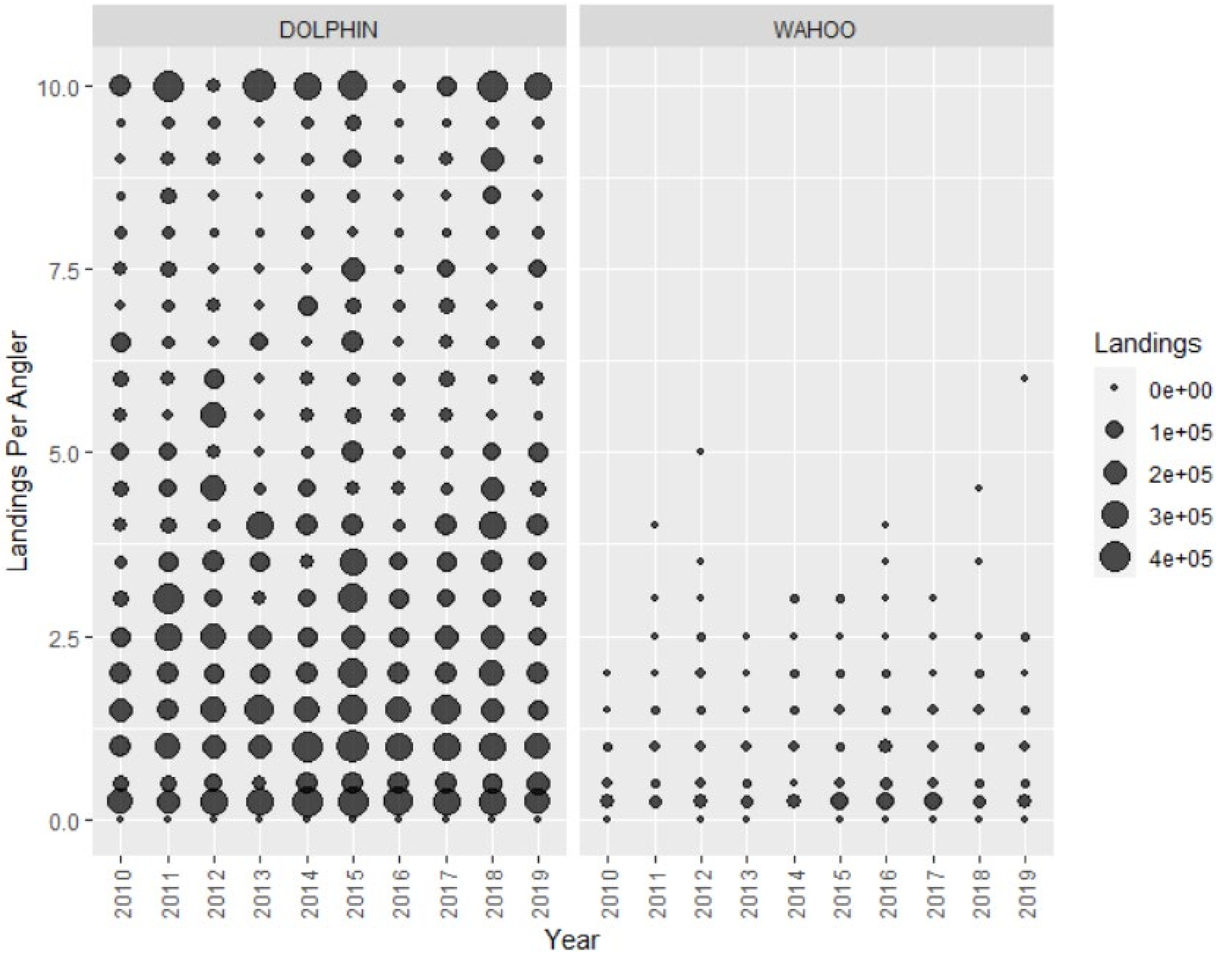


Figure 2. Landings per angler for Dolphin and Wahoo in the South Atlantic region including Monroe County from 2010 to 2019. The size of the circle is scaled to the total landing reported through MRIP. Source: Personal communication from NMFS, Fisheries Statistics Division, January 11, 2021.

Using a catch estimate from a single year seems to be the simplest and easiest method to evaluate the need for accountability measures. However, to get management measures in place and let fishermen plan their trips for the following year, DW 10 includes a notification date of up to September 1 to announce the bag limits, trip limits, and season for dolphin (if needed). If restrictive measures are deemed necessary as part of the accountability measure after September 1, only a shortening of the season will be considered. This will require the National Marine Fisheries Service (NMFS) to use preliminary data for the earlier part of the year (Waves 1-3, 2010 to 2019 average: dolphin = 50%, wahoo = 38%) and potentially project catches from previous years for the later part of the year (Waves 4-6, 2010 to 2019 average: dolphin = 50%, wahoo = 62%). In this situation, NFMS will be relying on preliminary and prior years of data to evaluate the need for accountability measures and potentially implement temporary management measures. This approach could work well if the fishery and catch estimates are fairly stable from year to year and during the year. Due to the uncertainty in recreational catch estimates, it carries risk of implementing an accountability measure that is not necessary if the estimated catch exceeds the true catch, or of allowing overfishing if the estimated catch is below the true catch. Using a multi-year averaging approach may help address this risk as well as address the added uncertainty provided by preliminary data or the unusual wave estimates that happen far too often.

Additionally, unlike the commercial sector or headboat component of the recreational sector, whose landings are based on a census of trips, the private and charter components of the recreational sector are based on a surveyed subset of trips. Since surveys do not encompass all trips, there is potential for random deviations from the true value. The presence of random deviations was investigated by calculating the mean catch over a decade for 17 species managed by the SAFMC from 2010 to 2019 (sub-region 6 in MRIP) and identifying years with catch estimates two standard deviations above or below the mean, defined as a spike for above and dip for below. Species with a trend in the data were removed since the trend in the data could cause certain years to exceed the standard deviations. Seven of the 13 remaining species had one spike over the decade (blueline tilefish and yellowtail snapper were only 2,000 lbs away from having a spike in one year) and none were identified with a dip (**Table 1**). Since the spike is not combined or smoothed with other years like the two methods described in more detail below, using a single point estimate results in the highest risk of implementing an accountability measure due to error in the survey.

Table 1. Seventeen species managed by the SAFMC to determine if spikes or dips (two standard deviations above and below the mean) were presented compared to the mean catch from 2010 to 2019. Species with trends were not compared to the mean.

Species	Trend in Landings	Spike	Dip
BLUELINE TILEFISH		*	
DOLPHIN		1	
GRAY TRIGGERFISH		1	
GREATER AMBERJACK		1	
HOGFISH		1	
RED GROUPER			
RED PORGY		1	
SCAMP		1	
SNOWY GROUPER			
SPANISH MACKEREL			
WAHOO		1	
WHITE GRUNT			
YELLOWTAIL SNAPPER		*	
BLACK SEA BASS	Yes	N/A	N/A
GAG	Yes	N/A	N/A
KING MACKEREL	Yes	N/A	N/A
VERMILION SNAPPER	Yes	N/A	N/A

*Indicates a point in the landings stream did not exceed the threshold but were within 2,000 lbs of being classified as a spike.

The Council has recognized these spikes in the recreational data in the past and expressed concern over the validity of the estimates when such spikes occur. To avoid the spikes in the data causing accountability measures to be implemented, multi-year approaches such as the sum over three years or a geometric mean could be used to smooth out or average potential spikes over a series of years. It should also be added that dips would be affected as well. The sum of the most recent three years of catch compared to the annual catch limit for three years is essentially an average (arithmetic mean) if the ACL does not change. Because spikes would be combined with two other years of recent catch, the spike would be reduced. Therefore, an estimate using the sum of the three most recent years of landings may not trigger an accountability measure even if a single annual catch estimate exceeds the limit. The amount of the reduction will vary, but combining over three years of landings should theoretically get closer to the true average since there are more data points used to estimate the landings. Since it is unknown if the spike is a true estimate of catch, the sum of three years has an intermediate risk of overfishing among the alternatives. The risk of implementing an accountability measure when it is not needed is also moderate among the alternative approaches because the smoothed estimate reduces the influence of a single year. One issue with using the sum of three years is the assumption that the data have a normal distribution, meaning that it is just as likely that when a potential error occurs in the estimation of landings it occurs on both the low and high sides of the mean. Based on information described above, it is more likely that spikes occur and the data are not likely normally distributed. The geometric mean (the cubic root of the product of three landings estimates in the case being examined) is an averaging technique used to estimate the likely average when data are skewed on the high side of the mean. Because the geometric mean is designed to reduce the influence of high values, it reduces the spike more than the average and is lower than the average described above in all situations. This results in the geometric mean having the highest risk of overfishing among the alternatives if the spike in the data is a true observation but lowest risk of implementing accountability measures if the spike in the data is due to random error. It is worth noting that multi-year approaches could result in triggering an AM over multiple years if the spike is large enough to drive the sum or geometric mean over the trigger level during the entire three-year period it remains part of the calculation. This is one reason even multi-year approaches may not function as hard and fast rules to simply set and forget.

To help display the points above, accountability measure triggers were developed for recreational landings of dolphin and wahoo from 2010 to 2019 based on the different methods. In this example, the point value is that year's landings. The average and geometric means were calculated with the point value and two years prior to the year (e.g., 2010 would include 2008, 2009, and 2010). The recreational dolphin landings typically remained between 11,000,000 and 17,000,000 million pounds (**Figure 3**). The 2015 landings estimate was approximately 8,000,000 pounds higher or 52% to 127% greater than any other point in the time series. The increase in 2015 over the all the other years results in the data distribution being non-normal (Shapiro-Wilk normality test p=0.0291). In this dolphin example, catch returned to below the annual catch limit without management measures. Since the effort (targeted trips) in the fishery (**Figure 1**) and landings per angler in 2015 (**Figure 2**) did not shift substantially, it seems unlikely that estimated landings should have increased by 50% over the preceding and following years. Therefore, point estimate accountability measures, which would have been triggered due to the annual catch limit being exceeded by 11%, would have been unnecessary. Both the average and geometric mean would have prevented a need to implement an accountability measure.

The landings data for the wahoo fishery had three years in a row when the estimated catch exceeded the potential new annual catch limit (**Figure 4**). The second highest effort occurred in 2016 when the landings spiked and along with the 2015 landings estimate being higher than the annual catch limit; therefore, evidence would have indicated a potential need to consider management. Based on this scenario, a point estimate accountability measure would have been triggered for the fishing years 2016, 2017, and 2018 and the catch exceeded the potential annual catch limit by 5%, 79%, and 28%, respectively. An almost 100% increase in the landings over the previous year (2016 compared to 2015) is unlikely for this fishery because it is a highly specialized fishery and recreational fishermen are limited to 2 fish per person. The landings per angler did not show a noticeable change in the number of wahoo kept per angler, just an increase in number of targeted trips.

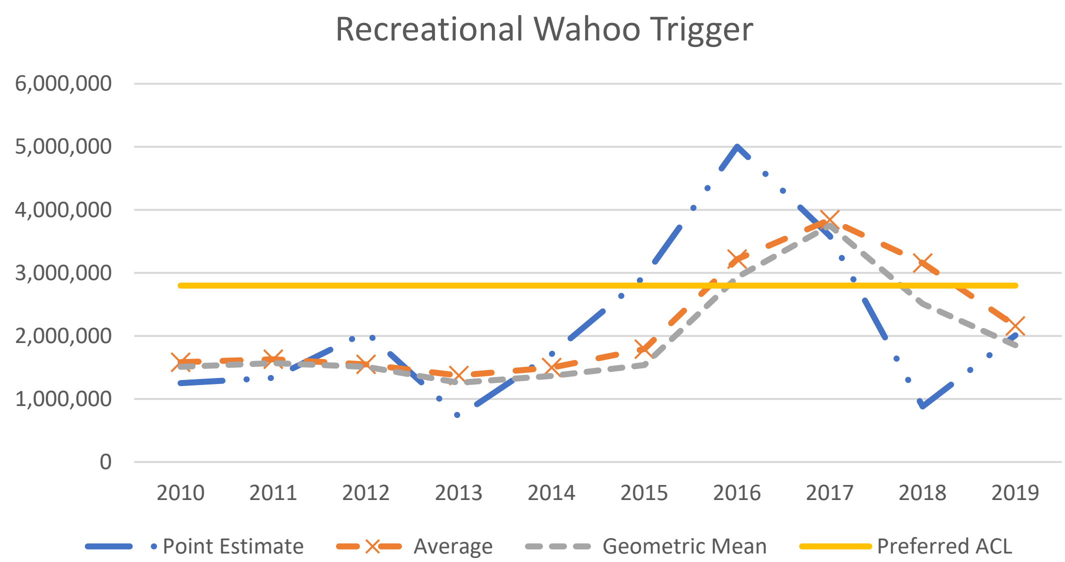
Both smoothing methods would have resulted in accountability measures beginning in 2017 after the spike in 2016. The average would have triggered accountability measures in 2017, 2018, and 2019, a year later than the point estimates, and the annual catch limit was exceeded by a smaller amount compared to the point estimate (15%, 37%, and 13%). The geometric mean would have resulted in the trigger being met in only two years (2017 and 2018, 5% and 34% over the annual catch limit).

It is important to consider the distribution of the data when determining which method is more appropriate to describe the central tendency of the estimated landings. The wahoo data did not appear to depart from normality (Shapiro-Wilk normality test $p=0.2042$). However, the distribution was skewed (skewness = 1.17) and log-transformation resulted in a more normal distribution (**Figure 5**). This indicates that geometric mean is the more appropriate method to smooth the data to describe central tendency particularly when catches spike.

The above examples just look a short time period to consider when different methods would have been triggered in the past for the recreational sector of the dolphin and wahoo fisheries. In conclusion, the Council is considering alternative methods of evaluating landings against accountability measures due to concerns about the reliability of the recreational catch estimates. Such issues are not being evaluated for the commercial sector accountability measures because the mandatory, census-style reporting by dealers and harvesters results in a more accurate and reliable data series. When examining triggers for recreational accountability measures, the Council should consider balancing its risk tolerance for potentially allowing overfishing with its desire to avoid imposing unnecessary, temporary, restrictive management measures due to skewed, outlier, or uncertain estimates. There are stock consequences to overfishing, and social and economic consequences to restrictive management measures.

Multi-year approaches tend to have a higher risk of allowing overfishing if data are accurate and precise, but they can help mitigate anomalies in the data when they are imprecise and, particularly spiky. Of the alternatives the Council is considering in DW 10 to trigger the recreational accountability measure, a geometric mean has the highest risk of allowing overfishing when a high catch estimate is legitimate, while providing the greatest likelihood of mitigating outliers and addressing issues with distribution of the data. A multi-year sum or arithmetic mean is likely a less risky approach in regard to overfishing but riskier in terms of implementing accountability measures unnecessarily. Single year approaches could be more likely to reduce the risk of overfishing when data are accurate and reliable, but do not address anomalies in the recreational data and assume a normal distribution, which, as exhibited in the landings data examined, is not always the case. The Council may wish to consider past performance of the catch estimate for a particular species before deciding how to balance the risk of overfishing with the risk of unnecessary regulatory changes. This could include the PSE of the estimates, their distribution and trends over the time series, and the frequency of spiky estimates as illustrated here. In the end, the Council can examine each species on a case by case basis to determine the most appropriate approach for triggering recreational accountability measures or develop a protocol to review recreational data prior to implementing recreational accountability measures.

Figure 3. The recreational dolphin annual accountability measure trigger values for a point estimate,



three-year average (defined as arithmetic mean), three-year geometric mean, and the Preferred sector Annual Catch Limit (ACL).

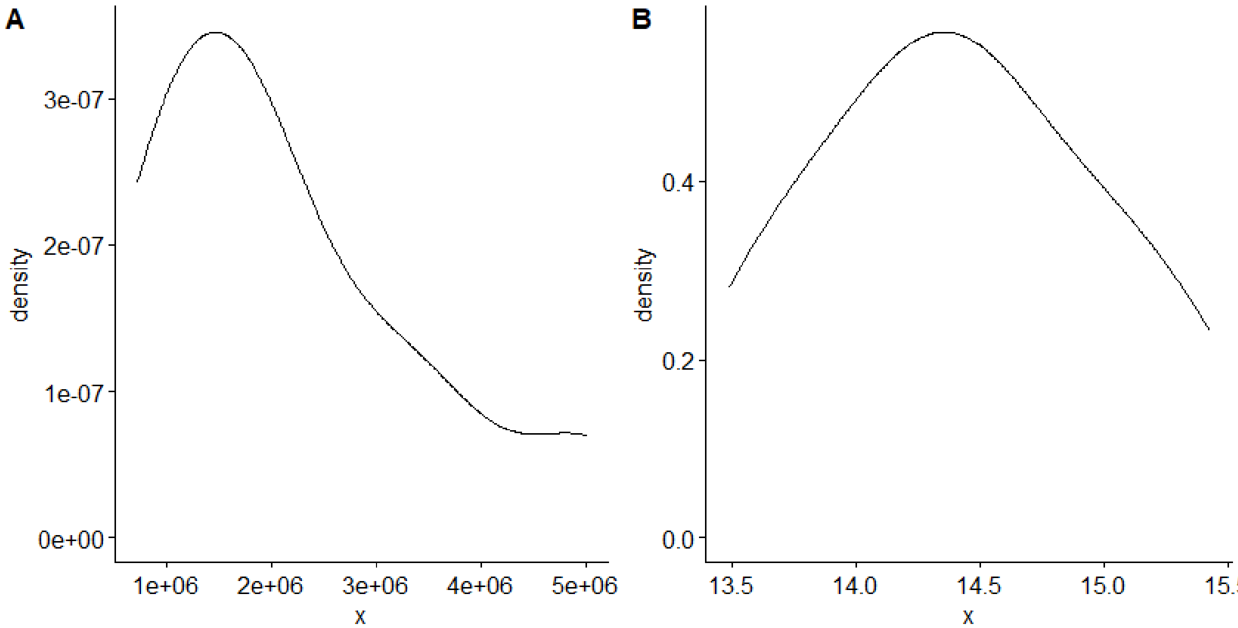


Figure 4. The recreational wahoo annual accountability measure trigger values for a point estimate, three-year average (defined as arithmetic mean), three-year geometric mean, and an average of Annual Catch Limits (ACL) (no preferred selected).

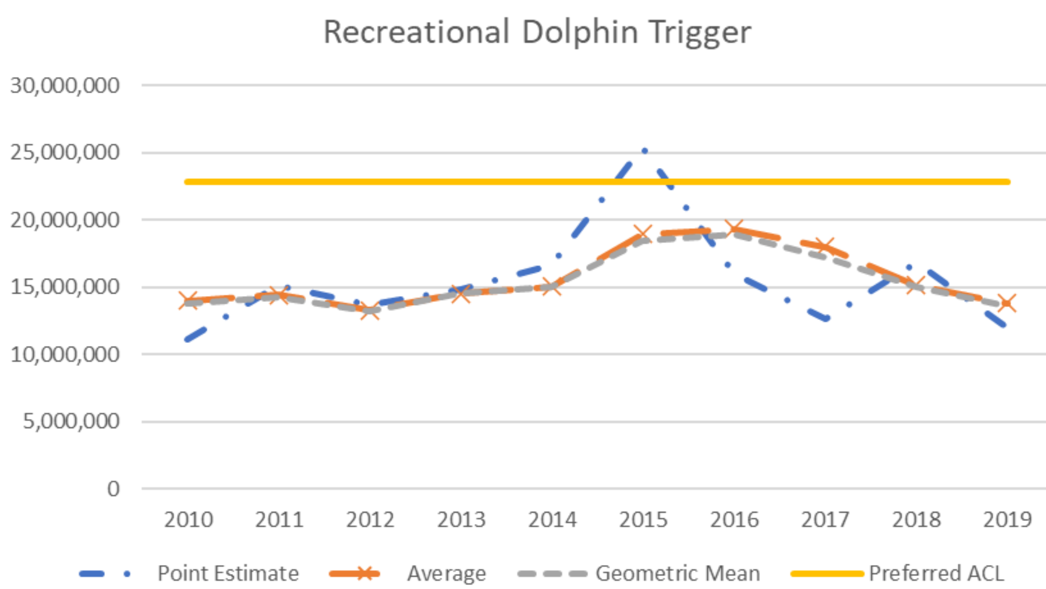


Figure 5. Wahoo landings density distribution (A) and log-transformed wahoo landings density distribution (B) for the South Atlantic region from 2010 to 2019.