from the 1:1 relationship line because of these outliers. However, there were very few observations where reader 2 estimated the fish to be 3 (1 observation) or 7 (2 observations) from the otolith, so this very small sample size can heavily affect the data. When ignoring these points, reader 1's average age estimates when reader 2 determined the age class to be 4, 5, and 6 respectively follow very closely to the 1:1 relationship line. This indicates much more precision when 2 readers estimate the age of the same otolith, therefore still suggesting that otoliths are the better aging structure to use for American shad in the Penobscot River.

Aside from this, one interesting finding is that when compared to reader 2, reader 1 consistently overestimated the age of younger fish and underestimated the age of older fish for both scales and otoliths even though the magnitude of this over/underestimation was much greater for scales. While I don't have enough information to conclude this entirely due to having too few readers, this does suggest that the way readers interpret scales and otoliths is consistent, and although this interpretation may differ from other readers, you might be able to better understand biases in the scale data without having to involve other aging structures by taking these consistent discrepancies into account. This hypothesis would need further testing with a comparison of the age estimates of many more readers, but this could be a direction to take the study in the future.

Even though the otoliths were determined to be better aging structures in American shad in this study, there could still be some circumstances in which using scales to estimate age would be acceptable. As seen in figure 7, the overall frequencies as determined by reader 1's scale and otolith age estimates are relatively similar for each age class. While this does not mean that the scale age estimate almost always matched the corresponding otolith age estimate, it does mean that despite these differences, the scale and otolith aging methods both came to relatively similar results overall. If this was consistent, then you might argue that using a different aging method would be unnecessary because you obtain the same results either way. However, this trend is not consistent with reader 2. The overall frequency of fish aged to be 5 years old is similar for both aging methods, but the other age classes differ a lot more. Due to the fact that one out of two readers do not follow this trend, I cannot conclude that it would persist with other readers. This could be further studied by conducting a similar study with more readers. While this is an avenue for further study, I still conclude that otoliths are much better aging structures for American shad based on the data available.

Since I could conclude that otoliths were better aging structures for our study species in our study system, the next step was to compare the scale age estimate to the otolith age estimate for each individual fish in order to understand potential bias. When I used the otolith age estimate as a standard to compare the scale age estimates to, they had perfect agreement less than half of the time for both readers. This suggests that scale age estimates determined from readers are inaccurate over 50% of the time. In turn, this suggests that the long-term American shad age data in the Penobscot River is highly biased. This high disagreement rate makes it even more important to understand the trends in the bias in order to get a better picture of what the past data should look like.

Again, most scale and otolith age estimates were within two years of each other, however, a single year difference in the age estimate compared to the real age of the fish can have major implications. A lot of management decisions are based on age-based stock assessments, so it is important that this data is as accurate as possible. Given how low the agreement is between scales and otoliths, this would mean relying on otoliths for age estimates or continuing to conduct studies like this where otoliths are used to understand the bias to calibrate scale data.

When looking more specifically at discrepancies between scale and otolith age estimates by age class, the percent agreement for the otoliths that were determined to be age class 5 is much higher for both readers compared to the other age classes. For this age class, both readers had a much higher percent agreement than their own respective average percent agreement among all age classes. This is significant because as seen in figure 7, both readers determined over half of all the age estimates for both scales and otoliths to be in age class 5. Additionally, 5-year-old fish should also be the most abundant age coming back to spawn. It is ideal that the aging method for scales is most accurate for the most abundant age class, but even though this percent agreement is much higher than the percent agreement for other age classes, it is still less than optimal. This percent agreement between 50% and 60% for both readers still indicates that a large portion of the time (40–50%), the scale age estimate is inaccurate for this age class, and this does not even take into account the lower accuracy of the method for other age classes.

There were also major trends of under and overestimation when the scale age estimate was compared to the otolith age estimate that was held as the standard. For both readers, the scale age estimates overestimated the age of younger fish and underestimated the age of older fish when compared to the otolith age estimates. This makes sense because as American shad spawn, their scales start to recede backwards, making annuli less visible and making it easier to age a fish younger than what it really is. Additionally, since scales are not significantly smaller for younger fish, it is a lot easier to count a “false annulus” in all of the extra space that is there in younger fish’s scale, therefore aging the fish as older than it really is when it is young. While it is not ideal that there are major biases for the aging structure that is much easier to collect, it is good that these trends are consistent. This data can easily be applied to the long-term American shad age data. For example, there are consistent trends where the scale age estimate is 5 when the corresponding otolith determined the age to be 4. Given the scale-otolith percent agreement and this average difference between the scale and otolith age estimates for that age class, we can calibrate the past data to better represent the actual population and not this bias.

CONCLUSION

As expected, otoliths are the better aging structure for American shad in the Penobscot River during our study period. I can conclude this based on the higher precision between the age estimates of both readers for a single otolith compared to a single scale. This precision varies among age classes, with the age estimates of younger and older fish being less precise between readers for both scales and otoliths, but overall, otolith age estimates are more precise than scale age estimates for all age classes. After concluding this, I could consider the otolith age estimate as the standard and therefore be able to test the scale age estimates for accuracy. Using different statistical methods, I determined that the scale age estimates were not very accurate at all, as they agreed with the otolith age estimates less than half of the time and had high error values. As other studies have suggested, the scales overestimated the age of younger fish and underestimated the age of older fish. While the scale age estimates from fish in the 5-year-old age class were more accurate, they were still not accurate enough to be able to rely on scale data alone. These major inconsistencies illustrate the need to use otoliths originally to collect age data on American shad or use otoliths in a comparison study like this one to understand the bias in the scale age estimates. This data will be very helpful when analyzing the scale age data from this past long-term study so that there is both reliable data and a sufficient amount of data. There may be some conditions where the scale age estimates are sufficient for studies on American shad in the Penobscot River (e.g., overall similar total counts for each age class for both scales and otoliths), but more research would need to be done to see if these conditions are really met.

REFERENCES

Campana, S. E. “Accuracy, Precision and Quality Control in Age Determination, Including a Review of the Use and Abuse of Age Validation Methods.” Journal of Fish Biology, vol. 59, no. 2, 2001, pp. 197–242, doi:10.1006/jfbi.2001.1668.

Cating, James P. “Determining Age of Atlantic Shad from Their Scales.” Fishery Bulletin, vol. 54, 1953, pp. 187–99.

Duffy, W. J., R.S. McBride, S.X. Cadrin, and K. Oliveira. “Is Cating’s method of transverse groove counts to annuli applicable for all stocks of American Shad?” Transactions of the American Fisheries Society, vol. 140, no. 4, 2011, pp. 1023–1034.

Elzey, Scott P., Katie A. Rogers, Kimberly J. Trull. “Comparison of 4 Aging Structures in the American Shad (Alosa Sapidissima).” Fishery Bulletin, 2015, doi:10.7755/FB.113.1.5.

Elzey, Scott P., Katie A. Rogers, Kimberly J. Trull. “Massachusetts Division of Marine Fisheries Age and Growth Laboratory: Fish Aging Protocols.” no. TR-58, 2015, pp. 1–43.

Government of Maine. “American Shad Species of Greatest Conservation Need Report.” 2016, https://www.maine.gov/ifw/wildlife/reports/pdfs/SGCN_Reports/SGCN/American%20Shad__Alosa%20sapidissima.pdf.

Grote, Ann B., Michael M Bailey, Joseph D. Zydlewski. “Movements and Demography of Spawning American Shad in the Penobscot River, Maine, Prior to Dam Removal.” Transactions of the American Fisheries Society, vol. 143, no. 2, 2014, pp. 552–63, doi:10.1080/00028487.2013.864705.

Hall, Carolyn J., Adrian Jordaan, Michael G. Frisk. “The Historic Influence of Dams on Diadromous Fish Habitat with a Focus on River Herring and Hydrologic Longitudinal Connectivity.” Landscape Ecology, vol. 26, no. 1, 2011, pp. 95–107, doi:10.1007/s10980-010-9539-1.

Izzo, Lisa K., et al. “Upstream Movements of Atlantic Salmon in the Lower Penobscot River, Maine Following Two Dam Removals and Fish Passage Modifications.” *Marine and Coastal Fisheries*, vol. 8, no. 1, Taylor & Francis, 2016, pp. 448–61, doi:10.1080/19425120.2016.1185063.

Limburg, Karin E., Kathryn A. Hattala, Andrew Kahnle. “American Shad in Its Native Range.” American Fisheries Society Symposium, vol. 2003, no. 35, 2003, pp. 125–40.

Marcy, C. “Age Determinations from Scales of Alosa Pseudoharengus (Wilson) and Alosa Aestivalis (Mitchill) in Connecticut Waters.” no. 57, 1961, pp. 622–30.

McBride, Richard S., Michael L. Hendricks, John E. Olney. “Testing the Validity of Cating’s (1953) Method for Age Determination of American Shad Using Scales.” 2005, www.fisheries.org.

NOAA Fisheries. “Milford Fish Lift Penobscot River.” Facebook, 29 June 2017, https://www.facebook.com/NOAAFisheriesNEMA/videos/milford-fish-lift-penobscot-river/1906120242959359/. Accessed 19 Mar. 2022.