Proposal

Background

Age estimation of fish is necessary for the assessment of fish population dynamics (e.g., recruitment, growth, and mortality) and stock structure. Ages are most often estimated from calcified structures. Those ages may vary among readers, among calcified structures, or among readings by the same reader. It is important to understand both the accuracy and precision of age estimates to better understand how biases and variability may affect estimates of population dynamic metrics.

Accuracy is related to estimating the true age of the fish. Accuracy is related to validation which was extensively discussed by Campana (2001) and won't be discussed further here.

Precision is related to the consistency or repeatability of age estimates among readings (i.e., among readers or among readings by the same reader at different times). Precision was also discussed extensively in Campana (2001) but will be discussed further here given comments and questions in Bauerlien et al. (2018).

Campana (2001) described three measures of precision for age estimation analyses. Percent agreement is the percentage of fish for which the readers' estimates of age perfectly agreed. Some authors have extended this concepts to the percentage of fish for which the readers' estimates of age matched within a certain tolerance range (e.g., within two years).

Absolute percent error (APE) is

$$APE = 100 * \frac{\sum_{j=1}^{n}\sum_{i=1}^{R}\frac{|x_{ij} - \bar{x}_j|}{\bar{x}_j}}{nR}$$

where x_{ij} is the ith age for the jth fish, $\bar{x}_j$ is the mean age for the jth fish, R is the number of times that each fish was aged (assumed to be the same for all fish), and n is the number of aged fish in the sample. Thus, the APE is the average absolute deviation in ages among readers scaled (i.e., divided) by the mean age among readers averaged across all n fish.

The coefficient of variation, which Ogle (2016) termed average coefficient of variation (ACV), is

$$ACV = 100 * \frac{\sum_{j=1}^{n}\frac{s_j}{\bar{x}_j}}{n}$$

where s_j is the standard deviation of the R age estimates for the jth fish. Thus, the ACV is the coefficient of variation (i.e., the standard deviation scaled by the average age estimate) averaged across all n fish.

The ACV is greater than the APE by a factor of $\sqrt(2)$ when $R = 2$ (Kimura and Anderl 2005). Thus, functionally the APE and ACV are interchangeable when $R = 2$ (McBride 2015). However, ACV may be more familiar and easier to interpret (Kimura and Anderl 2005) and appears to be the more common choice among analysts.

Bauerlien et al. (2018) introduced the mean absolute deviation (MAD), which we will call the average absolute deviation (AAD) here for consistency, as

$$AAD = 100 * \frac{\sum_{j=1}^{n}\sum_{i=1}^{R}|x_{ij} - \bar{x}_j|}{nR}$$

The AAD is very similar to the APE, except that the absolute deviations have not been rescaled by the mean age estimate.

For completeness, one could also consider the average standard deviation (ASD) as

$$ASD = 100 * \frac{\sum_{j=1}^{n} s_j}{n}$$

The ASD is very similar to the ACV, except that the standard deviation has not been rescaled by the mean age estimate.

Problem Statement

Precision is (anecdotally … see below) most often measured by the ACV. The ACV is designed to address the issue that differences among age estimates of say one year for young fish is more important than the same differences for old fish. In other words, the measure of precision should be adjusted for the (estimated) age of the fish. However, two possible issues arise with the use of ACV. First, interpretation is made more difficult. For example, an ACV of 10% is more difficult to interpret than say an ASD of 1.2. In the latter, 1.2 can be interpreted as ages differing by 1.2 years on average. However, the analyst likely will convert the ACV to actual age differences for interpretation (e.g., "an ACV of 10% means differences of 1 year for age 10 fish but 2 years for age 20 fish"). Second, the ACV is meant to remove the effect of age on the measure of precision. However, it is possible that ACV is still related to the age of the fish. If ACV is related to the age of the fish then the practice of reporting a single ACV value (averaged across all fish and, thus, all ages) is misleading. As suggested by Bauerlien et al. (2018), the analyst should assess the relationship between any of the precision metrics and age before summarizing with a single value.

Questions that we wish to consider are:

1. Which precision metric is most commonly used in the literature?
2. What is the distribution of precision metrics found in the literature?
3. Do authors test to determine if their chosen precision metric is related to age or not?
4. How often are precision metrics (especially ACV) related to age?
5. What kind of relationship (linear, quadratic, or more complicated) between precision metrics and age are most common?
6. What are the relationships between the various precision metrics?
7. Do other metrics of precision exist and are they useful to fisheries biologists?
8. What advice should be given to fisheries biologists when choosing a precision metric?

Research Direction