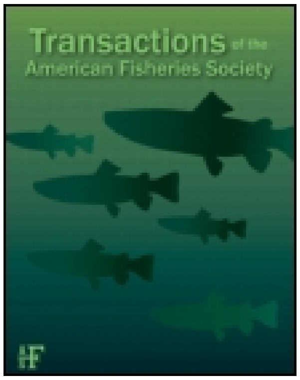

Transactions of the American Fisheries Society
Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/utaf20

Quantifying the Effects of Aging Bias in Atlantic Striped Bass Stock Assessment

Hongsheng Liao [a] , Alexei F. Sharov [b] , Cynthia M. Jones [a] & Gary A. Nelson [c]

[a] Center for Quantitative Fisheries Ecology, Old Dominion University, 800 West 46th Street, Norfolk, Virginia, 23508, USA

[b] Fisheries Service, Maryland Department of Natural Resources, 580 Taylor Avenue, Annapolis, Maryland, 21401, USA

[c] Division of Marine Fisheries, Massachusetts Department of Fish and Game, 30 Emerson Avenue, Gloucester, Massachusetts, 01930, USA

Version of record first published: 27 Dec 2012.

To cite this article: Hongsheng Liao , Alexei F. Sharov , Cynthia M. Jones & Gary A. Nelson (2013): Quantifying the Effects of Aging Bias in Atlantic Striped Bass Stock Assessment, Transactions of the American Fisheries Society, 142:1, 193-207

To link to this article: http://dx.doi.org/10.1080/00028487.2012.705255

Transactions of the American Fisheries Society 142:193–207, 2013

ISSN: 0002-8487 print / 1548-8659 online
DOI: 10.1080/00028487.2012.705255

ARTICLE

Quantifying the Effects of Aging Bias in Atlantic Striped Bass Stock Assessment

Hongsheng Liao*
Center for Quantitative Fisheries Ecology, Old Dominion University, 800 West 46th Street, Norfolk, Virginia 23508, USA

Alexei F. Sharov
Fisheries Service, Maryland Department of Natural Resources, 580 Taylor Avenue, Annapolis, Maryland 21401, USA

Cynthia M. Jones
Center for Quantitative Fisheries Ecology, Old Dominion University, 800 West 46th Street, Norfolk, Virginia 23508, USA

Gary A. Nelson
Division of Marine Fisheries, Massachusetts Department of Fish and Game, 30 Emerson Avenue, Gloucester, Massachusetts 01930, USA

Abstract

The effects of aging bias on stock assessment have been evaluated in many simulation studies. However, simulated data cannot be used to quantify the magnitude of aging biases in a real data set and thus their effects on an ongoing stock assessment. In this study, we validated scale and otolith aging using Striped Bass *Morone saxatilis* of known age collected from the Chesapeake Bay between 1997 and 2006. Otoliths provided more accurate (74% agreement) and precise (average CV = 1.9%) age estimates than scales, which overestimated the ages of young fish and underestimated the ages of old fish (22% agreement; average CV = 9.8%). Based on their accuracy, otolith ages were subsequently used to quantify aging biases in the scale ages and effects of bias in the stock assessment of Atlantic Striped Bass. We converted scale-age to otolith-age input using paired scale and otolith ages of Striped Bass collected by commercial fisheries in Virginia waters of the Chesapeake Bay and the mid-Atlantic from 1999 to 2010. A statistical catch-at-age model was run with the scale- and otolith-age input data and the results were compared. We found that aging biases from scale ages resulted in underestimates of population abundance (15%) and female spawning stock biomass (19%) and overestimates of fishing mortality in the terminal year (19%) and made strong age-1 recruitment years appear weaker and weak ones stronger. Our study demonstrates how aging bias in a large scale-age sample can be corrected with a relatively small sample of paired scale–otolith ages and provides fisheries management with a quantitative evaluation of aging bias and its effects on the assessment of the Atlantic Striped Bass stock.

The Striped Bass *Morone saxatilis* is one of the most important species targeted by recreational and commercial fisheries on the Atlantic coast of the United States. The population was depleted during the early 1980s but was fully restored by 1995 (Richards and Rago 1999). Currently, the Atlantic States Marine Fisheries Commission (ASMFC) conducts Striped Bass stock assessments using a statistical catch-at-age (SCA) model (ASMFC 2011). Although the SCA model is an age-structured forward-projection model that is capable of incorporating errors in catch-at-age data (CAA) into one of its objective functions,

---

*Corresponding author: hliao@odu.edu

Received December 22, 2011; accepted June 13, 2012

it could still provide biased estimates of population parameters due to biased estimates of ages (hereafter referred as aging bias) in the CAA. Therefore, quantification of the effects of aging bias on stock assessment becomes crucial to understanding and interpreting the population parameters estimated by the SCA model, along with improving stock assessment and providing more effective fisheries management advice.

It is difficult to quantify the effects of aging bias on population parameter estimates derived from age-structured models because true-age data with a reasonable age range and a sufficiently long time series are often lacking for species of interest. Therefore, simulation studies have become a popular alternative in examining the potential effects of aging bias on stock assessment. For example, simulation studies have reported that aging biases account for biased estimates of population parameters derived from age-structured stock assessment models (Lai and Gunderson 1987; Rivard 1989; Bradford 1991; Reeves 2003; de Pontual et al. 2006; Bertignac and Pontual 2007). However, because simulation studies assume aging biases, they lack the ability to specifically examine and quantify the effects of aging biases embedded in real data from an ongoing stock assessment.

A scale-based CAA has been used in the Striped Bass stock assessment (ASMFC 2011). However, previous studies found that scales provide less accurate and precise age estimates than otoliths for many fish species, including Striped Bass (Erickson 1983; Libby 1985; Boxrucker 1986; Barber and McFarlane 1987; Heidinger and Clodfelter 1987; Welch et al. 1993; Lowerre-Barbieri et al. 1994; Secor et al. 1995a; Brown et al. 2004; Brouder 2005; Decicco and Brown 2006). Therefore, otoliths should be a more reliable structure for aging and stock assessment. Due to their low cost, minimal impact on fish, and the requirement for a large sample, scales are used for aging by most state and federal agencies in the stock assessment of some species (such as Striped Bass and Summer Flounder *Paralichthys dentatus*), while available otolith-based age samples are insufficient to detect temporal and spatial variations in the age distributions. One possible solution to this problem would be to examine and quantify aging biases induced by scales and their effects on stock assessment with a smaller data set of paired otolith and scale ages.

The stock assessment of Atlantic coast Striped Bass provides a good example of how to empirically quantify the effects of aging bias on stock assessment for three reasons: (1) the sample size and age range of known-age fish have increased since 1995 (Secor et al. 1995b); (2) we have collected large samples of paired scale and otoliths ages for the past 12 years; and (3) the Atlantic coastal Striped Bass stock has varied dramatically over the past 29 years (Richards and Rago 1999).

In this study, we examined aging biases in scale and otolith ages using known-age Striped Bass and then compared the population parameters estimated using both scale and corrected-scale input data with those estimated using otolith-based input data in the SCA model. Our objectives were to (1) validate the scale and otolith aging of Atlantic Striped Bass using known-age fish to see whether the otolith ages can be treated as approximations of the true ages; (2) quantify the level of bias in Atlantic Striped Bass catch-at-age data introduced by scale aging; and (3) quantify the level of bias in estimates of Atlantic Striped Bass population abundance, female spawning stock biomass, fishing mortality for ages 8–11, age-1 recruitment, and retrospective patterns from the scale-based SCA model.

METHODS

*Data collection.—*Striped Bass were collected by the Virginia Marine Resources Commission (VMRC) from the commercial fishery in Virginia waters of the Chesapeake Bay and the Atlantic coastal area from 1999 to 2010. Fish were captured mainly with pound nets (23%) and gill nets (58%) from January to December each year. Fish were measured to the nearest millimeter of total length and weighed to the nearest 0.5 g, and both scales and otoliths were collected from each fish. We define this collection as the Virginia data. We also obtained scales and otoliths of known-age Striped Bass from the Maryland Department of Natural Resources (MD DNR). These fish have been coded-wire-tagged and released from the MD DNR's Joseph Manning Hatchery since 1985 and were recaptured through commercial gear monitoring and fishery-independent surveys in Maryland waters of the Chesapeake Bay and Patuxent River between 1997 and 2006.

*Preparation of hard parts.—*Scales were impressed on acetate slides with a hydraulic heated press for 7 min with 1.03×10^8 Pa at 77°C. Otoliths were thin-sectioned (0.5 mm) with a Buehler Isomet low-speed saw. The otolith section was baked in a Thermolyne 1400 furnace at 400°C for 2.5 min to a light caramel color and then mounted on a microscope glass slide. Lerner Laboratories flo-texx was used as the mounting medium to increase the clarity of the otolith section.

*Aging.—*The acetate impressions of the scales were viewed with a standard Bell and Howell R-735 microfiche reader equipped with 20-mm and 29-mm lenses. Sectioned otoliths were aged with a Leica MZ-12 dissecting microscope with transmitted light and dark-field polarization at between 8 and 100 × magnification. Our aging procedure consisted of two rounds. In the first round, two readers aged all samples separately in chronological order based on collection date without any knowledge of specimen lengths. In the second round, both readers sat together to re-age the fish for which there were different ages in the first round. Consensus ages between two readers were assigned to the fish. When a consensus could not be reached due to the quality of the scales or otoliths, the fish was excluded from further analysis. The percentage agreement between the two readers was calculated by dividing the number of fish for which there was agreement in the first round by the total number of fish aged in both rounds. This agreement was used as an indicator of the difficulty of aging a specific calcified structure (scale or otolith).

Conversion of scale-based SCA model input data.—The ASMFC used an adaptive framework of virtual population analysis model for the coastwide stock assessment of the migratory stock of Atlantic Striped Bass between 1996 and 2007 and switched to the SCA model in 2008 (ASMFC 2011). The CAAs and most abundance indices were collected between 1982 and 2010 from the area over which the stock is distributed, which ranges from Maine to North Carolina. Some young-of-the-year (age-0) and age-1 abundance indices were collected as early as 1970. All of the age-specific input data were derived from scale ages and therefore are referred to as scale-based data.

We developed a year-specific conversion matrix from the Virginia data. A conversion matrix contains the frequencies of otolith ages given a scale age collected during a calendar year. Each element in a conversion matrix is calculated as

$$p_{ji} = \frac{A_{ij}}{T}, \tag{1}$$

where p_{ij} is the proportion of fish with otolith age i and scale age j, A_{ij} is the number of fish in the data set with otolith age i and scale age j, and T is the total number of fish aged for the calendar year.

We used this conversion matrix to convert the scale-based CAAs collected from the coastwide assessment to the otolith-based CAAs. To do so, we assumed that there was no spatial variation in the relationship between scale and otolith ages among states. This assumption was based on the results from two Striped Bass workshops held by the ASMFC in 2003 and 2008. The 2003 workshop found no evidence that any state aged Striped Bass scales significantly more accurately than the others (ASMFC 2003). The 2008 workshop reported a similar relationship between scale and otolith ages in Striped Bass collected in Massachusetts, New Jersey, and Maryland in 2008 (Hoover and Sharov 2009).

The specific conversion is as follows:

$$\mathrm{CAA}_{ij} = C \times p_{ij} \tag{2}$$

$$\mathrm{CAA}_{i} = \sum_{j=\mathrm{min}}^{\mathrm{max}} \mathrm{CAA}_{ij}, \tag{3}$$

where C is the total catch in the year, CAA_{ij} is the catch with scale age j and converted otolith age i, and CAA_i is the catch with converted otolith age i.

Unlike the adaptive framework of virtual population analysis model, the SCA models an annual total index of relative abundance (I) and scale-based age composition separately. Therefore, we first estimated an index (I_{ij}) with scale age j and converted it to otolith age i by replacing C with I in equation (2). Second, we estimated an index (I_i) with otolith age i by replacing CAA_{ij} with I_{ij} in equation (3). Finally, the otolith-based age composition was calculated by dividing I_i by I. The indices of abundance and otolith-based age composition were modeled separately.

Because of cohort progression and recruitment fluctuation over time, pooling all paired scale–otolith age data over time will smooth the magnitude of different cohorts, making strong cohorts appear weaker and vice versa. Therefore, we did not pool Virginia scale–otolith age data to make a year-aggregated matrix for converting the scale-based CAAs to otolith-based ones for the period before 1999. The conversion matrices could not be applied to the age-aggregated, age-0, and age-1 indices. Because there were no age-1 fish collected in Virginia waters, the number of age-1 fish in the scale-based CAAs could not be converted to otolith-based age-1 fish. As a result, the age-aggregated, age-0, and age-1 indices and age-1 fish in the CAAs were kept as scale based. Besides the age-0 and age-1 indices, other indices are from the Marine Recreational Fisheries Statistics Survey (MRFSS), Northeast Fisheries Science Center (NEFSC), Connecticut CPUE (CT CPUE), Connecticut trawl (CT trawl), Delaware spawning stock number (DE SSN), Maryland spawning stock number (MD SSN), New Jersey trawl (NJ trawl), and New York ocean haul seine (NY OHS).

The otolith-based weight at age (WAA) between 1999 and 2010 were directly borrowed from weight at otolith age in the Virginia data. Because Virginia otolith-based WAAs before 1999 and age-1 fish for all years were not available, the WAAs for those fish were kept as scale based.

SCA modeling.—We conducted four SCA modeling runs by combining different time series and conversion completions of the CAAs, WAAs, and age composition indices: (1) the long base run, in which the CAAs, WAAs, and age composition indices were scale based from 1982 to 2010 (this is the ASMFC stock assessment setting); (2) the long corrected run, in which the data were scale based from 1982 to 1998 and otolith based from 1999 to 2010; (3) the short base run, in which the data were scale based from 1999 to 2010; and (4) the short corrected run, in which the data were otolith based from 1999 to 2010. Because the forward-projection procedure in the SCA model requires recruitment indices for 12 years prior to the first year for the CAAs (1982 and 1999 in the long and short runs, respectively), Maryland age-0 and age-1 indices from 1970 to 2010 and from 1987 to 2010 were used in the long and short runs, respectively. The parameters compared among the four runs were the population abundance, female spawning stock biomass, fishing mortality, and recruitment. The effective sample size (ESS) was used for data-weighting in the SCA model. We used an iterative weighting approach in which the ESS was updated until the input and output ESS values converged, assuming a different uncertainty in age composition for each run.

Diagnostics of model fit.—We examined the model fit for the indices and total catch within each run by means of root mean square errors (RMSEs) and that for the CAAs by means of annual residual sums of squares (RSSs) and their total across all years.

Retrospective analysis.—To compare the stability of the SCA model's performance in estimating population parameters among the four runs, we first conducted retrospective analyses of the population abundance and fishing mortality from the four runs from 2010 back to 2005 with the built-in retrospective analysis function in the SCA model. Then we compared the retrospective patterns among the four runs using the Mohn's ρ statistic (Mohn 1999), namely,

$$\rho = \sum_{y=1}^{npeels} \frac{R_{2010-y} - S_{2010-y}}{S_{2010-y}}, \qquad (4)$$

where R and S are estimates of the population parameter of interest for a modeling year from reduced and full time series data, respectively. The term *npeels* is the number of years backward from 2010 (five in this study). When the ρ statistic equals 0, there is no retrospective pattern. The larger the ρ statistic deviates from 0, the stronger the retrospective pattern is. A negative value of ρ indicates that the parameter is underestimated in the terminal year, a positive value that it could be overestimated.

Data analysis.—Three statistical analyses were used to examine the accuracy and precision of scale and otolith ages and their relationship. First, we used ordinary-least-squares regression analysis to compare the scale and otolith ages with the true ages obtained from Maryland known-age Striped Bass. Then we compared the mean annual coefficients of variation (CVs) between the scale and otolith ages of Virginia Striped Bass from 1999 to 2010 using a paired t-test. The CV is defined as follows (Campana 2001):

$$\mathrm{CV}_j = 100 \times \frac{\sqrt{\sum_{i=1}^{2}(X_{ij} - X_j)^2}}{X_j}, \qquad (5)$$

where X_{ij} is the ith age determination ($i = 1$ and 2 for reader 1 and 2, respectively) of the jth fish, X_j is the mean age estimate between the age estimates of reader 1 and 2 for the jth fish. Therefore, mean annual CV is:

$$\mathrm{CV} = \frac{1}{N} \sum_{j}^{N} \mathrm{CV}_j, \qquad (6)$$

where N is the total number of fish aged with both scales and otoliths in a particular year. Finally, we examined the relationships between the scale and otolith ages of Virginia Striped Bass by year from 1999 to 2010 using ordinary-least-squares regression. All statistical analyses were conducted with SAS (SAS Institute 1988) statistical packages.

RESULTS

The MD DNR collected 50 scales and 53 otoliths from known-age Striped Bass between 1997 and 2006. The true ages of these fish ranged from 5 to 19 years. The agreement between our otolith age estimates and the true ages was 74%, with an average CV of 1.9%. The very low CV of 1.9% indicates not only low percentage disagreement but also small differences between the true and otolith ages. For example, only 2 of 53 (3.8%) fish ages differ by more than 1 year. However, the agreement between our scale age estimates and the true ages was 22%, with an average CV of 9.8%. The intercept and slope of the regression of the otolith ages on the true ages were not significantly different from zero ($P = 0.9843$) and 1 ($P = 0.9348$), respectively, whereas the intercept and slope of the regression of the scale ages on the true ages were significantly different from zero ($P = 0.0056$) and one ($P < 0.0001$) (Figure 1), indicating a systematic bias in the scale age estimates. The scale ages tended to overestimate the ages of younger fish and to underestimate the ages of older fish.

The VMRC collected 4,469 Striped Bass from Virginia waters of the Chesapeake Bay and Atlantic coastal area from 1999 to 2010, from which paired scales and otoliths were obtained. Of these pairs, 60 (1.3% of the total collection) were removed due to poor scale and/or otolith quality, leaving 4,409 for further analysis (Table 1). The average agreement between the two readers was 55% (average CV = 5.6%) for scale ages, higher than the average agreement of 48% (average CV = 9.3%) among the readers from six Atlantic states participating in the 2003 ASMFC Striped Bass Aging Workshop, but less variable (Table 2). However, the average agreement between the two readers was 81% (average CV = 1.7%) for our otolith ages (Table 1). The precision was significantly higher for the estimates of the otolith ages than for those of the scale ages ($t = 11.18$, df $= 11$, $P < 0.0001$). The intercepts and slopes of the regressions of the scale and otolith ages by year were significantly different from zero (all Ps < 0.0001) and one (all Ps < 0.0001), respectively (Figure 2). The patterns of the differences were similar to those between the scale and true ages of the Maryland known-age fish (intercept > 0 and slope < 1), where the scale ages tended to overestimate the ages of younger fish and to underestimate those of older fish.

Strong and weak year-classes were more identifiable in the otolith-based CAAs than in the scaled-based CAAs (Figure 3). For example, the 1993 year-class was identified as a very strong cohort by the Maryland and Virginia juvenile seine indices (ASMFC 2011). Its progression could be followed in the otolith-based CAAs until 2004, when it was 11 years old. In contrast, the progression of this cohort in the scale-based CAAs was difficult to follow beyond 2000, when it was only 7 years old. The magnitude of different cohorts was smoothed by using scale-based CAAs, whereas several strong and weak cohorts could be clearly followed by using otolith-based CAAs.

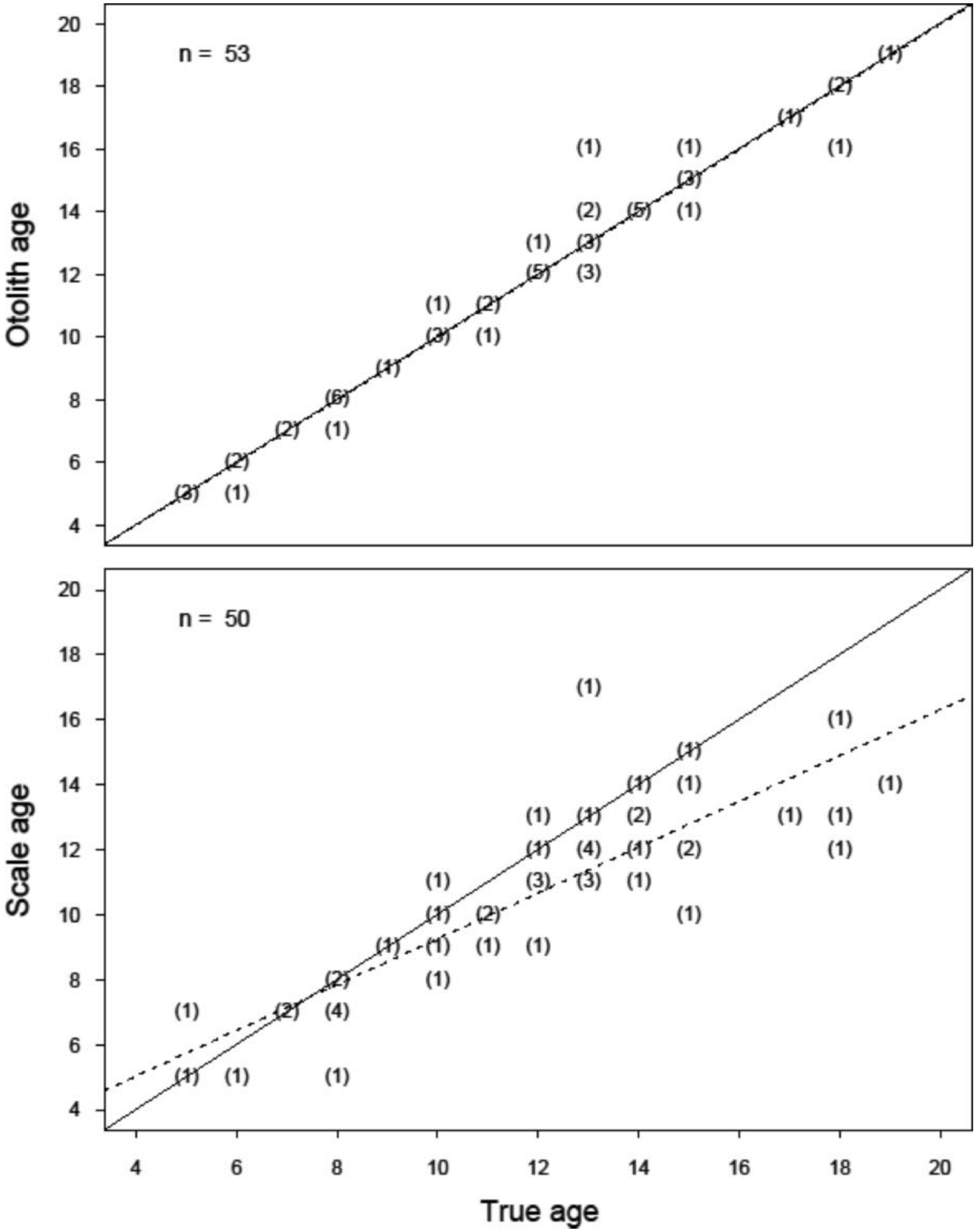

FIGURE 1. Results of regressions of scale and otolith ages on the true ages of known-age Striped Bass released since 1985 and recaptured between 1997 and 2006 by the Maryland DNR. The numbers of fish aged are given in parentheses. The solid lines represent one-to-one correspondences between the scale or otolith ages and the true ages. The dashed lines are the regression curves. The solid and dashed lines overlap in the first panel; *n* = total sample size.

TABLE 1. Agreement between two readers and average annual coefficients of variation (CVs) for scale and otolith ages of Virginia Striped Bass from 1999 to 2010. The number of fish analyzed is the number for which both scale and otolith ages were used to estimate agreement; the number of fish removed is the number whose ages could not be estimated due to damage to their scales and/or otoliths.

	Scale		Otolith		Number of fish	
Year	Agreement (%)	CV (%)	Agreement (%)	CV (%)	Analyzed	Removed
1999	46	7.9	78	2.1	347	5
2000	52	5.7	79	1.7	345	5
2001	63	4.3	78	2.5	241	3
2002	62	4.2	87	1.4	284	5
2003	67	3.8	86	1.5	448	8
2004	63	5.3	83	1.7	571	8
2005	40	8.6	77	2.1	324	8
2006	38	7.5	76	2.0	334	1
2007	60	4.4	78	1.8	532	3
2008	65	3.9	84	1.1	258	11
2009	49	6.6	82	1.3	261	1
2010	53	5.5	84	1.3	464	2
Mean	55	5.6	81	1.7	4,409[a]	60

[a] Total.

In the long base run, the terminal year estimates of population abundance and female spawning stock biomass were underestimated by 15% and 19%, respectively, while fishing mortality was overestimated by 19% (Figure 4). In addition, a maximum underestimation of 25% in the long base run for the population abundance occurred in 1994 when the strongest cohort (the 1993 year-class) appeared as age-1 recruitment for the first time since the population depletion in the early 1980s (Figure 4d). A

TABLE 2. Agreement (%; coefficients of variation in parentheses) between readers and states. The results are from the 2003 ASMFC Striped Bass Aging Workshop, which representatives from five states attended. There were two readers each from states A, B, D, and F and one reader each from states C and E. State D also reported its final readings which were agreements between reader 1 and reader 2 whereas states A, B, and F didn't report their final readings. (By permission of ASMFC.)

State and reader	State A, reader 1	State A, reader 2	State B, reader 1	State B, reader 2	State C, final	State D, reader 1	State D, reader 2	State D, final	State E, final	State F, reader 1	State F, reader 2
State A, reader 1	100 (0.0)										
State A, reader 2	69.9 (5.2)	100 (0.0)									
State B, reader 1	45.6 (8.0)	38.8 (9.3)	100 (0.0)								
State B, reader 2	55.3 (6.6)	54.4 (6.7)	40.8 (10.6)	100 (0.0)							
State C, final	28.2 (12.0)	46.6 (7.4)	28.2 (15.0)	38.8 (11.2)	100 (0.0)						
State D, reader 1	64.1 (4.2)	56.3 (5.4)	60.2 (6.8)	58.3 (7.1)	29.1 (12.8)	100 (0.0)					
State D, reader 2	62.1 (4.3)	56.3 (5.2)	60.2 (7.1)	58.3 (7.4)	35.9 (11.5)	72.8 (4.2)	100 (0.0)				
State D, final	63.1 (3.9)	59.2 (4.3)	60.2 (6.3)	57.3 (7.1)	34.0 (11.3)	80.6 (2.7)	89.3 (1.2)	100 (0.0)			
State E, final	53.4 (5.4)	54.4 (5.5)	52.4 (9.0)	50.5 (8.7)	35.9 (11.3)	71.8 (3.8)	69.9 (4.5)	72.8 (4.1)	100 (0.0)		
State F, reader 1	26.2 (13.1)	35.0 (9.8)	28.2 (16.9)	35 (11.1)	52.4 (7.3)	35.9 (12.0)	35.0 (13.4)	35.9 (12.6)	38.8 (12.3)	100 (0.0)	
State F, reader 2	34.0 (12.8)	25.2 (15.8)	52.4 (10.6)	38.8 (14.2)	15.5 (21.4)	45.6 (11.5)	42.7 (12.7)	43.7 (11.9)	34.0 (14.1)	18.4 (19.5)	100 (0.0)

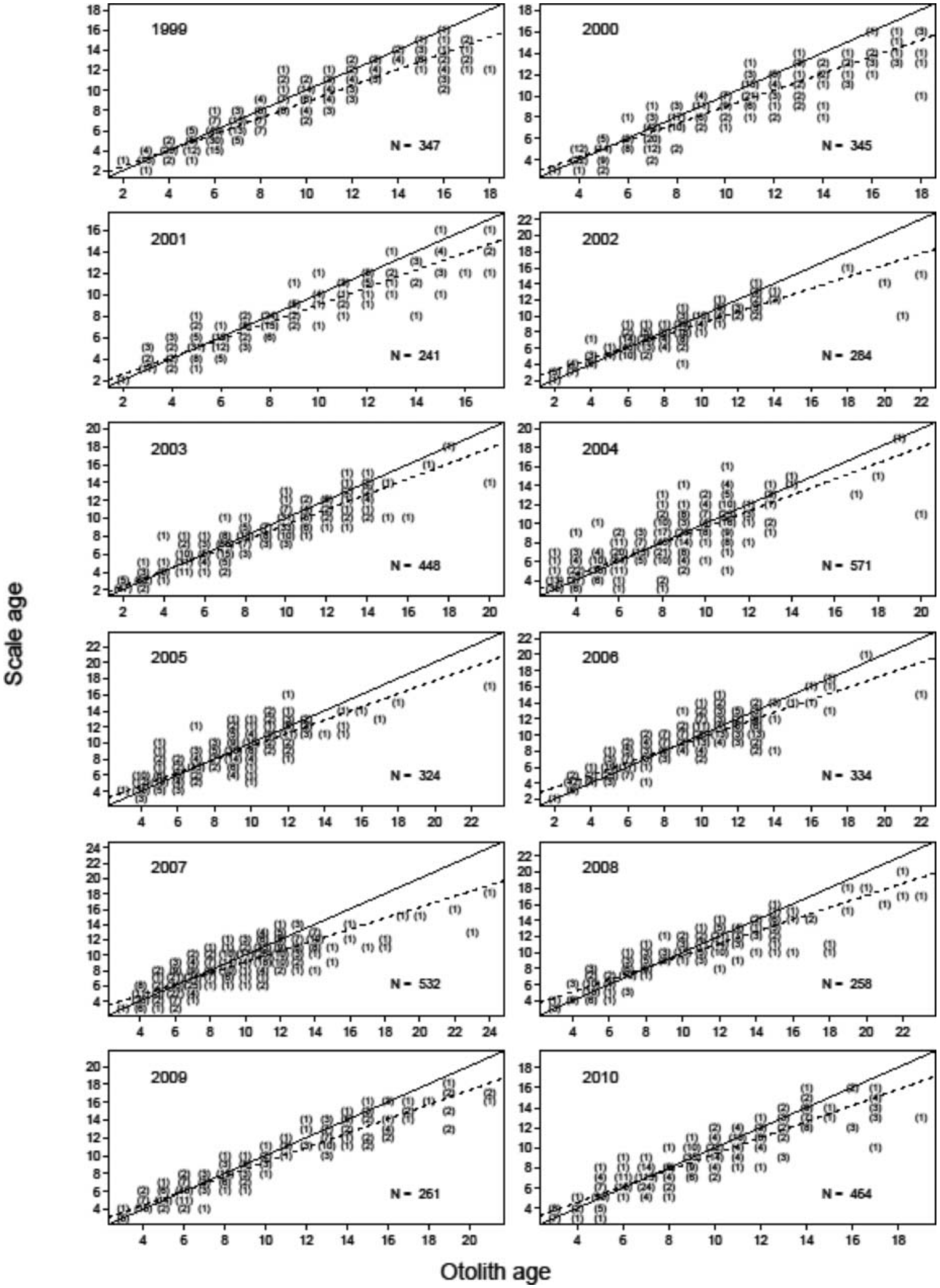

FIGURE 2. Results of regressions of the scale ages of Striped Bass on the otolith ages for fish collected by the VMRC in the Chesapeake Bay and Virginia waters of the Atlantic from 1999 to 2010. See Figure 1 for more information.

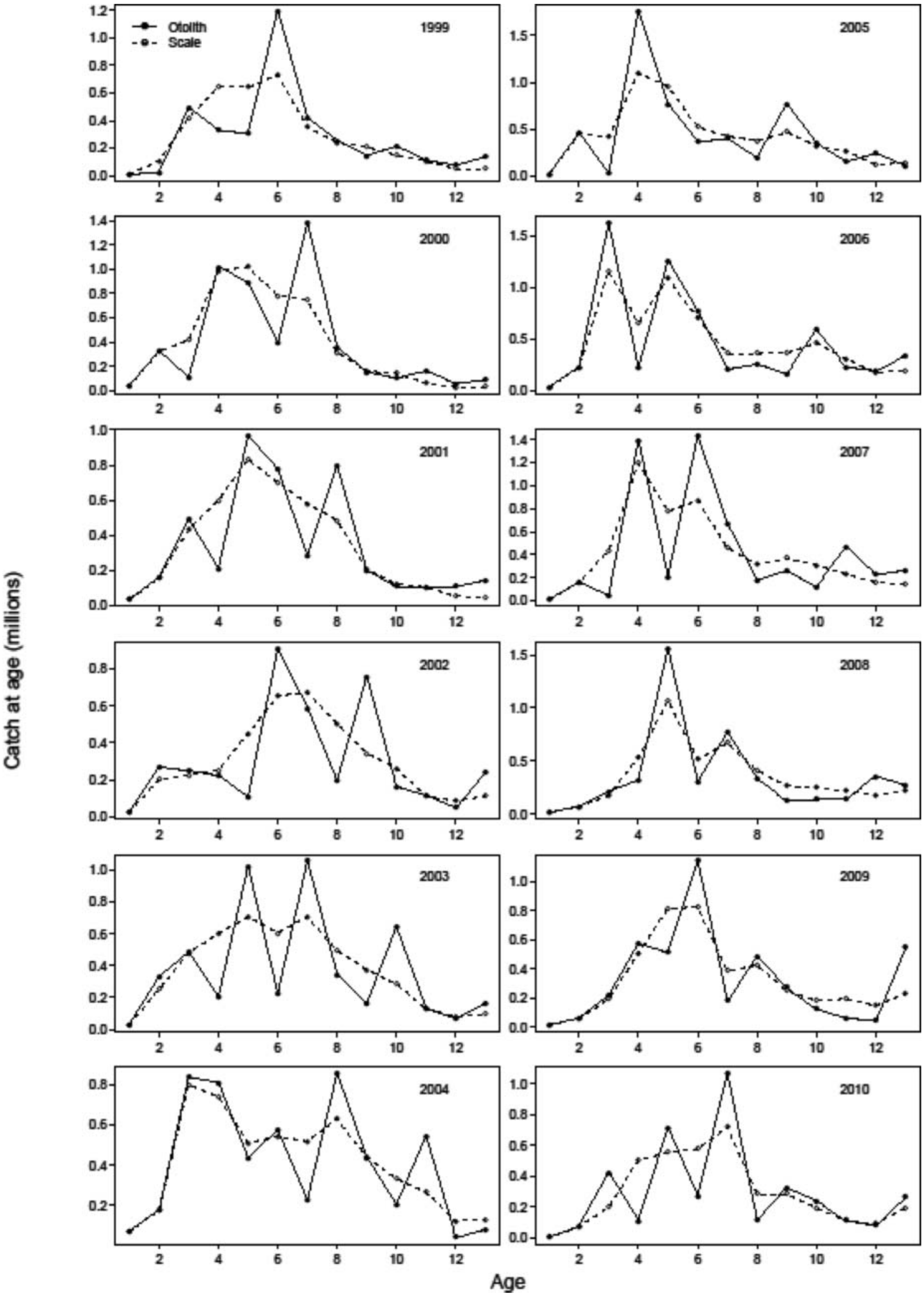

FIGURE 3. Comparisons of the scale-based and otolith-based catches at age (CAAs) of Striped Bass between 1999 and 2010. The otolith-based CAAs were converted from the scale-based CAAs using year-specific conversion matrices.

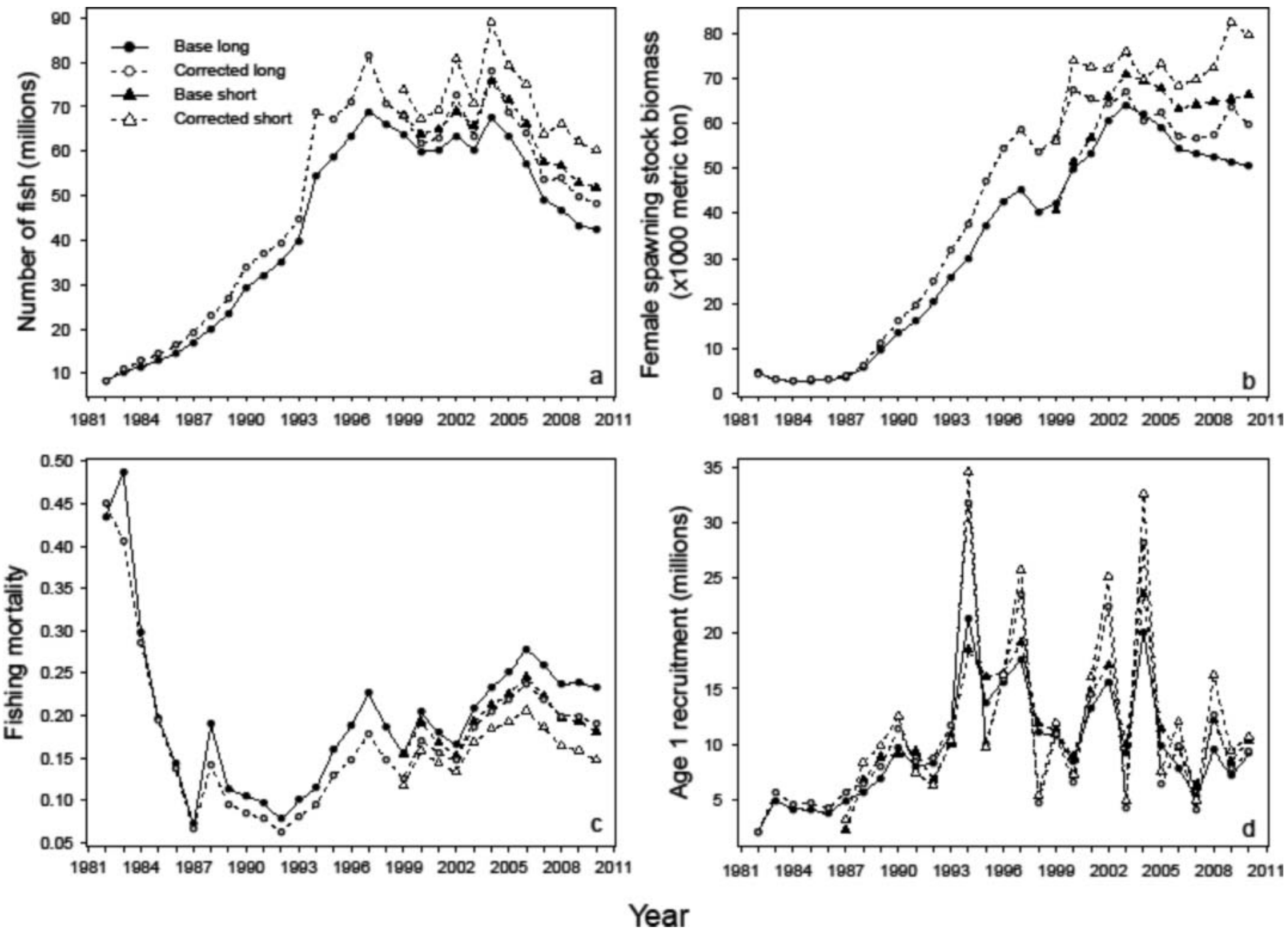

FIGURE 4. Estimates of **(a)** Striped Bass population size, **(b)** female spawning stock biomass, **(c)** fish mortality, and **(d)** age-1 recruitment from the long base, long base–corrected, short base, and short base–corrected runs of the SCA model (see text for details).

maximum underestimation of 33% in the long base run for the female spawning stock biomass occurred in 2000, when about 90% of the females in the 1993 year-class became mature (Figure 4b; ASMFC 2011). The magnitude and direction of difference in the estimates of population abundance, female spawning stock biomass, and fishing mortality between the short base and corrected runs were very similar to those between the long base and corrected runs for the period 1999–2010. However, the differences in those estimates between the long and short runs (both base and corrected) were slightly larger than those between the base and corrected runs (Figure 4a–c), most likely because of the significantly shorter time series in the short runs than in the long runs.

The magnitudes of the annual fluctuation in recruitment were dramatically different between the long base and corrected runs (Figure 4d). First, recruitment was very low with little annual fluctuation and similar magnitude between the long base and corrected runs from 1982 to 1993. Second, recruitment fluctuated more dramatically in the long corrected run than in the long base run between 1994 and 2004, which resulted from the strong year-classes being stronger and the weak year-classes being weaker in the long corrected run. More specifically, the strong year-classes of 1993, 1996, 2001, and 2003 increased 45, 34, 42, and 40%, respectively, whereas the weak year-classes of 1994, 1997, 1999, and 2002 decreased 27, 50, 21, and 50% from the long base to the corrected run. Finally, recruitment in both runs decreased dramatically, with less annual fluctuation, from 2005 to 2010. The magnitudes of the annual fluctuation in recruitment were very similar between the long and short base runs and between the long and short corrected runs (Figure 4d), indicating that the length of the time series had less effect on the estimates of recruitment than on those of population abundance, female spawning stock biomass, and fishing mortality.

The total RSSs across all years for the CAAs in the long and short corrected runs were smaller than in the long and short base runs (Table 3). More specifically, the annual RSSs in 20 of 29 years were smaller in the long corrected run than in the long base run whereas in 7 of 12 years they were smaller in the short

TABLE 3. Sums of squared residuals of CAAs for 1982–2010. The terms "long" and "short" refer to the models from 1982 to 2010 and 1999 to 2010, respectively. The terms "base" and "corrected" are defined in the text; "difference" = the value from the base model less the value from the corrected model.

	Long			Short		
Year	Base	Corrected	Difference	Base	Corrected	Difference
1982	13.1	10.6	2.5			
1983	27.3	18.3	9			
1984	41.5	24.6	16.9			
1985	38.4	25.6	12.8			
1986	16.4	9.9	6.5			
1987	50.7	34.4	16.3			
1988	24.8	16.1	8.7			
1989	39	28.2	10.8			
1990	16.8	11.7	5.1			
1991	17	11.7	5.3			
1992	4.7	3.2	1.5			
1993	10.9	6.7	4.2			
1994	21.7	16.3	5.4			
1995	19.7	7.7	12			
1996	24.9	27.6	−2.7			
1997	27	44.5	−17.5			
1998	3.8	22.2	−18.4			
1999	10.9	16.1	−5.2	7.3	14.4	−7.1
2000	28.6	18.7	9.9	27.3	13.1	14.2
2001	12.9	11.7	1.2	11.5	10.1	1.4
2002	35.4	20.3	15.1	33.5	15.4	18.1
2003	9	25.4	−16.4	11.6	20.4	−8.8
2004	7.8	16.4	−8.6	6.4	11.9	−5.5
2005	9.6	21.1	−11.5	8.5	18.7	−10.2
2006	12.8	17.6	−4.8	11	14.2	−3.2
2007	19.5	21.7	−2.2	24.4	16.6	7.8
2008	25.6	22.7	2.9	29.1	16.7	12.4
2009	30.4	21.8	8.6	38.1	17.2	20.9
2010	27.5	13.1	14.4	39.3	10.8	28.5
Total	627.7	546.3	81.4	248	179.4	68.6

corrected run than in the short base run (Table 3). The RMSEs for the total catch were smaller in the long and short corrected runs than in the long and short base runs (Table 4). The RMSEs for 7 of 14 indices were smaller in the long corrected run than in the long base run, and the same change was found between the short corrected and base runs (Table 4). However, there were decreases in the RMSEs all four age-specific indices (NY OHS, NJ trawl, MD SSN, and DE SSN) in the long corrected run and in 3 of the 4 age-specific indices (NJ trawl, MD SSN, and DE SSN) in the short corrected run. Interestingly, the uncertainties in the age composition were larger in the corrected runs (ESS = 172 in the short run and 212 in the long run) than in the base runs (ESS = 335 in the long run and 393 in the short run).

There were very similar retrospective patterns in the estimates of population abundance and fishing mortality between the long base and corrected runs (Figure 5). However, the retrospective pattern was stronger in the fishing mortality estimates (Mohn's $\rho = 1.252$ in the long base run and 1.129 in the long corrected run) than in the population abundance estimates (Mohn's $\rho = –0.341$ in the long base run and –0.344 in the long corrected run). Population abundance in the terminal year tended to be underestimated (negative Mohn's ρ), whereas fishing mortality tended to be overestimated (positive Mohn's ρ) in both the long base and corrected runs. However, the retrospective patterns in the estimates of population abundance and fishing mortality were more variable in the short base and corrected runs (Figure 6). The retrospective pattern was the weakest for population abundance (–0.02) in the short base run, followed by fishing mortality (0.178) in the short corrected run, population abundance (0.246) in the short corrected run, and fishing mortality (0.465) in the short base run.

TABLE 4. Root mean square errors for indices and total catch. Abbreviations are as follows: NY = New York, NJ = New Jersey, MD = Maryland, VA = Virginia, CT = Connecticut, DE = Delaware; MRFSS = Marine Recreational Fisheries Statistics Survey, NEFSC = Northeast Fisheries Science Center, OHS = ocean haul seine, and SSN = spawning stock number. See Table 3 for more information.

	Long			Short		
Component	Base	Corrected	Difference	Base	Corrected	Difference
Total catch	1.04	0.85	0.20	0.41	0.07	0.34
NY, age 0	7.96	8.56	−0.60	6.80	7.20	−0.40
NJ, age 0	2.89	3.54	−0.65	3.04	3.65	−0.61
MD, age 0	4.01	4.08	−0.07	3.25	3.00	0.25
VA, age 0	4.10	4.51	−0.40	3.87	4.20	−0.32
NY, age 1	2.50	3.11	−0.61	2.37	3.19	−0.82
MD, age 1	2.47	2.24	0.23	2.13	1.65	0.47
MRFSS	1.56	1.51	0.05	2.19	2.14	0.06
CT CPUE	2.15	2.18	−0.03	2.34	2.36	−0.03
NEFSC	2.37	2.43	−0.06	2.40	2.44	−0.04
CT trawl	2.78	3.17	−0.38	2.93	3.42	−0.49
NY OHS	6.29	6.16	0.14	2.19	2.21	−0.01
NJ trawl	2.32	2.27	0.05	2.83	2.81	0.02
MD SSN	3.03	3.00	0.03	1.30	1.24	0.06
DE SSN	2.70	2.49	0.21	1.55	1.54	0.01

DISCUSSION

Following Secor et al.'s (1995b) validation work on scale and otolith aging for Striped Bass, our study enhanced validation by taking advantage of the larger sample size and the broader age range of known-age Striped Bass collected for more than a decade. Our results are consistent with those of Secor et al. (1995); otoliths provided more accurate and precise age estimates for Striped Bass, whereas scales overestimated the ages of younger fish and underestimated the ages of older fish. The overestimation for younger fish is probably due to the misinterpretation of false annuli on the scales of those fish and the underestimation for older fish to the difficulty of interpreting narrow annuli on the peripheral fields of older scales (Barnes and Power 1984; Lowerre-Barbieri et al. 1994; Secor et al. 1995b).

The low agreements and CVs between readers found in this study and the 2003 ASMFC aging workshop indicate that it is very difficult to obtain accurate estimates of Striped Bass ages with their scales and emphasize the necessity and importance of correction for the aging bias inherent in scale age data. Most importantly, our validation with the true ages provides evidence that the otolith ages of Striped Bass have much smaller errors than the scale ages. Therefore, our validation of otolith aging allows us to quantify the effects of aging bias in the scale ages on the SCA stock assessment of Atlantic Striped Bass.

The ability of CAA to track the temporal progression of year-classes plays a crucial role in determining the reliability of population parameter estimates in stock assessment (Megrey 1989). Our study found that the otolith ages of Striped Bass are incomparably superior to the scale ages in tracking year-class progression. Aging bias erroneously assigns fish to adjoining year-classes and results in the strong cohort appearing weaker and the weak cohort stronger (Bradford 1991). As a result, a substantial amount of information on cohort progression could be lost in the scale-based CAAs and the SCA model would provide biased estimates of the population parameters for Striped Bass. In contrast, otolith-based CAAs permit the SCA to model cohort progress more accurately. Likewise, the SCA is better able to predict future declines in a timely fashion should these occur.

A simulation study is an effective tool with which to explore the relationships between aging bias and the estimation of population parameters in general. Heery and Berkson (2009) found that when length frequency was biased toward small fish fishing mortality would be overestimated and thus fisheries policy would be more restrictive than necessary. Yule et al. (2008) used otolith ages as the reference and simulated scale ages to quantify the effects of aging bias in scale ages on the stock assessment of Ciscoes *Coregonus artedi*. They concluded that underestimation of Cisco ages could be responsible for the collapse of these fisheries in Lake Superior. Reeves (2003) reported that aging bias could significantly influence catch forecasts and management advice, such as total allowable catch and effective fishing mortality. Bradford (1991) warned that aging bias not only caused biased estimates of recruitment but also masked environmental effects on recruitment. Using Monte Carlo simulation, Lai and Gunderson (1987) reported that aging bias could result in biased estimates of length at age and further change estimates of the age composition in a Walleye Pollock *Theragra chalcogramma* population, accounting for biased estimates of its population parameters. However, no simulation studies were capable of quantifying the actual aging bias embedded in a particular set

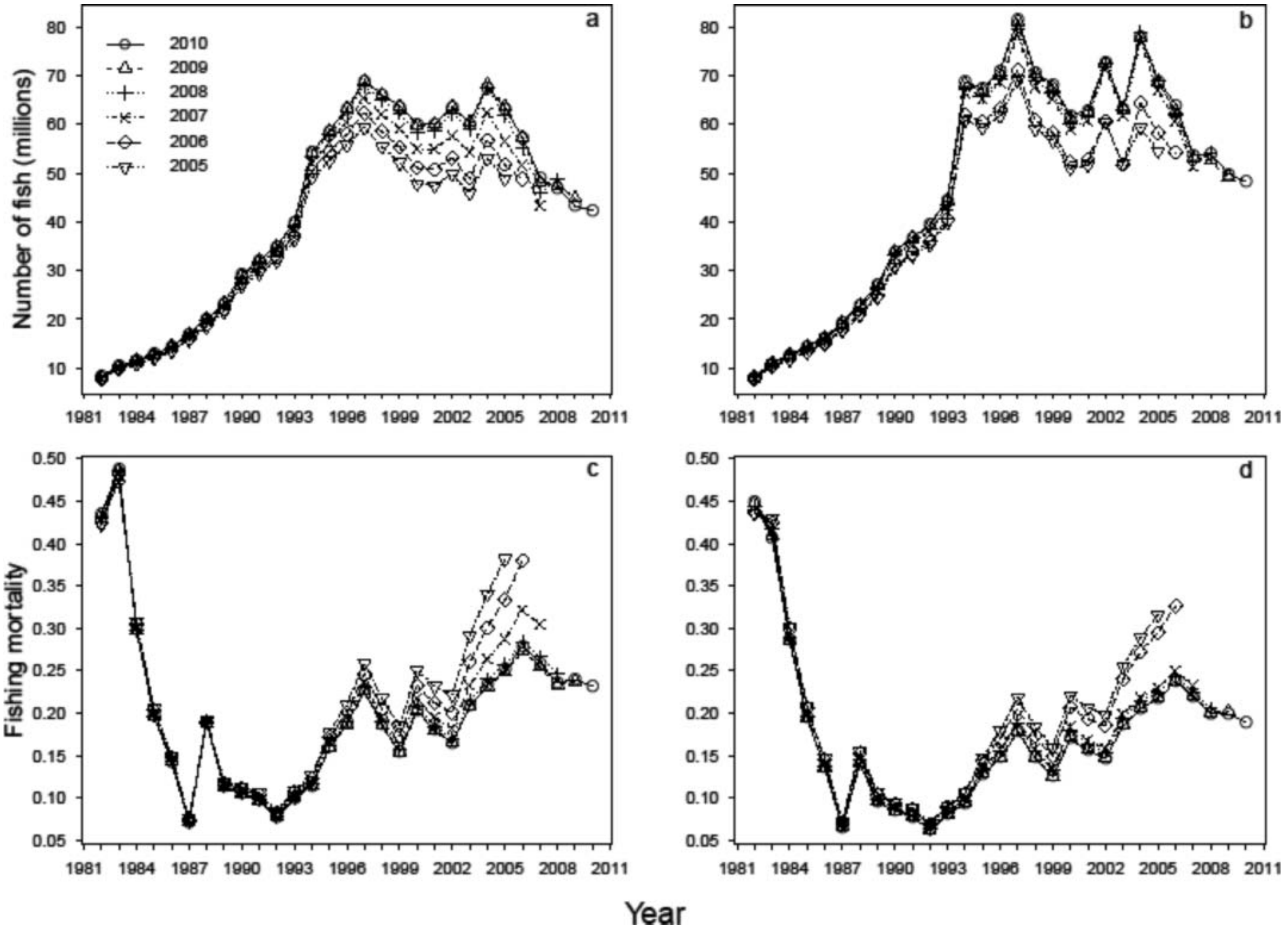

FIGURE 5. Retrospective analyses of the estimates of population size from **(a)** the long base and **(b)** long base–corrected runs of the SCA model and those of fishing mortality from **(c)** the long base and **(d)** long base–corrected runs. The data for 2010 are from the full time series and those for the years before 2010 from reduced time series. The years on the *x*-axis are those for which *N* was estimated from a time series.

of age data used for stock assessment. As a result, fisheries scientists still face challenges in assessing the effects of aging bias on estimates of population parameters derived from an ongoing stock assessment. Our study specifically targeted those challenges using paired otolith and scale ages from the same specimens with reasonable sample sizes across multiple years.

The comparisons between the long base and corrected runs and between the short base and corrected runs quantified the effects of aging bias in terms of underestimation of population abundance and spawning stock biomass, overestimation of fishing mortality, and fluctuation of recruitment over time. Such quantification is especially important when the estimate of a parameter in the terminal year plays a crucial role in fisheries management. The 2011 stock assessment with the long base run concluded that Striped Bass were not overfished and that overfishing was not occurring (ASMFC 2011). At this time, the estimates of female spawning biomass and fishing mortality for the terminal year from both the long base and corrected runs are above and below their thresholds, respectively. Therefore, no immediate management actions are required. However, it would be challenging to the management of Striped Bass if the thresholds had fallen between the estimates of a population parameter from the two runs. This once again emphasizes the importance of quantifying aging bias in an age-structured stock assessment.

There was minimal difference between the estimates of recruitment from the long and short runs. Therefore, we will focus our discussion only on the differences between the base and corrected runs. The low recruitment before 1994 and after 2003 may be responsible for the smaller differences in estimated recruitment between the base and corrected runs, indicating that the SCA may not be able to detect the effects of aging bias on recruitment estimates when recruitment is low and more constant over time. During the period 1994–2003, when recruitment was high and variable, both runs were able to identify strong and weak recruitment well. However, the base run tended to make strong recruitment appear weaker and weak recruitment appear stronger

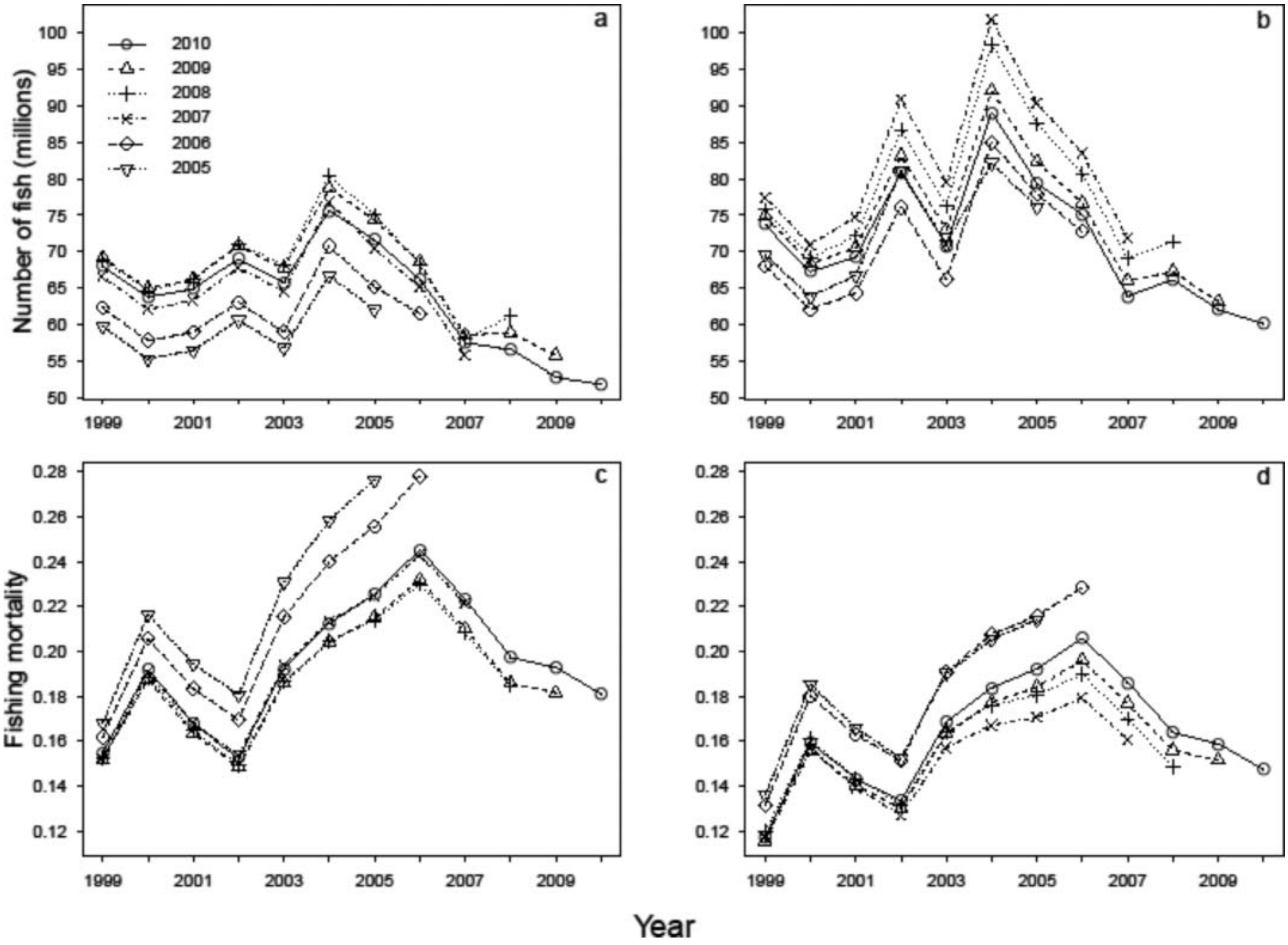

FIGURE 6. Retrospective analyses of the estimates of population size from **(a)** the short base and **(b)** short base–corrected runs of the SCA model and those of fishing mortality from **(c)** the short base and **(d)** short base–corrected runs. See Figure 5 for more information.

because of the aging bias. Bradford (1991) reported that aging biases could significantly reduce the variance in recruitment in sequential population analysis. Our result is not only consistent with Bradford's study (1991) but also indicates that age-structured models such as the SCA can be made more sensitive to a sudden change in recruitment (recovery or failure) between years by correcting the aging bias embedded in the input data.

The aging bias induced by scale ages could result in two opposite possibilities in Atlantic Striped Bass stock assessment and management. When the Striped Bass population was recovering from depletion, the underestimation of abundance, female spawning stock biomass, and the first strong recruitment could have resulted in management advice which allowed the population to have more time to recover in 1980s and early 1990s before exploitation occurred. However, once the population was completely restored by the mid-1990s, management advice based on such biased estimates of population parameters would have unnecessarily constrained Striped Bass harvests.

The correction of aging bias influenced the results of the SCA modeling. First, the smaller RMSEs for the total catch and age-specific indices and the smaller RSSs for most of the CAAs in the long and short corrected runs indicated that correction of aging bias could improve the model fit of age-specific modeling components. However, the model fit of those non-age-specific indices was not relevant to the correction of aging bias with an increase of the RMSEs for some indices and a decrease for the others in the long and short corrected runs. Second, we found that the retrospective patterns were more consistent between the long base and corrected runs than between the short base and corrected runs. We believe that the length of time series input data makes a more significant contribution to the retrospective pattern than does aging bias in the SCA modeling. In addition, Cadigan and Farrell (2005) reported that several factors (such as natural mortality) rather than biases in catch-at-age data could cause retrospective patterns in age-structured population analysis. Finally, the age composition was subject to larger uncertainties in the corrected runs than in the base runs. However, such an increase of uncertainty in the corrected runs is not necessarily due to aging error because uncertainty in age composition in age-structured modeling can also be due to measurement error, observation error, process error, and model specification error (Hulson et al. 2012). It is not clear at this point what errors caused the increase in uncertainty in the corrected runs.

Three concerns raised in this study may deserve more attention in future stock assessments. First, a composite 13+ age-group is used in Atlantic Striped Bass stock assessments because scale ages older than 13 are unreliable (ASMFC 2008). Because the purpose of this study was to examine the effects of aging bias instead of the effects of different age-groups, we made a parallel comparison by keeping the same 13+ age-group in all the runs. However, our study indicates that otolith ages are reliable up to age 19. Therefore, future stock assessments may explore a variety of plus age-groups (up to 19+) and identify the one that provides the most reliable estimates of population parameters in the SCA modeling. Second, although the Striped Bass aging workshops suggest that there is no spatial variation in the relationship between scale and otolith ages among states (ASMFC 2003; Hoover and Sharov 2009), we believe that direct comparison of the conversion matrices developed by different states will enhance understanding of how our empirical method works and how stock assessments can be improved. Finally, the short time series of CAA resulted in much higher estimates of population abundance and spawning stock biomass and lower fishing mortality in the short base and corrected runs than in the long base and corrected runs. These differences are not relevant to the correction of aging bias but could have a significant influence on stock assessment and fisheries management. Therefore, it may be necessary for stock assessment scientists to evaluate the effect of the length of the time series on estimates of population parameters in an age-structured model.

Our study is the first to empirically quantify the effects of aging bias on an ongoing stock assessment (Striped Bass fisheries on the Atlantic coast) using real data sets (paired scale and otolith ages from each specimen). When more accurate age estimates are limited temporally, spatially, and financially, aging bias correction can be used to improve age-specific input data for an age-structured model. More reliable management advice and policies will then be possible.

ACKNOWLEDGMENTS

We thank Beth Versack, Eric Zlokovitz, Eric Durell, Harry Hornick and Lisa Warner at the Maryland DNR for their hard work collecting and processing data. Steve Bobko, James Davies, Christina Morgan, Eric Robillard, and Karen Underkoffler at the Center for Quantitative Fisheries Ecology, Old Dominion University, made significant contributions to aging Striped Bass scales and otoliths. Special thanks go to the Striped Bass Stock Assessment Subcommittee of the Atlantic States Marine Fisheries Commission for their valuable comments on an early draft of this paper. We thank the Virginia Marine Resources Commission and the U.S. Fish and Wildlife Service for their financial support for this project.

REFERENCES

ASMFC (Atlantic States Marine Fisheries Commission). 2003. Striped Bass aging workshop proceedings. ASFMC, Washington, D.C.

ASMFC (Atlantic States Marine Fisheries Commission). 2008. A report of the 46th Northeast regional stock assessment workshop, part A. Assessment report. ASMFC, Reference document: 08-03a, Woods Hole, Massachusetts.

ASMFC (Atlantic States Marine Fisheries Commission). 2011. Striped Bass stock assessment update. ASMFC, Washington, D.C.

Barber, W. E., and G. A. McFarlane. 1987. Evaluation of three techniques to age Arctic Char from Alaskan and Canadian waters. Transactions of the American Fisheries Society 116:874–881.

Barnes, M. A., and G. Power. 1984. A comparison of otolith and scale ages for western Labrador Lake Whitefish, *Coregonus clupeaformis*. Environmental Biology of Fishes 10:297–299.

Bertignac, M., and H. D. Pontual. 2007. Consequences of bias in age estimation on assessment of the northern stock of European Hake (*Merluccius merluccius*) and on management advice. ICES Journal of Marine Science 64:981–988.

Boxrucker, J. 1986. A comparison of otolith and scale methods for aging White Crappie in Oklahoma. North American Journal of Fisheries Management 6:122–125.

Bradford, M. J. 1991. Effects of aging errors on recruitment time series estimated from sequential population analysis. Canadian Journal of Fisheries and Aquatic Sciences 48:555–558.

Brouder, M. J. 2005. Age and growth of Roundtail Chub in the upper Verde River, Arizona. Transactions of the American Fisheries Society 134: 866–871.

Brown, P., C. Green, K. P. Sivakumaran, D. Stoessel, and A. Giles. 2004. Validating otolith annuli for annual age determination of Common Carp. Transactions of the American Fisheries Society 133:190–196.

Cadigan, N. G., and P. J. Farrell. 2005. Local influence diagnostics for the retrospective problem in sequential population analysis. ICES Journal of Marine Science 62:256–265.

Campana, S. E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abused of age validation methods. Journal of Fish Biology 59:197–242.

de Pontual, H., A. L. Groison, G. Pineiro, and M. Bertignac. 2006. Evidence of underestimation of European Hake growth in the Bay of Biscay and its relationship with bias in the agreed method of age estimation. ICES Journal of Marine Science 63:1674–1681.

Decicco, A. L., and R. J. Brown. 2006. Direct validation of annual growth increments on sectioned otoliths from adult Arctic Grayling and a comparison of otolith and scale ages. North American Journal of Fisheries Management 26:580–586.

Erickson, C. M. 1983. Age determination of Manitoban Walleyes using otoliths, dorsal spines, and scales. North American Journal of Fisheries Management 3:176–181.

Heery, E. C., and J. Berkson. 2009. Systematic errors in length frequency data and their effect on age-structured stock assessment models and management. Transactions of the American Fisheries Society 138:218–232.

Heidinger, R. C., and K. Clodfelter. 1987. Validity of the otolith for determining age and growth of Walleye, Striped Bass, and Smallmouth Bass in power plant cooling ponds. Pages 241–252 *in* R. C. Summerfelt, editor. Age and growth of fish. Iowa State University Press, Ames.

Hoover, A. K., and A. Sharov. 2009. Comparison of otolith aged to scale aged Striped Bass and implications to catch at age estimates. Report to the Striped Bass Technical Committee of the Atlantic States Marine Fisheries Commission. Atlantic States Marine Fisheries Commission, Washington, D.C.

Hulson, P.-J. F., D. H. Hanselman, and T. J. I. Quinn. 2012. Determining effective sample size in integrated age-structured assessment models. ICES Journal of Marine Science 69:281–292.

Lai, H., and D. R. Gunderson. 1987. Effects of aging errors on estimates of growth, mortality, and yield per recruit for Walleye Pollock (*Theragra chalcogramma*). Fisheries Research 5:287–302.

Libby, D. A. 1985. A comparison of scale and otolith aging methods for the alewife *Alosa pseudoharengus*. U.S. National Marine Fisheries Service Fishery Bulletin 83:696–701.

Lowerre-Barbieri, S. K., J. M. E. Chittenden, and C. M. Jones. 1994. A comparison of a validated otolith method to age Weakfish, *Cynoscion regalis*, with the traditional scale method. U.S. National Marine Fisheries Service Fishery Bulletin 92:555–568.

Megrey, B. A. 1989. Review and comparison of age-structure stock assessment models from theoretical and applied points of view. Pages 8–48 *in* E. F. Edwards and B. A Megrey, editors. Mathematical analysis of fish stock dynamics. American Fisheries Society, Symposium 6, Bethesda, Maryland.

Mohn, R. 1999. The retrospective problem in sequential population analysis: an investigation using cod fishery and simulated data. ICES Journal of Marine Science 56:473–488.

Reeves, S. A. 2003. A simulation study of the implications of age-reading errors for stock assessment and management advice. ICES Journal of Marine Science 60:314–328.

Richards, R. A., and P. J. Rago. 1999. A case history of effective fishery management: Chesapeake Bay Striped Bass. North American Journal of Fisheries Management 19:356–375.

Rivard, D. 1989. Overview of the systematic, structural, and sampling errors in cohort analysis. Pages 49–65 *in* E. F. Edwards and B. A. Megrey, editors. Mathematical analysis of fish stock dynamics. American Fisheries Society, Symposium 6, Bethesda, Maryland.

SAS Institute. 1988. SAS/STAT user's guide, release 6.03. SAS Institute, Cary, North Carolina.

Secor, D. H., J. M. Dean, and S. E. Campana. 1995a. Recent developments in fish otolith research. University of South Carolina Press, Columbia.

Secor, D. H., T. M. Trice, and H. T. Hornick. 1995b. Validation of otolith-based aging and a comparison of otolith- and scale-based aging in mark–recaptured Chesapeake Bay Striped Bass, *Morone saxatilis*. U.S. National Marine Fisheries Service Fishery Bulletin 93:186–190.

Welch, T. J., M. J. V. D. Avyle, R. K. Betsill, and E. M. Driebe. 1993. Precision and relative accuracy of Striped Bass age estimates from otoliths, scales, and anal fish rays and spines. North American Journal of Fisheries Management 13:616–620.

Yule, D. L., J. D. Stockwell, J. A. Black, K. I. Cullis, G. A. Cholwek, and J. T. Myers. 2008. How systematic age underestimation can impede understanding of fish population dynamics: lessons learned from a Lake Superior cisco stock. Transactions of the American Fisheries Society 137:481–495.