FISHstory Length Analysis

FISHstory

A SAFMC CitSci Project

SAFMC SSC Meeting
October 2020

Tentative Agenda

- FISHstory Background & Overview
- Length Analysis
  - Goals & objectives
  - Protocol overview
  - Precision & accuracy analyses
  - Resampling analyses
  - Feedback and additional ideas
- Discussion

Background

- Limited recreational fisheries data prior to 1970s
- For-hire fleets have tradition of displaying catch at dock for commemorative photos
- Historic photos untapped source of data

FISHstory Overview

- Pilot project uses historic dock photos to estimate for-hire catch & length composition
- 1,300 color and black and white photos from Daytona Beach, FL
- Photos from the 1940's through 70's

FISHstory Project Components

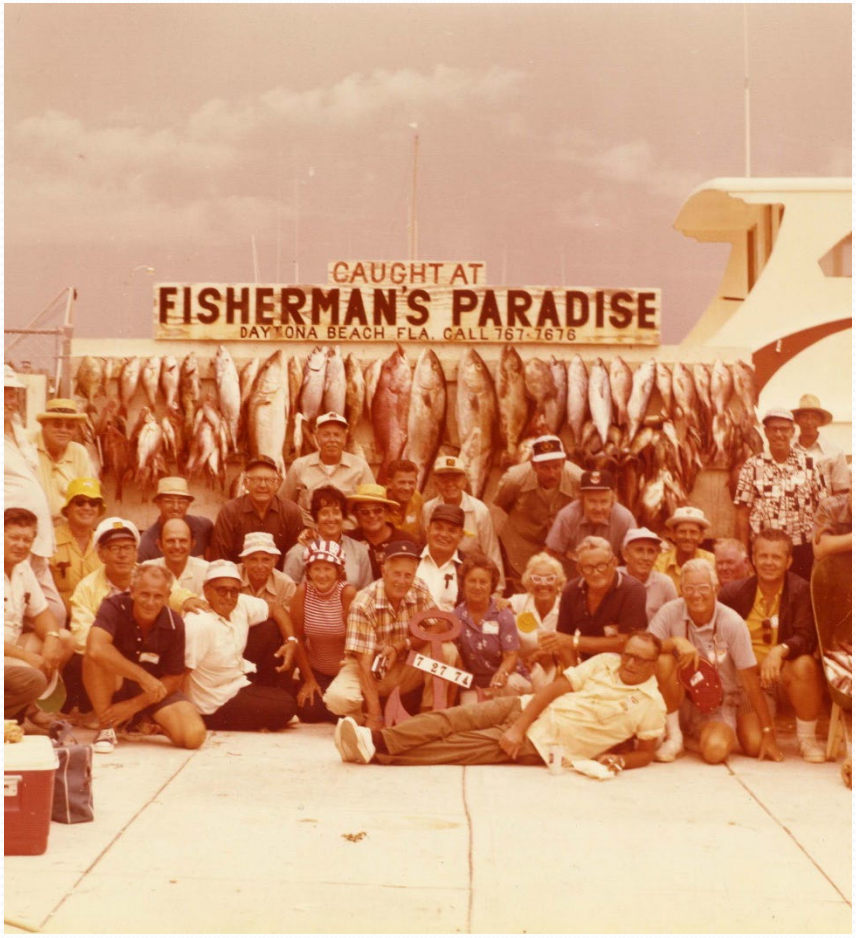
- **Digitizing historic fishing photos – *complete***
  - Photo archive created with corresponding metadata

- **For-hire catch composition - *underway***
  - Uses online crowdsourcing interface Zooniverse
  - Volunteers trained to identify and count fish and anglers in photos
  - Multiple volunteers classify each photo
  - Validation Team will help verify species identifications & counts when volunteers disagree

- **Method to estimate length composition – *today's review***
  - Estimate fish length using lumber in leaderboard as scale
  - Length measuring protocol developed using ImageJ software
  - Resampling method to produce length comps & associated uncertainty
  - Pilot tested on one species – King Mackerel

Length Analysis

Goal: Accurately describe lengths of fish from historic photos



- Test differences between readers
- Test accuracy of readers on known length objects
- Develop a protocol to collect data from historic photos
- Develop a length distribution with error estimates from historic photographs

Length Methods

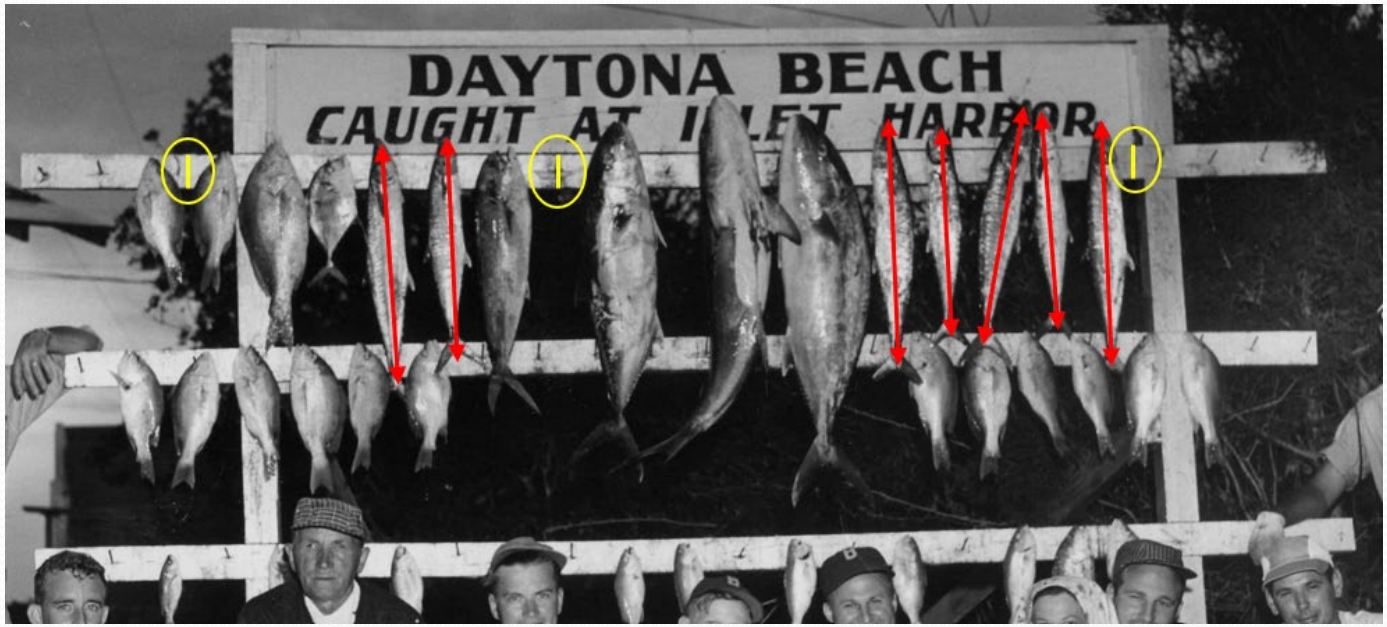
- Easy Part: Measure Fish
- Hard Part: How long is the line from snout to tail mean?
  - Need to convert unknown size of fish to length

Length Methods

- Staff worked with FISHstory Design Team to develop and pilot test draft protocol
- Detailed length analyst protocol document available
  - Includes descriptions of all data fields collected
- All data entered into Excel spreadsheet
- ImageJ used to get photo measurements
- 5 analysts trained for production length measurements
- All analysts must complete virtual training with staff & calibration photo set prior to production length measurements
- 2 individuals analyze each photo

Image J Measurements

- 3 ‘reference’ measurements taken per photo
- ‘Reference measurements taken right, middle, and left of leaderboard racks; try to bracket the fish being measured
- All King Mackerel in photo measured if possible; if can’t measure all KM in photo this denoted in spreadsheet



Scalar Development

- Compare length estimates among readers to known values
- Expansion from pixels to length
- Explore different items as scalars (labeled A-G)
- Identify scalar that performs best in estimating true length of boards

Scalar Development

Table 1. Results of regression analysis comparing predicted length with true length estimated using different scalars.

Scalar	True Length (inches)	Slope	SE	Adj R2	Slope Rank	SE Rank
Avg 2x3	2 7/16	0.99288	0.0023	0.9995	1	5
Letter H	3	0.9382	0.0038	0.9984	7	7
SAFMC Logo	5 7/8	0.95737	0.0021	0.9995	5	3
FISHstory Logo	8 1/2	0.95303	0.0022	0.9995	6	4
Wood Vertical	11	0.95765	0.0032	0.9989	4	6
Wood Horizontal*	41	0.97463	0.002	0.9998	3	2
White Board	81	0.97839	0.0016	0.9997	2	1

*Only one analyst measured this scalar

Scalar Development

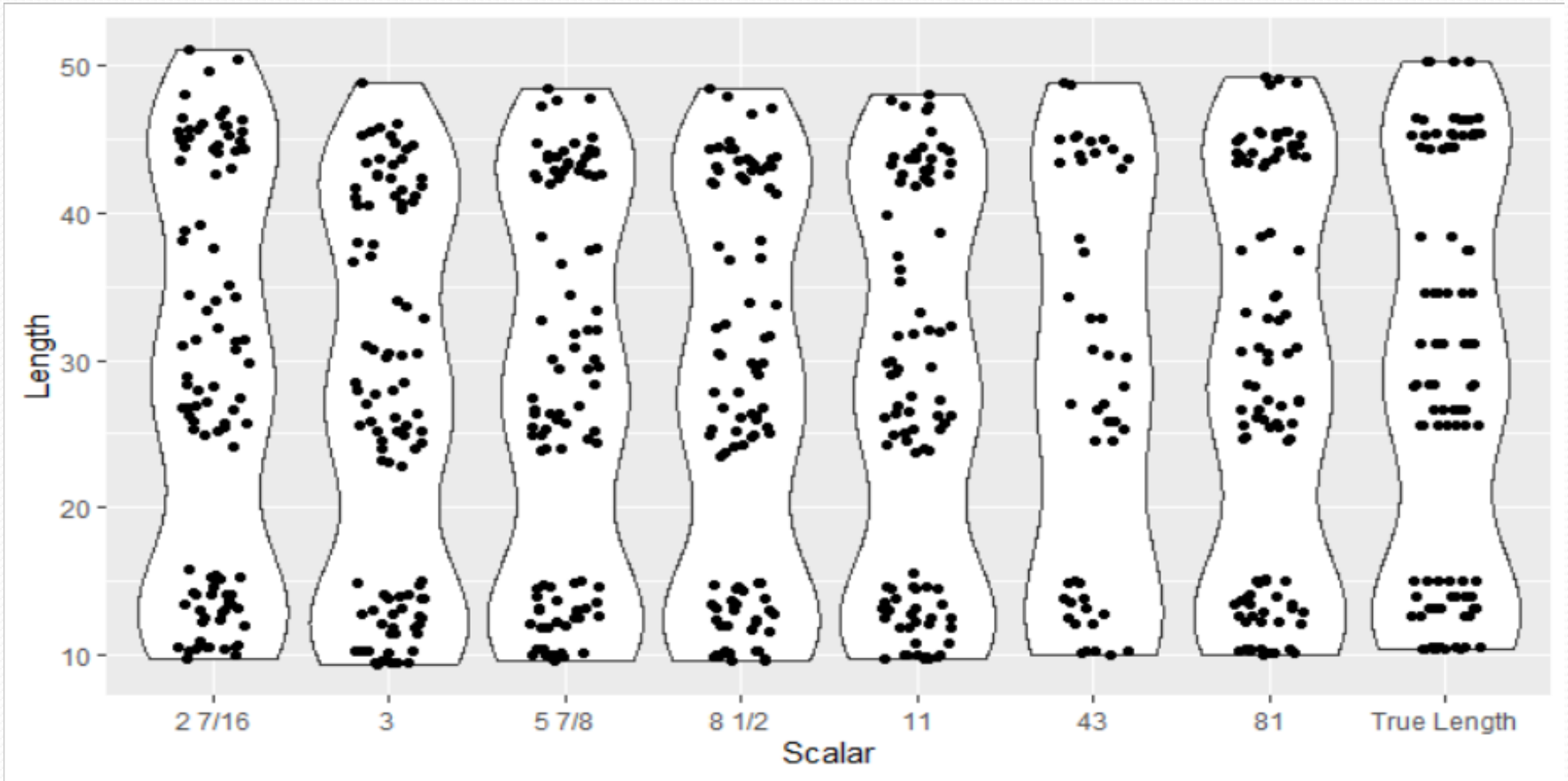


Figure 1. Distribution of board lengths estimated using different scalars.

Scalar Development

- Compare length distribution from a subset of photos estimated by two independent readers and keep as a training set
  - Kolmogorov-Smirnov
- Compare paired length estimates among readers
  - Paired T-Test
  - Wilcoxon Signed Rank Test

Table 2. Comparison of length estimates for primary readers using 2X3 scalar

Test	Statistic	p-value
Kolmogorov-Smirnov	.08333	0.9969
Paired T Test	-1.9716	0.05456
Wilcoxon Rank Sum (paired)	88	0.05557

Scalar Development
Accuracy

Goal: Accurately estimate length of fish within two inches based on the bin size for King Mackerel in SEDAR 38 Updated (SEDAR 2020)

Table 3. Comparison of percent of estimated length measurements within 2 inches of the true length measures for each scalar.

Error from True Length	2 7/16 (avg)	3	5 7/8	8 1/2	11	41*	81
< 2 inches	96%	66%	74%	78%	71%	100%	99%

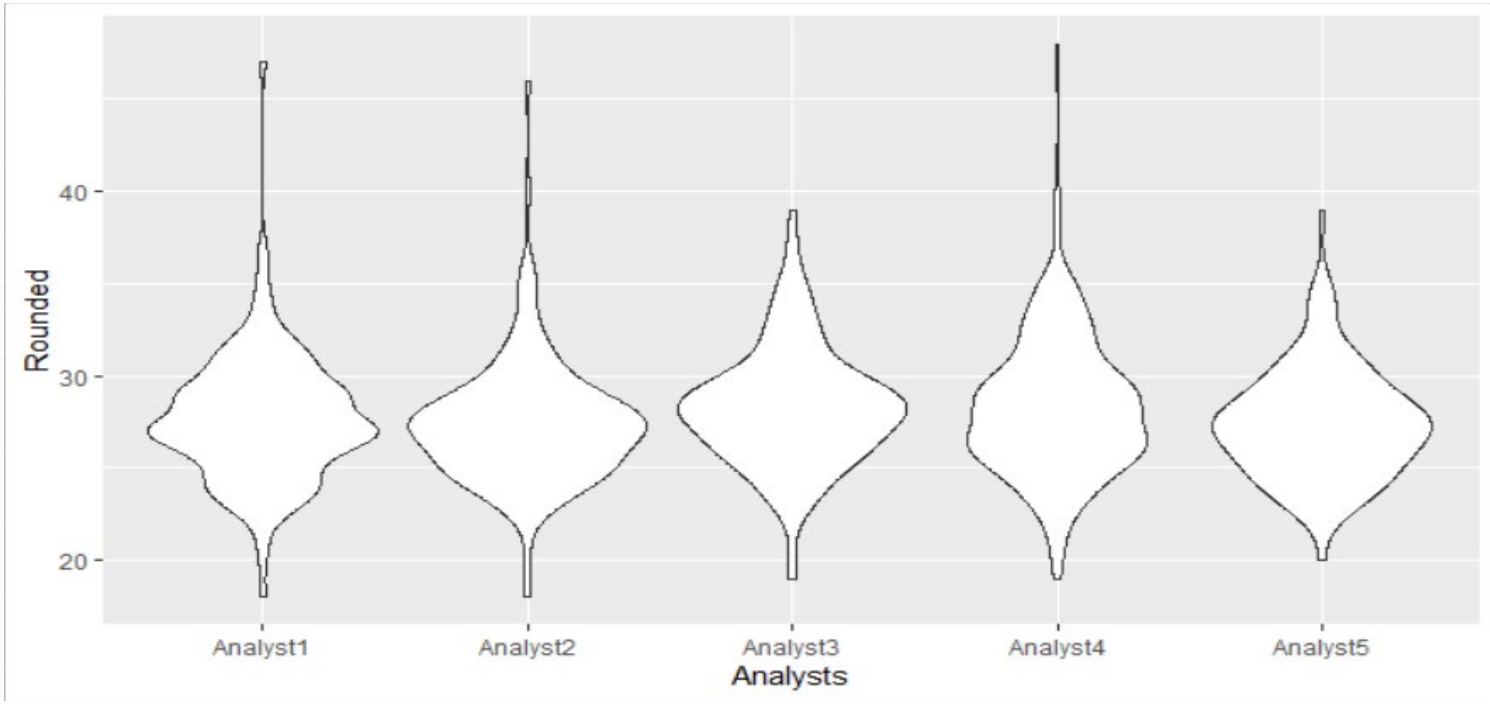
*Only one analyst measured this scalar

Scalar Development Selection

Average 2X3 selected as the preferred scalar method

- Had slope closest to 1 indicating no bias with increasing size
- Best captured full range of size distributions
- Length estimates developed by the readers were not significantly different
- Estimates from the readers were within 2 inches of actual size for 96% of the measured boards.

Comparison of Historical Photographs



- Visual analysis indicated difference in the size distribution was present among
- Variation seemed to be 1 inch different which is smaller than the 2-inch size bins.

Comparison of Historical Photographs

Table 4. P-values for the Komolgorov-Smirnov (left of diagonal) and Anderson-Darling tests (right of diagonal) used to compare the historical photographs length estimates by the different readers. Yellow indicates significant differences

KS/AD Test	Analyst 1	Analyst 2	Analyst 3	Analyst 4	Analyst 5
Analyst 1	-	0.547	0.024	0.1119	0.995
Analyst 2	0.5445	-	0.0173	0.053	0.7945
Analyst 3	0.0302	0.0025	-	0.467	0.031
Analyst 4	0.2281	0.0568	0.5061	-	0.0956
Analyst 5	0.995	0.9495	0.0265	0.1436	-

Comparison of Historical Photographs

Accuracy

Phyllis Hudson – known height

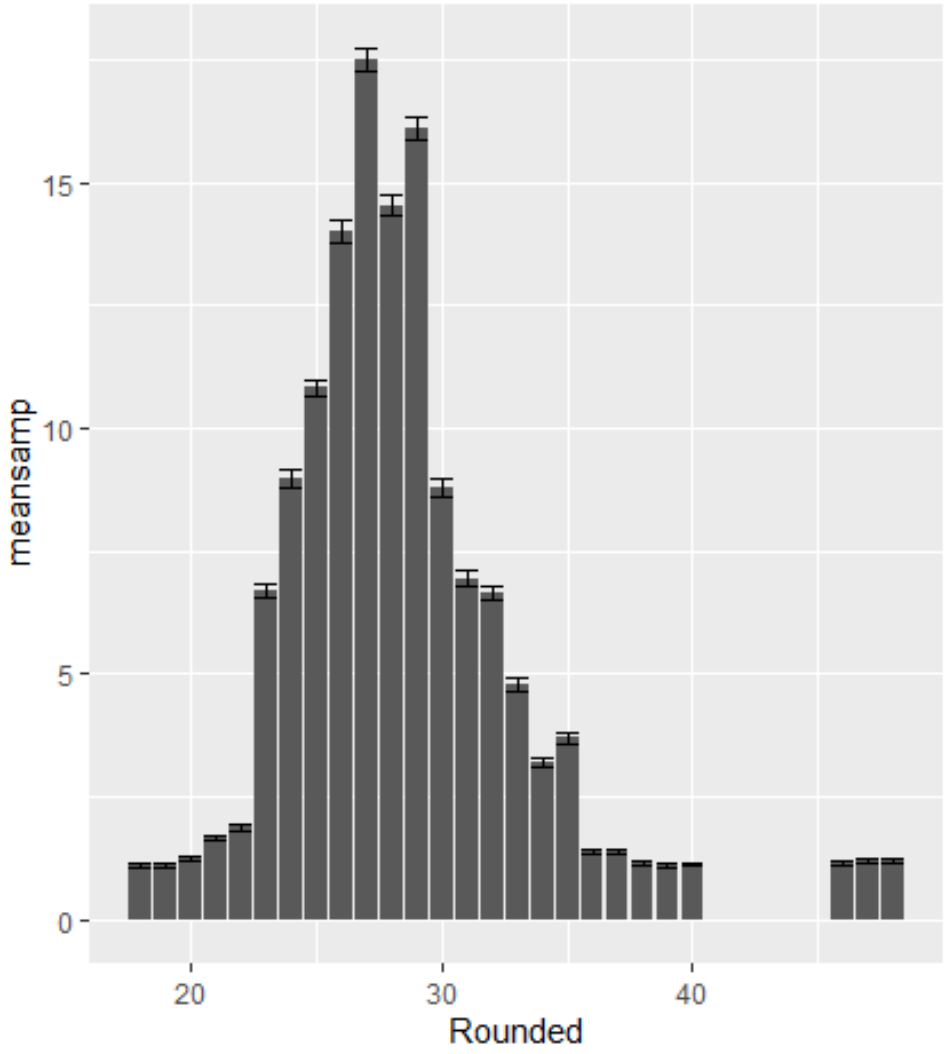
62.5% of length estimates within two inches (n=8)

Oil Barrel – standard height

71% of length estimates within two inches (n=49)

Length Distribution from Historical Photographs

- Bootstrap resampling based on photograph
- Combine lengths from each photograph
- Error bars derived from resampling



Length Distribution from Historical Photographs

- Significant differences in length estimates were observed among readers
- The accuracy of objects with known size were estimated with some success
  - All length estimates were within 5 inches of the true length
  - Most lengths (>60%) were within 2 inches of true length
- Length distribution
  - Lengths ranged from 18 to 48 inches fork length
  - Peak lengths ranged between 26 to 29 inches fork length

SSC Discussion Questions

- Is this methodology appropriate to use for measuring fish in pictures?
- Can an informative size composition of catch be derived using this methodology?
- Does the methodology adequately address uncertainty for the size composition?