The report "Improving estimation of rare event species in MRIP" was prepared by a team of statisticians from Westat in consultation with the NOAA Fisheries MRIP Working Group on Alternative Estimation Methods for Rare Event Species. The report represents findings based on an initial examination of multiyear averaging options to improve catch estimates. It should be noted that the estimates produced through these analyses are not official NOAA Fisheries recreational fishery catch estimates and are being provided as exemples of methods based on available MRIP data that may be used to improve the precision of catch estimates based on MRIP data. Official estimates may be found at - https://www.fisheries.noaa.gov/data-tools/recreational-fisheries-statistics-queries

Improving estimation of rare event species in MRIP

Jean Opsomer, Angela Chen, Clément Guillo
Westat

Project Overview

For a number of species that are rarely encountered in APAIS intercepts, the standard MRIP estimation approach results in estimates that are highly variable or missing in some years, making it difficult to use those in the setting of annual catch limits. The goal of the project was to investigate whether using the same data as collected during standard MRIP surveys but alternative estimation methods can provide estimates that are more stable over time.

The key strength of the standard MRIP estimation approach is that the resulting estimates are approximately unbiased. When sufficient data for a target species are available, the estimates for that species are sufficiently reliable to use in fishery monitoring. The alternative estimation methods proposed here are no longer guaranteed to be unbiased, but the expectation is that reductions in variance might be sufficient to result in estimates that have lower mean squared error (MSE) than the standard estimates. In evaluating the alternative estimates, we therefore wanted to gain insights in their bias and variance characteristics. We also wanted to find a method that would work well across a range of species, to avoid requiring customized estimation methods for individual species.

For this investigation, we considered 21 species, representing a mix of rarely and more commonly observed species:

- *Atlantic spadefish*
- *bar jack*
- *black grouper*
- *black sea bass*
- *blueline tilefish*
- *chub mackerel*
- *Cubera snapper*
- *little skate*
- *queen snapper*
- *red grouper*
- *red hake*
- *red porgy*
- *scamp*
- *silk snapper*
- *silver hake*
- *snowy grouper*
- *(golden) tilefish*
- *Warsaw grouper*

- *white hake*
- *windowpane flounder*
- *yellow-edge grouper*

Data from the time period January 2004 to December 2018 were used in this analysis. For this period, 1.4 million records of trips are available with an average of 95,000 trips per year which, using the MRIP sampling weights, result in an average fishing effort per year equal to 20 millions of trips.

Estimation approaches

The class of estimators that were considered for this analysis involved multi-year averages. We begin by introducing the notation used to define the estimators. The estimated number of trips that took place for a given geographic area (e.g. state) and time period (e.g. year, wave) is denoted generically by $\hat{N} = \sum_{s} w_i$, where w_i represents the weight for trip i and s is the sample of trips corresponding to that geographic area and time period. To denote the estimate and sample for a target area and period in what follows, we will write $\hat{N}^*$ and s^*, respectively, in that case. Let y_i represent a variable recorded for trip i, for instance the number of fish landed of a particular species. The traditional MRIP estimate for the total of that variable is $\hat{t}_y = \sum_{s} w_i y_i$, again denoted $\hat{t}_y^*$ if it is a target quantity. We write t_y^* for the unknown population target.

As noted above, the estimator $\hat{t}_y$ is (approximately) unbiased but imprecise (large variance) when the number of trips $i \in s$ with observed y_i is small. We will therefore construct alternative estimators as follows. For a given target area and period, the total number of trips $\hat{N}^*$ and sample s^* are maintained, but in order to increase the number of observed trips with catch, we estimate the average of y_i over a sample s^{**} that is larger than s^*. We will do so by combining the averages from samples that are expected to be "similar" to the target sample. Let $s_j, j = 1\ldots,J$ represent these samples, and we will assume that $s^*$is contained among the s_j. The combined average, denoted $\tilde{y}^*$, is a linear combination of the averages over the s_j, $\tilde{y}^* = \sum_{j=1}^{J} a_j \hat{y}_j$, with $\hat{y}_j = \sum_{s_j} w_i y_i / \sum_{s_j} w_i$. The coefficients a_j determine the linear combination of the J averages, and we write a^* for the coefficient of the target year. Finally, the alternative estimator for the target area and period is then defined as $\tilde{t}_y^* = \hat{N}^* \tilde{y}^*$.

To make this more specific, if $s^*$is the sample for a target state and year, then s^{**} covers the same state but multiple years, e.g. a 5-year period centered at the target year. Alternative versions, including having s^{**}cover a wider geographic area, are possible but not pursued here.

Instead, we will focus on evaluating different ways to expand s^{**} temporally only, with the J averages corresponding to the years in a neighborhood of the target year. We will also refer to y_i as the catch, but the approach can be used for any variable associated with an MRIP trip.

Two types of multi-year estimators were evaluated: moving averages and time series predictions. Moving averages use data from both prior and future years to estimate for the target year. In other words, the sample s^{**} is centered at the target year and the same number of years before and after that year are used in constructing $\hat{y}^{**}$. Time series predictions use only prior years to do so, which is more useful if the goal is to improve estimates on an on-going basis. In contrast, the moving average can be used to improve historical catch data.

Because the multi-year estimators are arithmetic functions of survey estimators, it is possible to study their statistical properties and to estimate their variance. For the latter, standard linearization-based or replication variance estimation methods are readily applied. Using linearization, the variance estimator of $\tilde{t}_y^* = \hat{N}^* \tilde{y}^*$ is

$$\hat{V}(\tilde{t}_y^*) = \hat{V}(\hat{t}_u^*) + \hat{N}^{*2} \sum_{j \neq j^*} \frac{a_j^2}{\hat{N}_j^2} \hat{V}(\hat{t}_{e,j})$$

with $u_i = a^* y_i + \sum_{j \neq j^*} a_j \hat{y}_j$ and $e_{ij} = y_i - \hat{y}_j$. The variance terms in this expression can be readily obtained for MRIP using standard survey software, once the new variables have been computed.

Regardless of the type of multi-year estimator, unless the average catch over the period covered by s^{**} is constant, $\tilde{t}_y^*$ will be biased for t_y^*. However, it is expected to be less variable than $\hat{t}_y^*$, so that it can result in a lower Mean Squared Error (MSE). Another advantage of $\tilde{t}_y^*$ over $\hat{t}_y^*$ is that it can provide an estimate even in years in which no catch is observed in s^*.

For both types of estimates, either 3 or 5 year time periods were considered. We first considered equal-weighting of the individual years, i.e. $a_j = 1/J$. However, for the 5-year moving averages, we looked at further alternatives that weight the individual years differentially (higher weight to current and neighboring years vs more distant years). Table 1 shows the estimators we evaluated in this study, together with their linear combination coefficients a_j.

Table1: Coefficients for original and multi-year estimators evaluated in this study.

Estimator	-4	-3	-2	-1	Target year (*)	+1	+2
Original estimator (ORN)					1		
Time series prediction, 3-year (TS3)			1/3	1/3	1/3		
Time series prediction, 5-year (TS5)	1/5	1/5	1/5	1/5	1/5		
Moving average, 3-year (MA3)				1/3	1/3	1/3	
Moving average, 5-year (MA5)			1/5	1/5	1/5	1/5	1/5
Moving average, 5-year, modified version 1 (MA5_1)			1/10	¼	3/10	¼	1/10
Moving average, 5-year, modified version 2 (MA5_2)			1/12	¼	1/3	¼	1/12
Moving average, 5-year, modified version 3 (MA5_3)			1/12	1/6	1/2	1/6	1/12

Each estimator was applied to the 21 species, for each year between 2004 and 2018 for which the necessary years of data were available, and compared to the original MRIP estimator in those years.

Assessment metrics

The "optimal" estimator among the seven considered is the one that achieves the smallest MSE across all species and time periods considered. The MSE is not available, so instead we evaluate the bias and variance performance of the different estimation approaches separately. We consider two performance measures:

- the fraction of times that the target estimator stays within the 95% confidence interval of the traditional MRIP estimates ("CI fractions"),
- the estimated standard deviation of the target estimator, expressed as coefficient of variation ("CV").

The latter is a direct measure of the variability of the target estimator and obtained directly from the design-based properties of the estimator. The former is an indirect measure of the bias of the estimator, which requires some additional explanation because it is non-standard. As already noted, the traditional MRIP estimator is approximately unbiased, and in addition, its confidence interval is a valid (but highly variable) measure of the unknown true catch. Hence, if the target estimator falls outside that confidence interval, this is an indication that it is likely to be subject to bias. The variability in these confidence intervals makes this interpretation unreliable for any given year. However, when considered across multiple years, the fractions of times the target estimator falls outside the confidence intervals provides a more stable indication of potential bias.

Results

Table 2 show the CI fractions of the original and multi-year estimators for the 21 species used in this evaluation, and Table 3 shows the average annual CVs for the estimators.

Table 2: Fractions of years in which each estimator lands within 95% confidence interval of MRIP estimator.

species	ORN	TS3	TS5	MA3	MA5	MA5_1	MA5_2	MA5_3
atlantic spadefish	100%	77%	82%	92%	73%	73%	91%	100%
bar jack	100%	38%	27%	46%	36%	36%	36%	36%
black grouper	100%	85%	82%	85%	82%	82%	82%	82%
black sea bass	100%	85%	73%	92%	100%	100%	100%	100%
blueline tilefish	100%	46%	55%	69%	55%	73%	73%	73%
chub mackerel	100%	54%	36%	38%	27%	36%	36%	36%
cubera snapper	100%	46%	64%	62%	45%	55%	55%	55%
little skate	100%	54%	55%	77%	64%	64%	64%	73%
queen snapper	100%	38%	27%	46%	18%	18%	18%	18%
red grouper	100%	77%	45%	85%	91%	91%	91%	91%
red hake	100%	92%	91%	85%	82%	91%	91%	91%
red porgy	100%	85%	73%	92%	91%	91%	91%	91%
scamp	100%	92%	91%	100%	91%	100%	100%	100%
silk snapper	100%	62%	36%	69%	36%	36%	45%	55%
silver hake	100%	85%	100%	62%	55%	64%	64%	64%
snowy grouper	100%	62%	45%	69%	64%	64%	73%	82%
tilefish	100%	38%	45%	54%	36%	36%	36%	36%
warsaw grouper	100%	69%	45%	54%	55%	55%	55%	73%
white hake	100%	69%	73%	85%	73%	73%	73%	73%
windowpane	100%	62%	45%	54%	55%	55%	55%	55%
yellowedge grouper	100%	77%	64%	69%	45%	55%	55%	55%

Table 3: Average annual estimated CV of each estimator.

species	ORN	TS3	TS5	MA3	MA5	MA5_1	MA5_2	MA5_3
atlantic spadefish	29%	18%	15%	18%	15%	15%	16%	17%
bar jack	66%	58%	56%	58%	56%	58%	58%	58%
black grouper	47%	28%	22%	28%	22%	24%	24%	26%
black sea bass	13%	8%	7%	8%	7%	7%	7%	8%
blueline tilefish	42%	28%	23%	28%	23%	24%	24%	25%
chub mackerel	70%	67%	61%	67%	61%	64%	64%	64%
cubera snapper	85%	67%	57%	67%	57%	57%	57%	57%
little skate	67%	54%	42%	54%	42%	44%	45%	46%
queen snapper	95%	92%	90%	92%	90%	90%	90%	91%
red grouper	19%	11%	9%	11%	9%	9%	10%	11%
red hake	46%	29%	25%	29%	25%	25%	26%	28%
red porgy	22%	13%	10%	13%	10%	11%	11%	13%
scamp	35%	23%	18%	23%	18%	20%	20%	22%
silk snapper	73%	66%	72%	66%	72%	73%	73%	74%
silver hake	68%	56%	47%	56%	47%	47%	47%	48%
snowy grouper	53%	38%	36%	38%	36%	36%	36%	36%
tilefish	59%	43%	40%	43%	40%	39%	39%	39%
warsaw grouper	69%	60%	54%	60%	54%	55%	55%	55%
white hake	57%	35%	28%	35%	28%	28%	28%	30%

windowpane	85%	72%	65%	72%	65%	64%	65%	66%
yellowedge grouper	77%	54%	52%	54%	52%	50%	49%	50%

In order to better assess the bias/variance tradeoff of the estimators, we computed the average CI fractions and average CV across the species and plotted them against each other. Figure 1 shows the results. The original estimator is exactly unbiased (average CI fraction equal to 100%), but has a CV well above those of the multi-year estimators. The multi-year estimators are all subject to bias, with their CI fractions mostly between 60 and 70%, but achieve smaller CV.

Figure 1: Performance of estimators on both evaluation measures

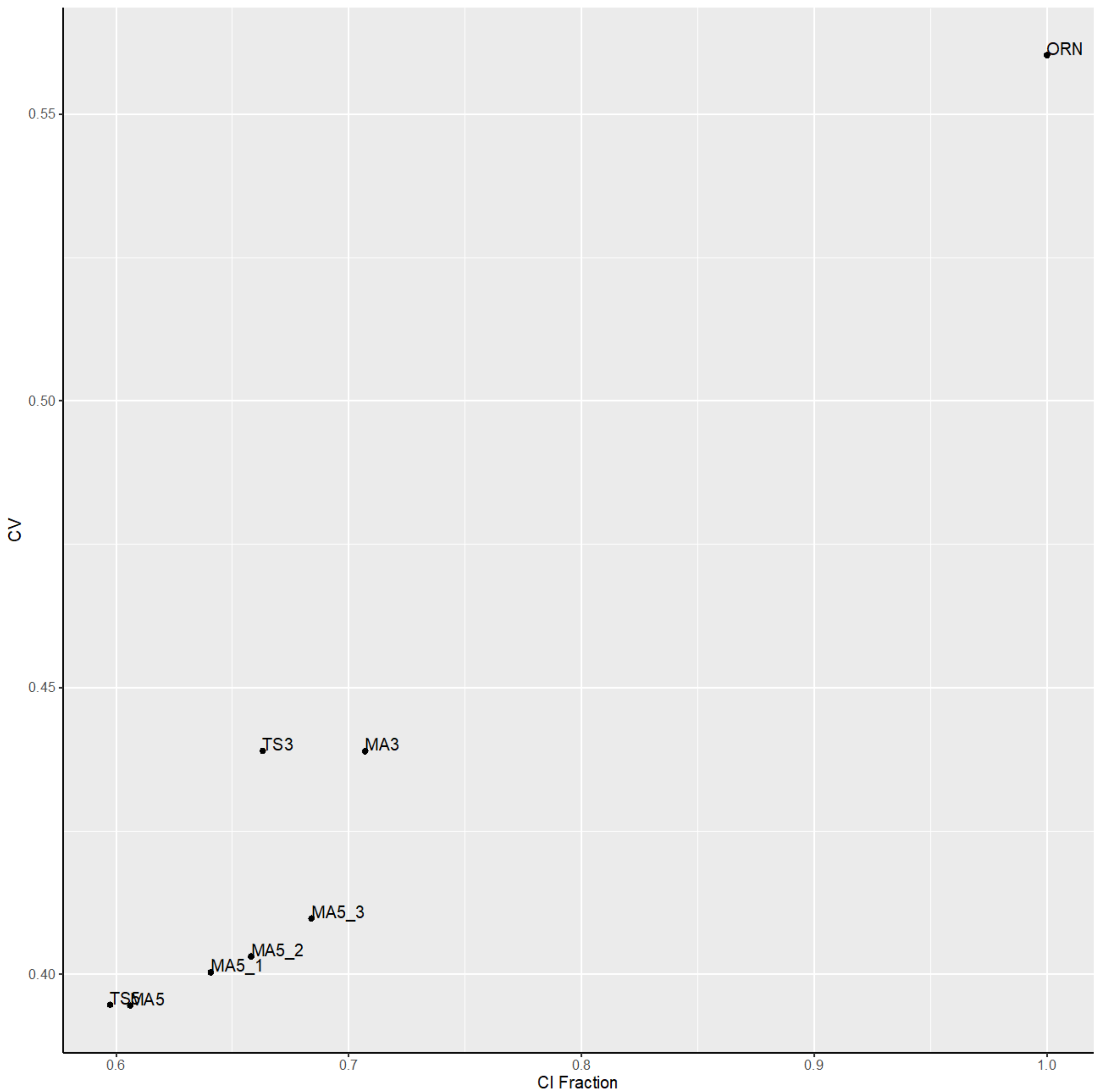
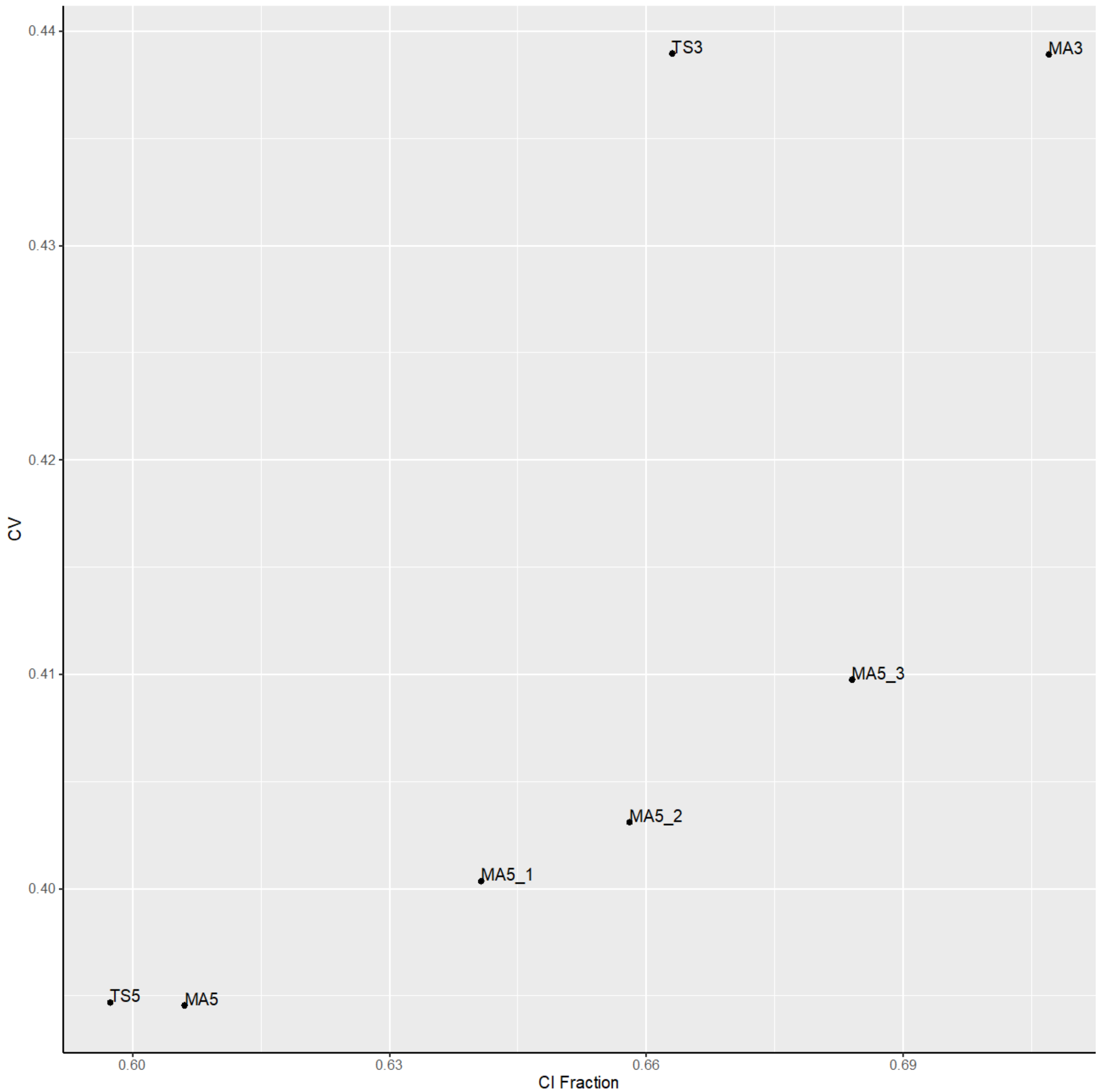


Figure 2 shows the same results but with the original estimator removed. Not surprisingly, the 3-year estimators have higher CVs than the 5-year ones. In each group (3-year, 5-year), the time series estimators have similar CVs but smaller CI fractions, indicating that they are likely to be more biased. Nevertheless, the 5-year time series estimator has results that are almost identical to those of the equal-weighted moving average estimator. Among the 5-year moving average estimators, the different sets of weighting coefficients lead to different bias and variance tradeoffs, with MA5_3 the lowest bias/highest variance and MA5 the highest bias and lowest variance. It should be noted however that the differences between them are modest.

Figure 2: Performance of multi-years estimators on both evaluation measures.



Conclusions

The analyses focused on a comparison of different estimators that combined catch data across multiple years to create annual catch estimates. The results indicate that combining more years appears preferable (comparing 3 and 5 year combinations), and that estimators that use data from both past and future years (MA) are preferable to those that only use past years (TS). The choice among the different weighting approaches for the 5-year moving average estimators impact the results but not sufficiently to decide which is best. On balance, we would recommend the MA5 for improving historical time series and the TS5 for setting catch limits for current years.