Comparison of Blueline Tilefish Otolith Derived Ages: Comparing Increment Counts Derived by Readers from NMFS SEFSC-Beaufort and SCDNR Age Laboratories

Joseph C. Ballenger

SEDAR50-DW13

Submitted: 6 January 2017

Please cite this document as:

Ballenger, J. C. 2017. Comparison of Blueline Tilefish Otolith Derived Ages: Comparing Increment Counts Derived by Readers from NMFS SEFSC-Beaufort and SCDNR Age Laboratories. SEDAR50-DW13. SEDAR, North Charleston, SC. 14 pp.

Comparison of Blueline Tilefish Otolith Derived Ages:

Comparing Increment Counts Derived by Readers from NMFS SEFSC-Beaufort and SCDNR Age Laboratories

Joseph C. Ballenger

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

SEDAR50-WPXX
SCDNR Reef Fish Survey Technical Report # 2016-006

Marine Resources Monitoring, Assessment and Prediction (MARMAP) program and the Southeast Area Monitoring and Assessment Program, South Atlantic (SEAMAP-SA) region. Both programs are sponsored by the National Marine Fisheries Service (Southeast Fisheries Science Center) and the South Carolina Department of Natural Resources.

Introduction

Recently the South Carolina Department of Natural Resources (SCDNR) Reef Fish Survey conducted an age validation study of Blueline Tilefish using bomb radiocarbon estimates derived from adult Blueline Tilefish otoliths cores (SEDAR 50-WPXX). Reef Fish Survey staff analyzed 40 blueline tilefish otolith cores for bomb radiocarbon analysis, selecting fish with estimated birth years ranging from 1952 to 1987 based on SCDNR age reader consensus (i.e., if two readers disagreed, a third consensus read was conducted). After bomb radiocarbon analysis, additional age determination reads of these same fish were performed by 1) the two SCDNR Reef Fish Survey staff responsible for aging Blueline Tilefish, denoted SCDNR Reader 1 and SCDNR Reader 2, respectively, 2) two National Marine Fisheries Service (NMFS) Southeast Fisheries Science Center (SEFSC), Beaufort Laboratory readers; one was the currently assigned age reader for Blueline Tilefish at the SEFSC-Beaufort lab (SEFSC Reader 1) and the other had previous experience aging Blueline Tilefish (SEFSC Reader 2), and 3) the primary Blueline Tilefish age reader from the Center for Quantitative Fish Ecology (CQFE) at Old Dominion University. In addition, historical age determinations made by previous employees of the SCDNR Reef Fish Survey were available for the majority of these fish. Each of these independent age determination reads were used to compare a bomb radiocarbon chronology derived from their estimated age estimates to reference chronologies. The bomb radiocarbon chronologies from the independent Blueline Tilefish age determinations were also compared. Full results of these analyses can be found in SEDAR 50-WPXX.

Preliminary bomb radiocarbon data available prior to the SEDAR 50 Stock ID workshop (held June 28-30, 2016 in Raleigh, NC) suggested a bias in ages between the current SEFSC Reader 1 for Blueline Tilefish and the ages derived from both the SCNDR Reef Fish Survey readers and the SEFSC Reader 2. SEFSC Reader 1 was under-aging relative to all others. Identification of this bias led to a recommendation from both the SCDNR Reef Fish Survey and SEFSC-Beaufort age laboratories to delay the SEDAR 50 Data Workshop until the nature of the bias could be further explored. An understanding of the bias was necessary before age data from both labs could be used in the SEDAR 50 Blueline Tilefish stock assessment. This recommendation led to the limited use of age data to explore the stock structure of Blueline Tilefish at the SEDAR 41 Stock ID Workshop, and to a recommendation for an additional age workshop involving all laboratories that were expected to provide age data to the SEDAR 50 Data Workshop (SCDNR Reef Fish Survey, SEFSC-Beaufort, CQFE, SEFSC-Panama City Laboratory, and the NEFSC-Woods Hole Laboratory). The SEDAR 50 Stock ID Workshop panelists stated the results of the bomb radiocarbon study were “useful to inform age readers on interpretation of the otolith structure” and that “this information will be used during the age workshop” (SEDAR 50 DW12). The SEDAR 50 Stock Id Workshop also concluded that Blueline Tilefish should be managed as a unit stock from throughout the Gulf of Mexico through mid-Atlantic waters; this necessitated the inclusion of two additional labs in the proposed age workshop, SEFSC-Panama City and Northeast Fisheries Science Center (NEFSC), Woods Hole.

Prior to the scheduled age workshop on August 29-31, 2016, SEFSC-Beaufort provided the SCDNR Reef Fish Survey with otoliths from 898 Blueline Tilefish previously aged by SEFSC-Beaufort Reader 1. The intent was to have SCDNR Reef Fish Survey Blueline Tilefish age readers age as many of these prior to the Blueline Tilefish age workshop as possible. This could provide needed information on possible bias in age determinations across the full age range of Blueline Tilefish among the two labs. This can serve as one of the initial discussion points for the age workshop. This working paper presents the results of those comparisons. Due to budget, staffing, and other constraints, the two labels current employ different methodologies to determine a final age; SCDNR has two readers and upon disagreement a consensus read is made while SEFSC-Beaufort has a single reader.

Materials and Methods

Samples

SEFSC-Beaufort Lab personnel provided the SCDNR Reef Fish Survey with otolith sections from 898 Blueline Tilefish collected through fishery-dependent sampling programs along the U.S. South Atlantic coast from 2012 through 2015 (Table 1). Each of these samples had been recently aged (e.g. increment count, edge type, and quality code) by SEFSC-Beaufort Reader 1. These specimens were subsequently aged by the two readers from the SCDNR Reef Fish Survey lab. The two SCDNR readers had recently completed (re-)aging all Blueline Tilefish otoliths available at the SCDNR in preparation for SEDAR 50.

Calculations

Average Percent Error (APE)

The average percent error for fish *i* (APE_i) was calculated as

$$APE_i = 100 * \frac{1}{R} * \sum_{j=1}^{R} \frac{|x_{i,j} - \bar{x}_i|}{\bar{x}_i},$$

where $x_{i,j}$ is the age for fish *i* from reader *j*, $\bar{x}_i$is the average age for fish *I,* and R is the number of times each fish is aged (always equals 2 in this analysis; Beamish and Fournier 1981). From this, the mean average percent error ($\overline{APE}$) observed across *n* individual fish (see Table 2 for n) was calculated as

$$\overline{APE} = \frac{\sum_i APE_i}{n},$$

where n is a comparison specific sample size. The $\overline{APE}$ represents an index of average percent error (Campana 2001).

Coefficient of Variation (CV)

The coefficient of variation for fish *i* (Chang 1982) was calculated as

$$CV_i = 100 * \frac{\sqrt{\sum_{j=1}^{R} \frac{(x_{i,j} - \bar{x}_i)^2}{R-1}}}{\bar{x}_i}$$

where CV_i is the age precision estimate for fish *i.* All other terms are as defined for APE. As with APE, the mean CV ($\overline{CV}$) observed across *n* individual fish was calculated as

$$\overline{CV} = \frac{\sum_i CV_i}{n}.$$

Percent Agreement

For each comparison, the percent agreement to within ±*y* years was calculated, where $y = 0, 1, \ldots, 10$.

Symmetry Test

Here we used Bowker's (1948) method to test the hypothesis that an *m* x *m* contingency table consisting of two classifications of a sample into categories (e.g., ages given by two readers) is symmetric about the main diagonal. This method to assess differences between two sets of age readings was first proposed by Hoenig et al. (1995). The test statistic is distributed as a chi-square variable with $m(m-1)/2$ degrees of freedom for a table that has no empty cells. The test statistic is

$$\mathrm{X}^2 = \sum_{i=1}^{m-1} \sum_{j=i+1}^{m} \frac{\left(n_{i,j} - n_{j,i}\right)^2}{n_{i,j} + n_{j,i}}$$

where $n_{i,j}$ is the observed frequency in the *i*th row and *j*th column and $n_{j,i}$ is the observed frequency in the *j*th row and *i*th column. The summation is over all the cells above the diagonal; these cells are paired with the corresponding cells below the diagonal. If X^2 is large, indicates that there is a systematic difference between ages. Finally, when comparing the observed test statistic to a chi-square distribution, the degrees of freedom is equal to the number of comparisons. See Hoenig et al. (1995) for a simple example.

Plots

Two different types of plots are provided that are useful in qualitatively assessing whether bias is present between two sets of age reads on the same individual fish, a distribution plot and a bias plot. The bias plot was proposed by Campana (2001). No bias is present if there is no significant devation form the expected 1:1 line. See Figures 1 and 2 for a complete description of the plots.

Comparisons

Three sets of comparisons were performed: SCDNR Reef Fish Survey Reader 1 vs SCDNR Reef Fish Survey Reader 2, SEFSC-Beaufort Reader 1 vs SCDNR Reef Fish Survey Reader 1, and SEFSC-Beaufort Reader 1 vs SCDNR Reef Fish Survey Reader 2. For each comparison, the $\overline{APE}$, the $\overline{CV}$, the percent agreement, and the symmetry test were performed. Further, for each comparison I provide reciprocal distribution and bias plots.

Results

SCDNR Reef Fish Survey Reader 1 vs SCDNR Reef Fish Survey Reader 2

Both readers from the SCDNR Reef Fish Survey assigned ages to 646 individual Blueline Tilefish (Table 2). The $\overline{APE}$ between the two readers was 11.80% and, given the complexity of the otolith structure, was considered acceptable. Further, percent agreement was greater than 90% by ±5 years, and exceeding 95% by ±7 years. The symmetry test suggests only marginal lack of symmetry between the readers ages (p = 0.0453).

Both the distribution (Figure 1) and bias (Figure 2) plots indicate that there is little to no bias between the two SCDNR Reef Fish Survey age reads. This is in contrast to comparisons involving SEFSC-Beaufort Reader 1.

SEFSC-Beaufort vs SCDNR Reef Fish Survey Reader 1

Both the SEFSC-Beaufort Reader 1 and SCDNR Reef Fish Survey Reader 1 assigned ages to 653 individual Blueline Tilefish (Table 2). The $\overline{APE}$ between the two readers was 26%, exceeding an acceptable level. In addition, percent agreement was less than 90% at ±8 years.

Both the distribution (Figure 3) and bias (Figure 4) plots indicate that the SEFSC-Beaufort Reader 1 is substantially under-aging Blueline Tilefish relative to the SCDNR Reef Fish Survey Reader 1 across most age classes. This under-aging is particularly evident in the bottom panel of both Figure 3 and Figure 4. The degree of under-aging relative to SCDNR Reef Fish Survey Reader 1 increases with age. The degree of under-aging by SEFSC-Beaufort Reader 1 increases from 1 year to approximately 5 years from age 5 to 11, being 6-7 years younger at ages 12-17, and 7 to 10 years younger at ages 18+.

SEFSC-Beaufort vs SCDNR Reef Fish Survey Reader 2

Both the SEFSC-Beaufort Reader 1 and SCDNR Reef Fish Survey Reader 2 assigned ages to 874 individual Blueline Tilefish (Table 2). The $\overline{APE}$ between the two readers was 26%, exceeded an acceptable level. This was very similar to the comparison of the SEFSC-Beaufort Reader 1 to SCDNR Reef Fish Survey Reader 1. As was the case with the previous comparison, percent agreement was less than 90% at ±8 years.

Both the distribution (Figure 5) and bias (Figure 6) plots indicate that SEFSC-Beaufort Reader 1 substantially under-aged relative to the SCDNR Reef Fish Survey Reader 2 across most age classes. As was the case with the previous comparison, this under-aging is particularly evident in the bottom panel of both Figure 5 and Figure 6. The degree of under-aging relative to SCDNR Reef Fish Survey Reader 2 increases with age. The degree of under aging by SEFSC-Beaufort Reader 1 steadily increases from 1 year to approximately 5 years from age 5 to 13, being 7-8 years younger from ages 14-19, and then generally 10 to 15 years younger at ages 20+.

Discussion

Results suggest that the two SCDNR age readers are interpreting Blueline Tilefish in a consistent manner, as evidenced by the lack of appreciable bias between their reads of 646 individual Blueline Tilefish. A similar result was obtained when the same readers read internal Blueline Tilefish calibration sets both prior to and during production aging of Blueline Tilefish in preparation for SEDAR 50 (data not shown).

The SCDNR Reef Fish Survey lab employs a protocol where every Blueline Tilefish is independently aged by two age readers. If there is disagreement with the independent reads, both readers perform a final consensus read of the individual fish together. Further, production aging of Blueline Tilefish begins only after the reader has successfully aged a standardized calibration set and obtained acceptable $\overline{APE}$ and bias plots indicated minimal bias assuming the calibration age represents the true age of that individual Blueline Tilefish. Given this protocol, even if marginal bias is present, consensus age reads will help to standardize the final age determinations provided to the stock assessment process. Further, the consensus reads allow for an ad-hoc SCDNR internal calibration during production aging; consensus reads allows for the identification of between reader drift in age determination methodology during production aging. While temporal drift (i.e., across months or years) may still occur, re-calibration to the standardized calibration set with previously agreed upon Blueline Tilefish ages minimizes this possibility.

Conversely, comparisons of ages determined by SEFSC-Beaufort Reader 1 and either SCDNR age reader suggests significant bias between the two labs. Such bias makes direct comparisons of age compositions and life history parameter estimates derived from the two laboratories problematic. Further complicating these results is that the primary source of age composition information for the population is derived from the SEFSC-Beaufort fishery-dependent sampling programs while much of the life history and fishery-independent age composition information is derived from the SCDNR Reef Fish Survey data. The source of this bias must be identified and addressed prior to the use of both data sets in the SEDAR 50 stock assessment.

The results of the SCDNR Reef Fish Survey bomb radiocarbon study of Blueline Tilefish (SEDAR 50-WPXX) suggests that ages derived from SCDNR, CQFE and SEFSC-Beaufort Reader 2 are more consistent with the true age of Blueline Tillefish. When compared to previously developed bomb radiocarbon reference chronologies for finfish and corals from across the world, the results suggest that the SCDNR Reef Fish Survey readers may still be slightly under-aging, not over-aging Blueline Tilefish (SEDAR 50-WPXX). When compared to bomb radiocarbon data collected for other U.S. South Atlantic deep water fish species (Filer and Sedberry 2008, Friess and Sedberry 2011, Lytton et al. 2016), the SCDNR Reef Fish Survey readers exhibit no bias (SEDAR 50-WPXX). The additional under-aging, compared to the SCDNR age readers, observed for SEFSC-Beaufort Reader 1 suggest a larger bias for those ages than those derived from SCDNR Reef Fish Laboratory readers.

Tables and Figures

Table 1: Number of Blueline Tilefish aged (n) by each reader. Also provided is the minimum and maximum age recorded for an individual Blueline Tilefish by each reader. Note, as of this report SCDNR Reader 1 and SCDNR Reader 2 had not examined 201 and 2 samples, respectively, for age determination. Additional differences in sample size between the SCDNR readers and the SEFSC-Beaufort reader derives from SCDNR readers identifying some fish as "un-readable" (e.g., quality code of A).

Reader	**n**	**Min. Age**	**Max. Age**
SEFSC-Beaufort	895	1	29
SCDNR Reader 1	653	1	35
SCDNR Reader 2	874	2	44

					Percent Agreement											Symmetry Test		
Reader 1	**Reader 2**	**n**	**APE**	**CV**	**±0**	**±1**	**±2**	**±3**	**±4**	**±5**	**±6**	**±7**	**±8**	**±9**	**±10**	χ^2	**df**	**p**
SEFSC	SCDNR R1	653	26.36	37.28	7.04	19.91	30.63	44.10	55.28	68.15	75.96	81.93	87.29	92.19	94.18	1072.72	298	0.0000
SEFSC	SCDNR R2	874	25.50	36.07	6.41	19.34	35.93	49.43	60.87	69.91	77.12	83.30	87.99	90.39	92.45	1432.67	368	0.0000
SCDNR R1	SCDNR R2	646	11.80	16.69	18.27	45.82	61.61	76.78	84.37	90.25	93.81	95.98	96.75	98.14	98.61	342.65	300	0.0453

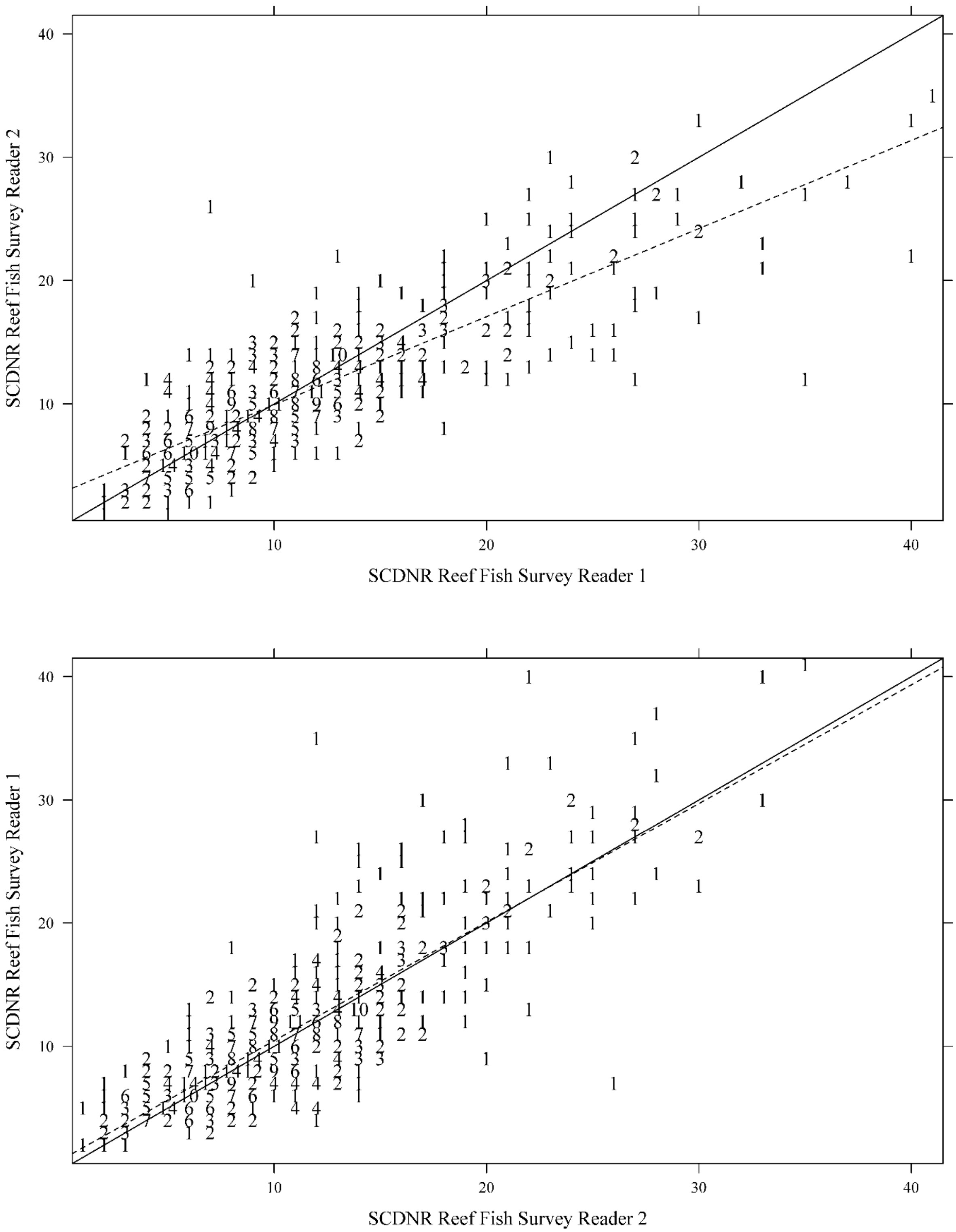


Figure 1: Distribution plot comparing SCDNR Reef Fish Survey Reader 1 and SCDNR Reef Fish Survey Reader 2 ages for the same Blueline Tilefish. Top panel – SCDNR Reef Fish Survey Reader 1 ages are considered the true ages; Bottom panel – SCDNR Reef Fish Survey Reader 2 ages are considered the true ages. Sold line – expected 1:1 line; Dashed line – linear regression model fit to the observed data; numbers – observed frequency of fish with an observed x,y combination.

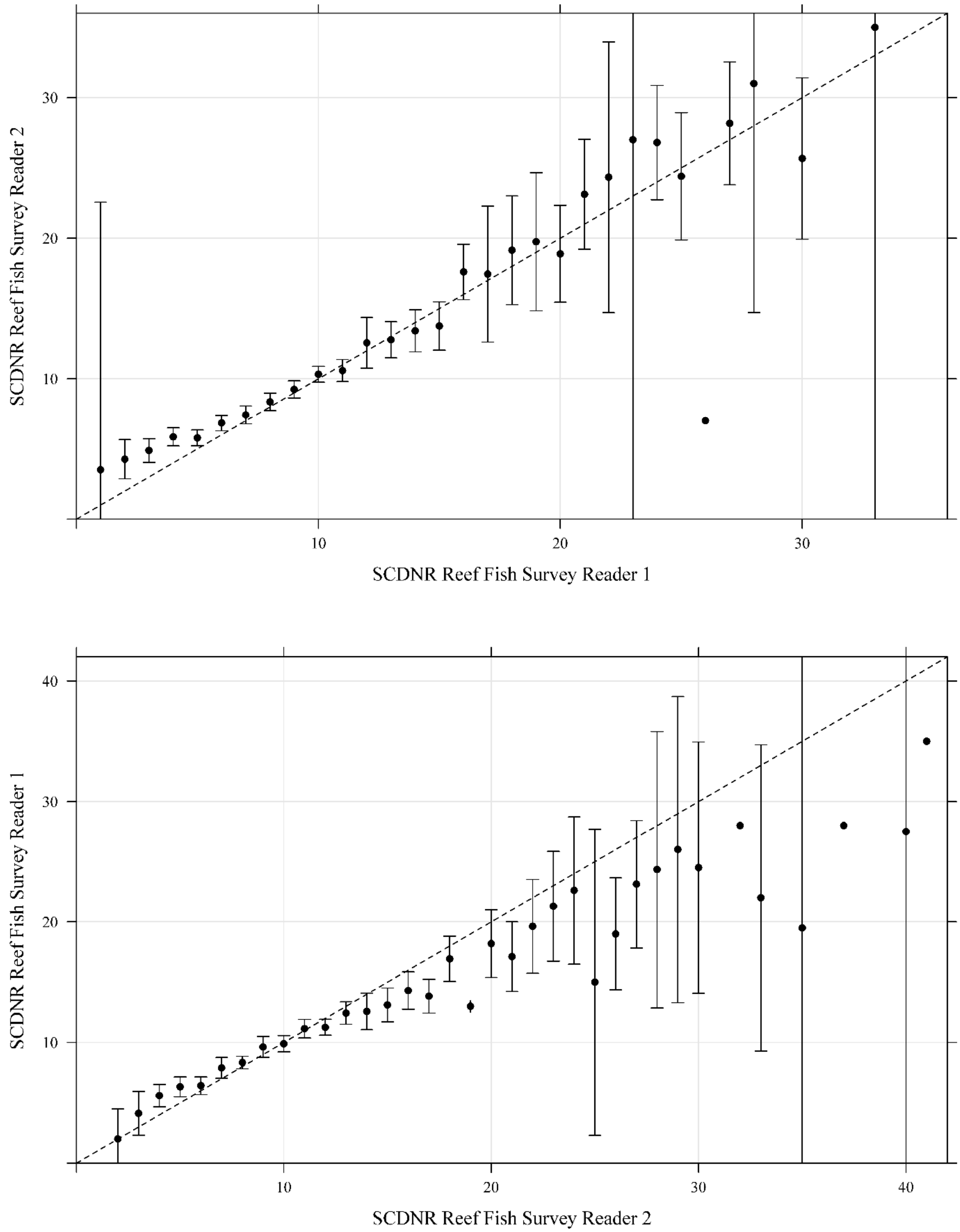


Figure 2: Bias plot comparing SCDNR Reef Fish Survey Reader 1 ages to SCDNR Reef Fish Survey Reader 2 ages for the same Blueline Tilefish. Top panel – SCDNR Reef Fish Survey Reader 1 ages are considered the true ages; Bottom panel – SCDNR Reef Fish Survey Reader 2 ages are considered the true ages. Dashed line – expected 1:1 line; Error bars represent the 95% confidence interval of y-axis reader ages at a given x-axis reader age.

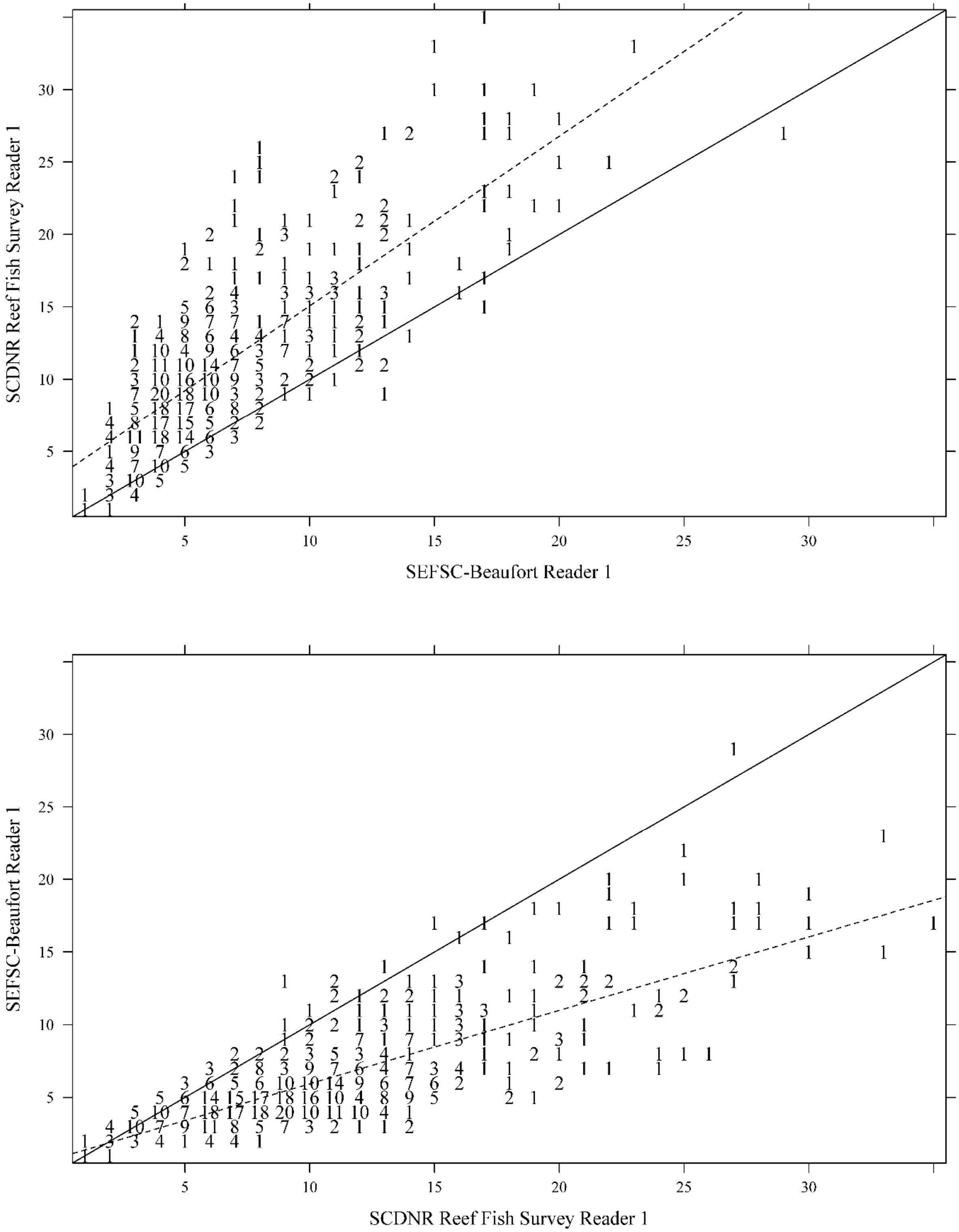
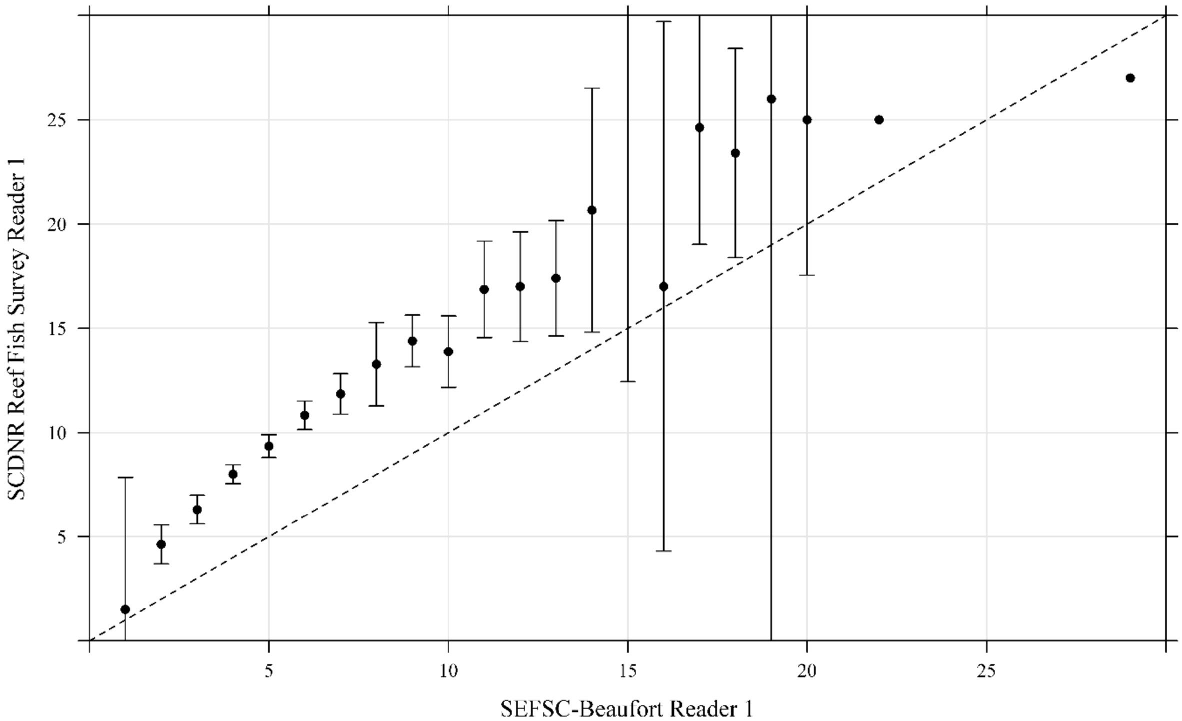


Figure 3: Distribution plot comparing SEFSC-Beaufort ages to SCDNR Reef Fish Survey Reader 1 ages for the same Blueline Tilefish. Top panel – SEFSC-Beaufort ages are considered the true ages; Bottom panel – SCDNR Reef Fish Survey Reader 1 ages are considered the true ages. Sold line – expected 1:1 line; Dashed line – linear regression model fit to the observed data; numbers – observed frequency of fish with an observed x,y combination.



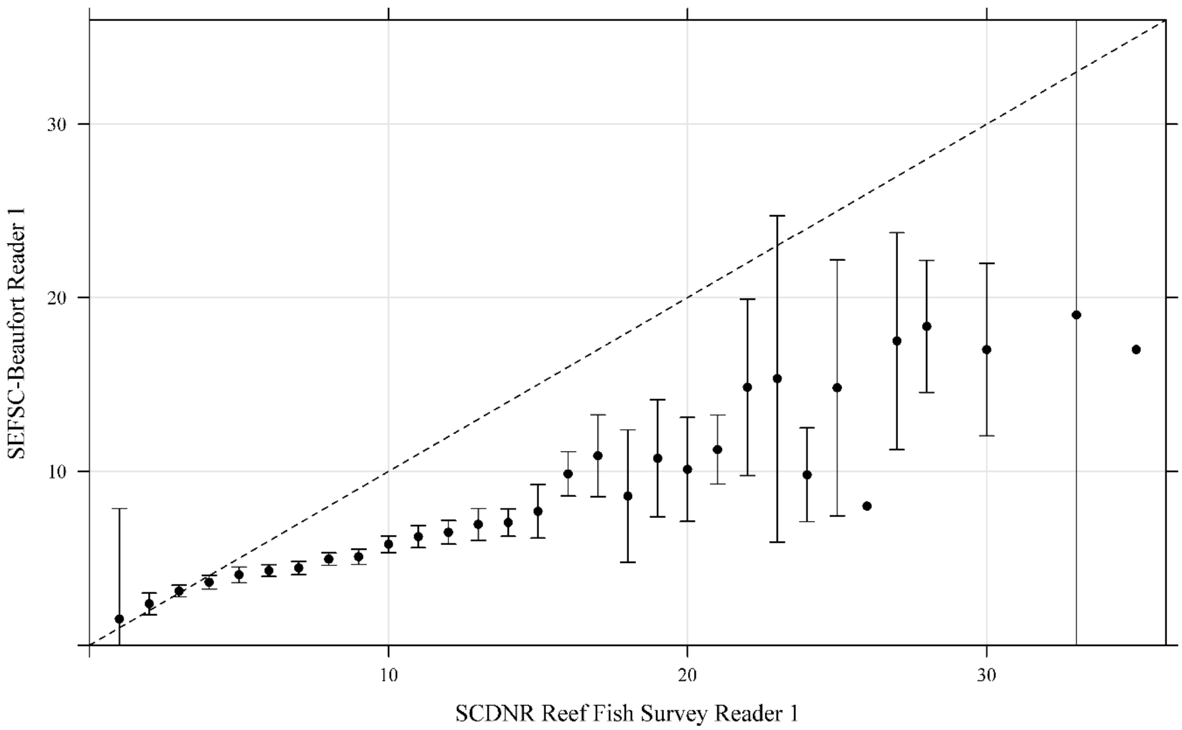


Figure 4: Bias plot comparing SEFSC-Beaufort ages to SCDNR Reef Fish Survey Reader 1 ages for the same Blueline Tilefish. Top panel – SEFSC-Beaufort ages are considered the true ages; Bottom panel – SCDNR Reef Fish Survey Reader 1 ages are considered the true ages. Dashed line – expected 1:1 line; Error bars represent the 95% confidence interval of y-axis reader ages at a given x-axis reader age.

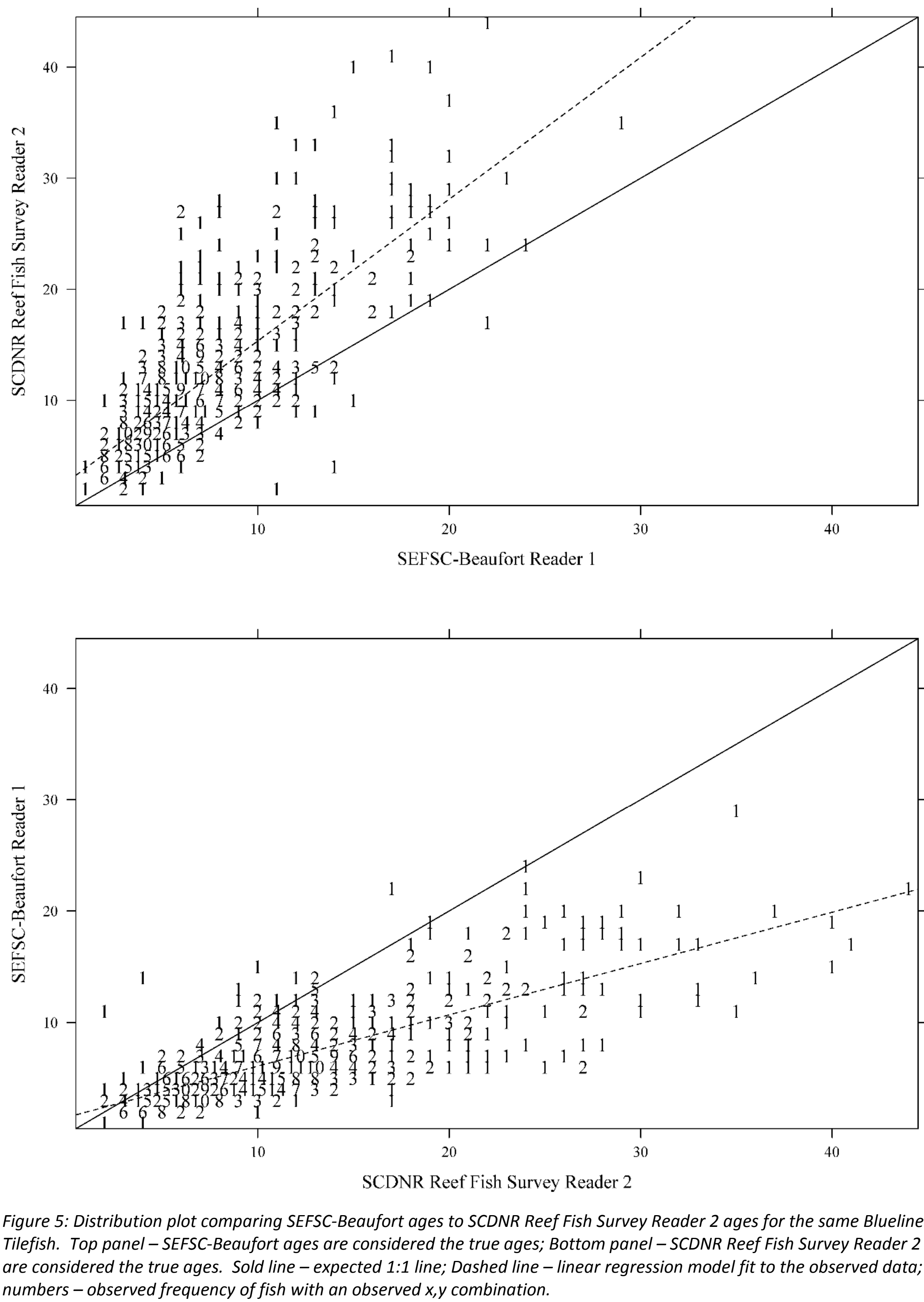


Figure 5: Distribution plot comparing SEFSC-Beaufort ages to SCDNR Reef Fish Survey Reader 2 ages for the same Blueline Tilefish. Top panel – SEFSC-Beaufort ages are considered the true ages; Bottom panel – SCDNR Reef Fish Survey Reader 2 ages are considered the true ages. Sold line – expected 1:1 line; Dashed line – linear regression model fit to the observed data; numbers – observed frequency of fish with an observed x,y combination.

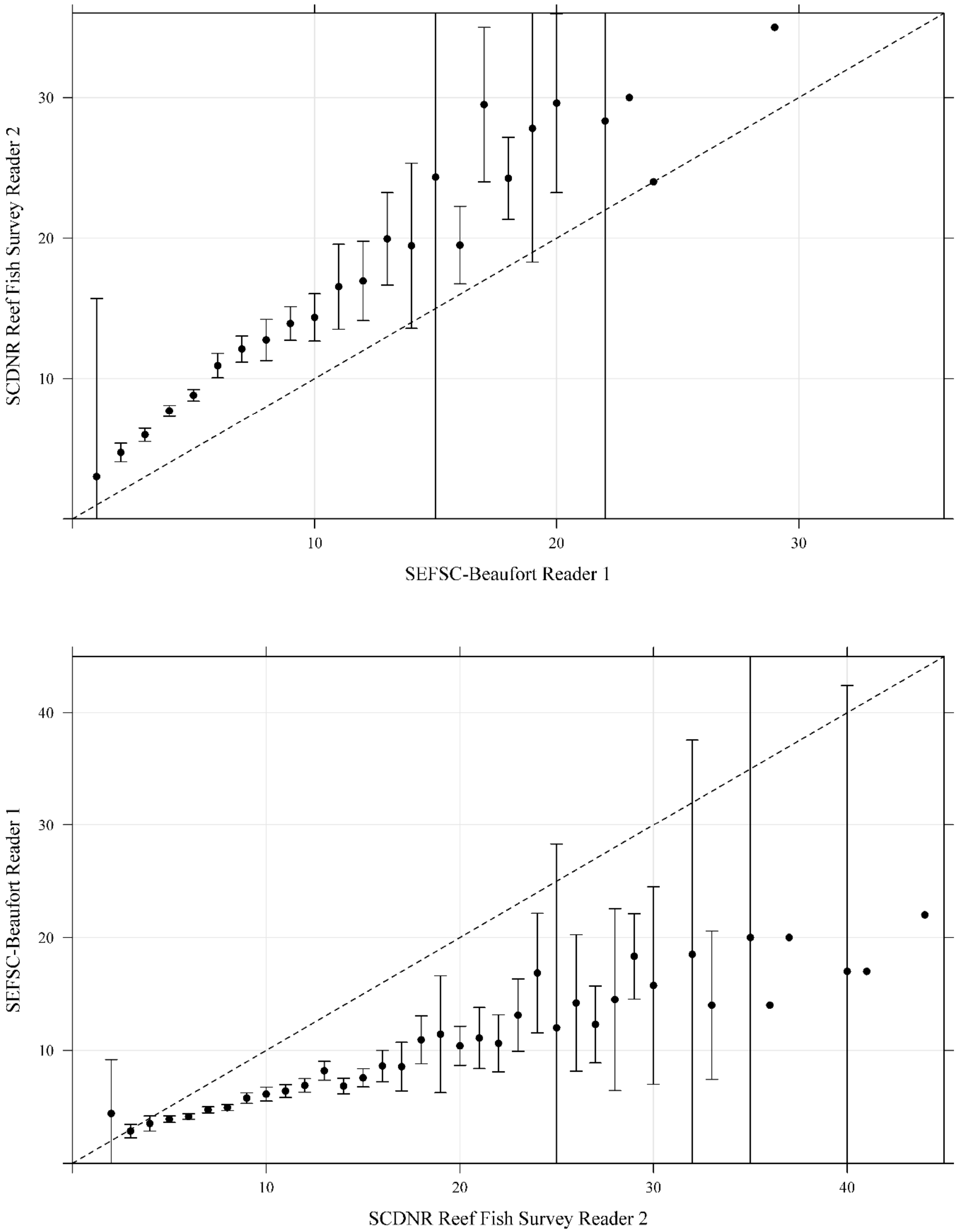


Figure 6: Bias plot comparing SEFSC-Beaufort ages to SCDNR Reef Fish Survey Reader 2 ages for the same Blueline Tilefish. Top panel – SEFSC-Beaufort ages are considered the true ages; Bottom panel – SCDNR Reef Fish Survey Reader 2 ages are considered the true ages. Dashed line – expected 1:1 line; Error bars represent the 95% confidence interval of y-axis reader ages at a given x-axis reader age.

Literature Cited

Beamish, R.J., and Fournier, D.A. 1981. A method for comparing the precision of a set of age determinations. Can. J. Fish. Aquat. Sci. **38**: 982-983.

Bowker, A.H. 1948. A test of symmetry in contingency tables. J. Am. Stat. Assoc. **43**: 572-574.

Campana, S.E. 2001. Accuracy, precision and quality control in age determination, including a review of the use and abuse of age validation methods. J. Fish Biol. **59**(2): 197-242.

Chang, W.Y.B. 1982. A statistical method for evaluating the reproducibility of age determination. Can. J. Fish. Aquat. Sci. **39**(1208-1210).

Filer, K.R., and Sedberry, G.R. 2008. Age, growth and reproduction of the Barrelfish *Hyperoglyphe perciformis* (Mitchill) in the western North Atlantic. J. Fish Biol. **72**(4): 861-882.

Friess, C., and Sedberry, G.R. 2011. Age, growth, and spawning season of Red Bream (*Beryx decadactylus*) off the southeastern United States. Fish. Bull. **109**(1): 20-33.

Hoenig, J.M., Morgan, M.J., and Brown, C.A. 1995. Analyzing differences between two age determination methods by tests of symmetry. Can. J. Fish. Aquat. Sci. **52**(2): 364-368.

Lytton, A.R., Ballenger, J.C., Reichert, M.J.M., and Smart, T.I. 2016. Age validation of the North Atlantic stock fo wreckfish (*Polyprion americanus*), based on bomb radiocarbon (^{14}C), and new estimates of life history parameters. Fish. Bull. **114**: 77-88.