MRFSS to MRIP Adjustment Ratios and Weight Estimation Procedures for South Atlantic and Gulf of Mexico Managed Species

Vivian M. Matter and Adyan Rios

SEDAR32-DW-02*

5 February 2013

*There is a corresponding spreadsheet that accompanies this working paper titled, "SEDAR32_DW02_Matter&Rios_mrip_adj_ratios".

Please cite this document as:

Matter, V.M. and A. Rios. 2013. MRFSS to MRIP Adjustment Ratios and Weight Estimation Procedures for South Atlantic and Gulf of Mexico Managed Species. SEDAR32-DW02. SEDAR, North Charleston, SC. 6 pp.

MRFSS to MRIP Adjustment Ratios and Weight Estimation Procedures for South Atlantic and Gulf of Mexico Managed Species

by Vivian M. Matter[1] and Adyan Rios[2]

NOAA Fisheries
Southeast Fisheries Science Center
[1]Fisheries Statistics Division
[2]Sustainable Fisheries Division
75 Virginia Beach Drive
Miami FL 33149

February 5, 2013
Sustainable Fisheries Division Contribution No. SFD-2013-006

Introduction

The Marine Recreational Information Program (MRIP) was implemented to provide more detailed, timely, and reliable estimates of marine recreational fishing catch and effort. MRIP's independently peer reviewed methodology replaced the Marine Recreational Fisheries Statistics Survey (MRFSS) program that had operated since 1981 (Visit http://www.st.nmfs.noaa.gov/st1/recreational/ for more information on MRIP and MRFSS). Official MRIP estimates are available from 2004 to 2011 and represent the best available data for use in stock assessments. To expand the MRIP time series to include years prior to 2004, we re-estimated catch and adjusted variance for years with MRFSS data (Rios 2012). In addition, we developed a standardized approach for obtaining weight estimates throughout the recreational time series.

MRFSS to MRIP Adjustment

Adjustment ratios were first calculated for Gulf of Mexico red snapper estimates in SEDAR 31. The methods and procedures in SEDAR31-DW25 (Rios et al 2012) were used to develop adjustment ratios for all South Atlantic and Gulf of Mexico managed species. Additional considerations and procedures are described here

Methods

Hierarchy rules

We attempted to calculate ratio estimators that were species, region, and mode specific. When ratio estimators and variances of ratio estimators were not available from the species-region-mode specific data, we calculated them from species-region specific data, and when necessary from only species-specific data.

The MRFSS-MRIP adjustment requires values for the following 6 variables:
1) the ratio estimator of the ab1 estimate (re_ab1)
2) the ratio estimator of the b2 estimate (re_b2)
3) the ratio estimator of the variance of the ab1 estimate (re_varab1)
4) the ratio estimator of the variance of the b2 estimate (re_varb2)
5) the variance of re_ab1 (varR_ab1)
6) the variance of re_b2 (varR_b2)

A matrix of possible variable outcomes, where 1 indicates that a value is obtained and 0 indicates that a value is missing, is shown in Table 1. The matrix focuses on scenarios where we are missing ratio estimates in numbers or variances of those ratios (variables 1, 2, 5, and 6). The outcomes of variables 3 and 4 (ratios of variance) are not included since they mirror those of variables 1 and 2 (ratios of numbers), respectively. To substitute or re-estimate missing ratio estimates in numbers or missing variances of the ratios, we followed the steps outlined in Table 1.

Charterboat/Headboat mode
Before applying mode-specific ratio estimators to data from the Gulf of Mexico and South Atlantic regions, we separated the 1981-1985 MRFSS combined charterboat/headboat mode following the methods described in SEDAR31-DW-25. In the Mid-Atlantic and North Atlantic regions the charterboat and headboat modes were combined in the MRFSS survey from 1981 to 2003. In 2004, MRFSS began calculating estimates for these modes separately. In these regions we combined the charter and headboat modes from MRFSS and MRIP between 2004 and 2011 and then calculated ratio estimators for the combined charterboat/headboat mode.

Rounding
In contrast to the methods used by Rios et al. (2012), which included rounding the ratio estimators to three decimal places and the variances of ratio estimators to six decimal places, we did not round ratio estimators or their associated variances.

Results

When an ab1 or b2 ratio was based on only a single year of non-zero data, the ratio did not have an associated variance. When possible, we substituted the missing variance with that of a corresponding b2 or ab1 ratio estimator. However, when variances were missing for both, the estimates could not be fully adjusted and the associated variances were flagged. The AB1 estimates with flagged variances accounted for 0.01% of the landings in 1981-2003. The B2 estimates with flagged variances accounted for 0.0001% of the discards in 1981-2003.

Ratio estimators could not be calculated for all species. The following South Atlantic and Gulf of Mexico managed species do not have a ratio estimator for the AB1 landings: blackline tilefish, Nassau grouper, and smallmouth grunt. These species' landings (1981-2003) make up 0.06% of the total landings from South Atlantic and Gulf of Mexico managed species between 1981 and 2003. The following South Atlantic and Gulf of Mexico managed species do not have a ratio estimator for the B2 discards: black snapper, cottonwick, dwarf sand perch, misty grouper, temperate bass genus, tiger grouper, and saucereye porgy. These species' discards (1981-2003) make up 0.03% of the total discards from South Atlantic and Gulf of Mexico managed species between 1981 and 2003.

Ratio estimators of the landings, ratio estimators of the variances, and variances associated with the ratio estimators of the landings for all South Atlantic and Gulf of Mexico managed species are available in the accompanying spreadsheet (mrip_adj_ratios.xlsx).

Weight Estimation Procedure

The MRFSS and the MRIP surveys use different methodologies to estimate landings in weight. To apply a consistent methodology over the entire recreational time series, the Southeast Fisheries Science Center (SEFSC) implemented a method for calculating average weights for the MRIP (and MRIP adjusted) landings. The SEFSC has used this method in the past for substituting missing weight estimates in the MRFSS data (i.e. when sample data in a given strata include an estimate of landings in number but not in weight due to missing weight samples).

The SEFSC method obtains average weights by aggregating MRFSS/MRIP data according to the following hierarchy: species, region, year, state, mode, wave, and area (SEDAR22-DW16). The minimum number of weights required at each hierarchy level is 30 fish, except at the final species level, where the minimum is 1 fish. When sample data include lengths without associated weights, weights are estimated from length-weight equations. Average weights are multiplied by the landings estimates in number to obtain estimates of landings in weight. These estimates are provided in pounds, whole weight.

The Office of Science and Technology, the Northeast Fisheries Science Center, and the Southeast Fisheries Science Center expect to work together in the future to establish a standard method to use in stock assessments, management, and in the MRIP survey.

References

Rios, A., V. Matter, J. Walter, N. Farmer, and S. Turner. 2012. SEDAR31-DW25. Estimated Conversion Factors for Adjusting MRFSS Gulf of Mexico Red Snapper Catch Estimates and Variances in 1981-2003 to MRIP Estimates and Variances. National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division, Fisheries Statistics Division, Miami, FL

Matter, V. and S. Turner. 2010. SEDAR22-DW16. Estimated Recreational Catch in Weight: Method for Filling in Missing Weight Estimates from the Recreational Surveys with Application to Yellowedge Grouper, Tilefish (golden), and Blueline Tilefish. National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division.

Table 1. Matrix of outcomes for the variables used in the MRFSS-MRIP adjustment (re_ab1: the ratio estimator of the ab1 estimate; re_b2: the ratio estimator of the b2 estimate; varR_ab1: the variance of re_ab1; varR_b2: the variance of re_b2). 1 indicates that a value is obtained and 0 indicates that a value is missing. Highlighted cells represent ratio estimators that could not be calculated.

Hierarchy level	re_ab1	re_b2	varR_ab1	varR_b2	next steps
species, region, mode	1	1	1	1	
	1	1	1	0	borrow varR_b2 from varR_ab1
	1	1	0	1	borrow varR_ab1 from varR_b2
	1	0	1	0	move b2 up hierarchy (all modes collapsed)
	0	1	0	1	move ab1 up hierarchy (all modes collapsed)
	1	1	0	0	no variances associated with ratio estimators- move up hierarchy (all modes collapsed)
	1	0	0	0	no variances associated with ratio estimators- move up hierarchy (all modes collapsed)
	0	1	0	0	no variances associated with ratio estimators- move up hierarchy (all modes collapsed)
	0	0	0	0	move up hierarchy (all modes collapsed)
species, region	1	1	1	1	
	1	1	1	0	borrow varR_b2 from varR_ab1
	1	1	0	1	borrow varR_ab1 from varR_b2
	1	0	1	0	move b2 up hierarchy (all modes and regions collapsed)
	0	1	0	1	move ab1 up hierarchy (all modes and regions collapsed)
	1	1	0	0	no variances associated with ratio estimators- move up hierarchy (all modes and regions collapsed)
	1	0	0	0	no variances associated with ratio estimators- move up hierarchy (all modes and regions collapsed)
	0	1	0	0	no variances associated with ratio estimators- move up hierarchy (all modes and regions collapsed)
	0	0	0	0	move up hierarchy (all modes and regions collapsed)
species	1	1	1	1	
	1	1	1	0	borrow varR_b2 from varR_ab1
	1	1	0	1	borrow varR_ab1 from varR_b2
	1	0	1	0	flag b2
	0	1	0	1	flag ab1
	1	1	0	0	no variances associated with ratio estimators- use ab1 and b2 ratio estimators; variance will not be adjusted fully (e.g. adjust variance using re_varab1 but not with varR_ab1); flag resulting variances
	1	0	0	0	flag b2; no variances associated with ratio estimators- use ab1 ratio estimator; variance will not be adjusted fully (e.g. adjust variance using re_varab1 but not with varR_ab1); flag resulting variance
	0	1	0	0	flag ab1; no variances associated with ratio estimators- use b2 ratio estimator; variance will not be adjusted fully (e.g. adjust variance using re_varb2 but not with varR_b2); flag resulting variance
	0	0	0	0	flag ab1 and b2

SEDAR 22-DW-16

Estimated Recreational Catch in Weight: Method for Filling in Missing Weight Estimates from the Recreational Surveys with Application to Yellowedge Grouper, Tilefish (golden), and Blueline Tilefish

by Vivian M. Matter and Stephen C. Turner

NOAA Fisheries
Southeast Fisheries Science Center
Sustainable Fisheries Division
75 Virginia Beach Drive
Miami FL 33149

March 15, 2010
Sustainable Fisheries Division Contribution No. SFD-2010-003

INTRODUCTION

Estimates of recreational catch in for marine fish species in the Gulf of Mexico beginning in 1981 are obtained by a combination of results from three surveys:

- the Marine Recreational Fishery Statistics Survey (MRFSS) conducted by the NOAA Fisheries (NMFS).
- the Texas Marine Sport-Harvest Monitoring Program by the Texas Parks and Wildlife Department (TPWD).
- the Headboat Survey (HBS) conducted by NMFS, Southeast Fisheries Science Center, Beaufort, NC.

Landings estimates are provided in numbers of fish from all surveys. Estimates of landings (A+B1 for MRFSS) in weight from the recreational surveys have typically not been used due to incompleteness. The TPWD survey does not provide estimates of catch in weight. The HBS and MRFSS do provide estimates of landings in weight. However, the MRFSS estimates must be treated with caution due to the occurrence of missing weight estimates in some strata. MRFSS weight estimates are calculated by multiplying the estimated number harvested in a cell (year/wave/state/mode/area/species) by the mean weight of the measured fish in that cell. When there are no fish measured in the cell (fish were gutted or too big for the sampler to weigh, harvest was all self-reported, etc) estimates of landings in number are provided but there are no corresponding estimates of landings in weight.

Due to these limitations in the weight estimates provided by the recreational surveys, landings estimates have typically always been provided in numbers of fish. However, management measures oftentimes require estimates in weight. In the past, the SEDAR process has calculated estimates of recreational landings in weight often using procedures developed by assessment scientists which might vary from species to species and assessment to assessment. The following is a proposed standardized method for filling in these missing weight estimates in the recreational data that can be applied to all species on a regular basis.

SUBSTITUTION SCHEMA

Sample Data

The intercept data from the MRFSS is compiled for all managed species across all years available (1981+) and for the entire Atlantic coast and Gulf of Mexico (sub regions 4-7). For the state of Texas two methods are used to compile weights. The TPWD survey provides only length measurements of fish in their sample data; weights are not recorded. The first and preferred method of obtaining weights is to convert lengths from the TPWD intercept data into weights using SEDAR endorsed length-weight equations. Since these conversions need to be done at a species specific level, only a small number of species have been calculated using this method at this time. The rest of the species' weights have been obtained using the second method, which is to use MRFSS weights from the Louisiana intercept data. The ultimate goal is to obtain all Texas weights using the length-weight equation approach for all species.

The sample data from the MRFSS and Texas (using either method 1 or 2, depending on the species) is then compiled into one datafile.

Estimating landings in weight

HBS provides estimates of landings in weight, so no substitutions are required. The MRFSS estimates of landings in weight are used when provided by the survey. In cases where there is an estimate of landings in number but not weight, the Southeast Fisheries Science Center has used the sample data discussed above to obtain an average weight using the following hierarchy: species, region, year, state, mode, and wave. The minimum number of weights used at each level of substitution is 30 fish, except for the final species level, where the minimum is 1 fish. For the TPWD survey average weights are calculated from the TPWD length samples using the same hierarchy as MRFSS except "area" is added to the finest level of substitution (species, region, year, state, mode, wave, and area). If there are not 30 fish size observations at the finest level, then the number of samples across areas in the same wave are examined for sufficient sample size. If there are not sufficient samples with the state (across modes, waves and areas) then regional information across states within the year are examined; this and later steps include size observations from both the TPWD survey and MRFSS.

Average weights are then multiplied by the landings estimates in number to obtain estimates of landings in weight. These estimates are provided in pounds whole weight (lbsest_SECwwt). Weight estimates for managed groupers and tilefish are also provided in pounds gutted weight (lbsest_SECgwt). The level of substitution used is recorded in the data file provided to user in the variable lbsest_SECsource, which has the following possible values and meanings:

Variable	Value	Definition
lbsest_SECwwt		estimated whole weight of landings (type A+B1) in pounds
lbsest_SECgwt		estimated gutted weight of landings (type A+B1) in pounds; available for red grouper, gag, black grouper, scamp, dwarf sand perch, sand perch, red hind, rock hind, yellowfin grouper, yellowmouth grouper, yellowedge grouper, warsaw grouper, snowy grouper, speckled hind, misty grouper, golden tilefish, anchor tilefish, blackline tilefish, goldface tilefish, blueline tilefish, queen snapper, and wenchman
lbsest_SECsource	MRFSSest	no substitution made; weight estimate as reported by MRFSS
	HBSest	no substitution made; weight estimate as reported by HBS
	srysmwa	average weight from intercept data by species, region, year, state, mode, wave, and area; minimum number of weights used is 30; used only for TPWD survey as first strata
	srysmw	average weight from intercept data by species, region, year, state, mode, and wave; minimum number of weights used is 30
	srysm	average weight from intercept data by species, region, year, state, and mode; minimum number of weights used is 30
	srys	average weight from intercept data by species, region, year, and state; minimum number of weights used is 30
	sry	average weight from intercept data by species, region and year; minimum number of weights used is 30
	sr	average weight from intercept data by species and region; minimum number of weights used is 30
	s	average weight from intercept data by species; minimum number of weights used is 1

LANDINGS ESTIMATES

Landings estimates in weight for yellowedge grouper, tilefish (golden), and blueline tilefish from recreational fisheries surveys in the Gulf of Mexico are presented in Tables 1-6; in many years there were no fish estimated to have been landed and those years are not shown in the tables. Weight estimates are provided for A+B1 landings using the methods discussed above. In Tables 1-3 the landings estimates in weight are provided by survey for each species. In Tables 4-6 the landings estimates in weight are presented by how the estimates were derived; as described above all HBS estimates were derived from the headboat survey. Using Table 4 (yellowedge grouper) as an example, one can see that in some years (2004) MRFSS estimates of landed weight were available for all strata and thus no substitution is necessary. In other years (2007), weight estimates are not available from the MRFSS from any strata and were derived entirely from a multi-year substitution at the species-region level. Still in other years (2005) MRFSS weight estimates were available in some starta, while in other strata a multi-year substitution was needed.

Tables 4-6 show that for yellowedge grouper and blueline tilefish species-region substitutions were used when MRFSS and TPWD weight estimates were not available; this indicates that in any given year there were never more than 29 fish of either of those species available. The golden tilefish appears to be rarer in the recreational fishery as shown by the fewer strata in which estimates occurred and by the fact that a species level average rather than a higher level average weight was used; this indicates that between 1981 and 2009 less than 30 golden tilefish had been measured in MRFSS and TPWD surveys.

DISCUSSION

The MRFSS uses a limited susbstitution scheme for average weights which is different from the one proposed in this paper. The scheme proposed in this paper follows a progression of eliminating on stratum at a time from most disaggregated to most aggregated: the progression is species, region, year, state, mode, and wave with species being the most aggregated and wave being the most disaggregated. If there is not already a MRFFS estimate of weight in a stratum , then substitution is used; for a stratum to be used for substitution there has to be at least 30 fish weights available from that stratum). The MRFSS uses a different approach and sets lower minimum numbers of observations (two fish). For the official MRFSS, if a cell (species/year/wave/state/mode/area) is missing a mean weight, then a state-wide average is used if there are at least two fish measured in the state (all fishing areas and modes combined). If there are not at least two fish at the state level then the subregion (all fishing areas, modes, and states combined) is used if possible. If there are not at least two measured fish at the subregion level no average weight is used and no weight is estimated and MRFSS leaves a missing weight estimate.

The proposed substitution scheme (species, region, year, state, mode, and wave) would benefit from analysis of patterns in average weights across strata to determine if alternative patterns might result in calculated average weights closer to the true average for the specific sampling stratum.

The MRFSS procedures allow the use of a relatively small number of fish in a stratum (two or more) for calculating an average weight for use in estimating the weight of the landings. There is concern that those small sample sizes might result in highly variable estimates of landed weight. It would be sensible to examine the impact of sample size on precision and accuracy of the calculated yield and consider adding the ability to replace MRFSS estimated landing weights in strata with small sample sizes with estimates from more aggregated strata with at least 30 observed sizes

---

General overview of the recreational surveys from the following:
Recreational Survey Data for Gag and Black Grouper in the Gulf of Mexico. Patty Phares, Vivian Matter, and Steve Turner. National Marine Fisheries Service, Southeast Fisheries Science Center, Sustainable Fisheries Division, January, 2006. Sustainable Fisheries Division Contribution No. SFD-2006-008.

Table 1. Estimated landings of fish (A+B1) in pounds whole weight and gutted weight by source survey for **yellowedge grouper** in the Gulf of Mexico.

	HBS		MRFSS		TPWD		Total	
YEAR	Whole weight	Gutted weight	Whole weight	Gutted Weight	Whole weight	Gutted weight	Whole weight	Gutted weight
1982			166,472	159,471			166,472	159,471
1984					218	209	218	209
1986	478	457	0	0	456	437	934	895
1987	1,152	1,103	11,064	10,599			12,216	11,702
1988	2,274	2,178					2,274	2,178
1989	766	734	17,289	16,562			18,055	17,296
1990	1,715	1,643					1,715	1,643
1991	1,390	1,331	4,633	4,438			6,023	5,769
1992	510	489					510	489
1993	347	333	5,069	4,856			5,417	5,189
1994	442	423	0	0			442	423
1995	632	605					632	605
1996	188	180	0	0			188	180
1997	386	369	2,410	2,308			2,795	2,678
1998	465	445	7,791	7,463			8,256	7,909
1999	56	53	1,028	985			1,084	1,038
2000	39	37	0	0			39	37
2001	52	50	1,433	1,373			1,485	1,422
2002	30	29	3,975	3,808			4,005	3,837
2003	95	91	401	384			496	475
2004	72	69	1,193	1,143			1,264	1,211
2005	148	142	59,357	56,861			59,506	57,003
2006	216	207	2,680	2,568			2,897	2,775
2007	211	202	1,207	1,156			1,418	1,358
2008	211	202	1,244	1,191			1,455	1,394
2009			5,920	5,671			5,920	5,671
Grand Total	**11,874**	**11,375**	**293,166**	**280,837**	**674**	**646**	**305,714**	**292,858**

Table 2. Estimated landings of fish (A+B1) in pounds whole weight and gutted weight by source survey for **tilefish (golden)** in the Gulf of Mexico.

	HBS		MRFSS		Total	
YEAR	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight
1981			179,080	159,893	179,080	159,893
1987			17,944	16,022	17,944	16,022
1990			4,419	3,946	4,419	3,946
1992	3	3	3,336	2,978	3,339	2,981
1995	2	2			2	2
1998	6	6			6	6
2000			197	176	197	176
2001	1	1	137	122	138	123
2005			5,453	4,869	5,453	4,869
2006			0	0	0	0
2008			216	193	216	193
Grand Total	**13**	**11**	**210,783**	**188,199**	**210,796**	**188,211**

Table 3. Estimated landings of fish (A+B1) in pounds whole weight and gutted weight by source survey for **blueline tilefish** in the Gulf of Mexico.

	HBS		MRFSS		Total	
YEAR	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight
1986	281	251			281	251
1987	671	599	2,739	2,446	3,410	3,045
1988	1,013	904			1,013	904
1989	678	605			678	605
1990	1,400	1,250			1,400	1,250
1991	462	412	0	0	462	412
1992	4	4			4	4
1993	78	70	3,706	3,309	3,784	3,379
1994	56	50			56	50
1995	18	16			18	16
1996	71	63			71	63
1997	28	25	669	598	697	622
1998	6	6			6	6
1999	5	5	3,480	3,108	3,486	3,112
2000	60	53	221	198	281	251
2001	11	10	639	571	650	580
2002	127	114	116	103	243	217
2003	32	29	4,084	3,646	4,116	3,675
2004	22	20	4,509	4,026	4,531	4,046
2005	74	66	2,283	2,038	2,357	2,104
2006	7	6	920	821	927	828
2007			11,580	10,339	11,580	10,339
2008	53	47	28,030	25,027	28,083	25,074
2009			17,696	15,800	17,696	15,800
Grand Total	**5,157**	**4,604**	**80,672**	**72,029**	**85,829**	**76,633**

Table 4. Estimated landings of fish (A+B1) in pounds whole weight and gutted weight by substitution level for **yellowedge grouper** in the Gulf of Mexico.

	HBSest		MRFSSest		Sr		Total	
YEAR	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight
1982					166,472	159,471	166,472	159,471
1984					218	209	218	209
1986	478	457			456	437	934	895
1987	1,152	1,103	11,064	10,599			12,216	11,702
1988	2,274	2,178					2,274	2,178
1989	766	734			17,289	16,562	18,055	17,296
1990	1,715	1,643					1,715	1,643
1991	1,390	1,331	4,633	4,438			6,023	5,769
1992	510	489					510	489
1993	347	333	340	325	4,730	4,531	5,417	5,189
1994	442	423					442	423
1995	632	605					632	605
1996	188	180					188	180
1997	386	369	2,410	2,308			2,795	2,678
1998	465	445	7,791	7,463			8,256	7,909
1999	56	53	1,028	985			1,084	1,038
2000	39	37					39	37
2001	52	50	1,433	1,373			1,485	1,422
2002	30	29	109	104	3,866	3,703	4,005	3,837
2003	95	91	401	384			496	475
2004	72	69	1,193	1,143			1,264	1,211
2005	148	142	58,938	56,460	419	402	59,506	57,003
2006	216	207	2,680	2,568			2,897	2,775
2007	211	202			1,207	1,156	1,418	1,358
2008	211	202			1,244	1,191	1,455	1,394
2009			586	562	5,334	5,109	5,920	5,671
Grand Total	**11,874**	**11,375**	**92,607**	**88,712**	**201,233**	**192,771**	**305,714**	**292,858**

Table 5. Estimated landings of fish (A+B1) in pounds whole weight and gutted weight by substitution level for **tilefish (golden)** in the Gulf of Mexico.

	HBSest		MRFSSest		S		Total	
YEAR	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight
1981			179,080	159,893			179,080	159,893
1987			17,944	16,022			17,944	16,022
1990			4,419	3,946			4,419	3,946
1992	3	3			3,336	2,978	3,339	2,981
1995	2	2					2	2
1998	6	6					6	6
2000			197	176			197	176
2001	1	1	137	122			138	123
2005			5,453	4,869			5,453	4,869
2008			85	76	131	117	216	193
Grand Total	**13**	**11**	**207,316**	**185,104**	**3,467**	**3,095**	**210,796**	**188,211**

Table 6. Estimated landings of fish (A+B1) in pounds whole weight and gutted weight by substitution level for **blueline tilefish** in the Gulf of Mexico

	HBSest		MRFSSest		Sr		Total	
YEAR	Whole Weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight	Whole weight	Gutted weight
1986	281	251					281	251
1987	671	599	221	197	2,518	2,249	3,410	3,045
1988	1,013	904					1,013	904
1989	678	605					678	605
1990	1,400	1,250					1,400	1,250
1991	462	412					462	412
1992	4	4					4	4
1993	78	70			3,706	3,309	3,784	3,379
1994	56	50					56	50
1995	18	16					18	16
1996	71	63					71	63
1997	28	25	669	598			697	622
1998	6	6					6	6
1999	5	5	3,480	3,108			3,486	3,112
2000	60	53	221	198			281	251
2001	11	10	136	122	503	449	650	580
2002	127	114	116	103			243	217
2003	32	29	4,084	3,646			4,116	3,675
2004	22	20	1,479	1,320	3,030	2,705	4,531	4,046
2005	74	66	1,229	1,097	1,054	941	2,357	2,104
2006	7	6	920	821			927	828
2007			11,580	10,339			11,580	10,339
2008	53	47	4,290	3,831	23,740	21,196	28,083	25,074
2009			17,438	15,570	258	230	17,696	15,800
Grand Total	**5,157**	**4,604**	**45,864**	**40,950**	**34,808**	**31,078**	**85,829**	**76,633**

SEDAR67-DW-xx

Sample size sensitivity analysis for calculating MRIP weight estimates

Kyle Dettloff and Vivian Matter

National Marine Fisheries Service
Southeast Fisheries Science Center

Fisheries Statistics Division
75 Virginia Beach Drive
Miami, FL 33149

August 15, 2019

INTRODUCTION

The Marine Recreational Information Program (MRIP), formerly the Marine Recreational Fishery Statistics Survey (MRFSS), was implemented in 1981 to provide regional based catch and effort estimates of marine finfish in United States recreational fisheries. The survey provides species specific catch estimates by stratum (species/year/wave/sub-region/state/mode/area) in numbers of fish, however corresponding weight measurements are not always available due to sampling constraints or incomplete self-reporting. Therefore, recreational landings estimates were historically provided in numbers of fish for stock assessments. When management measures, such as ACL monitoring, began requiring estimates in weight on a routine basis, a standard methodology was developed to estimate missing recreational weights (Matter and Turner, 2010). This method follows a hierarchy that requires a sample size of at least 30 fish to calculate mean weight at the stratum level, sequentially pooling samples from coarser strata until this number is achieved. The calculated mean weight is then multiplied by the estimated number of fish in a cell to calculate the missing total weight.

In 2012, as part of the MRIP re-estimation project, MRIP developed a new weight estimation methodology that was applied to weight estimates back to 2004. A mix of hot and cold deck imputation and length-weight modeling was used to fill in missing weight estimates (NOAA Fisheries). However, this still left some missing weight estimates at the stratum level. In order to provide consistent weight estimates throughout the time series, SEFSC implemented the weight estimation procedure, previously used to fill in missing weights, to produce weight estimates for all strata back to 1981 (Matter and Rios, 2013). This methodology, which uses a minimum sample size of 30 fish to calculate an average weight, was considered an improvement over the previous survey method which required weights of only two fish to calculate stratum level weight estimates. The impact of requiring 30 samples, however, has never been formally assessed.

This paper formally examines the stability of weight estimates as a function of sample size for a variety of ACL managed species. Based on this analysis, a reasonable sample size threshold can be determined at which the precision of the estimates stabilizes and further samples result in diminishing improvement. Requiring a sample size no larger than this cutoff will ensure more accurate weight estimates within strata as potentially less aggregation of samples across coarser levels will be necessary to meet the minimum sample size threshold.

METHODS

A non-parametric bootstrapping approach was used to calculate the variability in mean stratum-level weight estimates using random samples drawn from 2018 MRIP source weights (all sub-regions) of size one to thirty. Only ACL managed species with at least 30 recorded weights in a particular stratum were included in the analysis. At each stratum level, 50,000 bootstrap replicates were run for each sample size (i.e., the mean weight was calculated for 50,000 random samples of size one to thirty drawn with replacement, and the standard deviation of those 50,000 means was taken to obtain the precision of the estimate at each sample size).

$$SE_n = std((\frac{1}{n}\sum_{i=1}^{n} X_i)_j)$$

$$n = 1, \dots, 30; j = 1, \dots, reps\ (50{,}000)$$

$$X = random\ stratum\ level\ fish\ weight\ selected\ with\ replacement$$

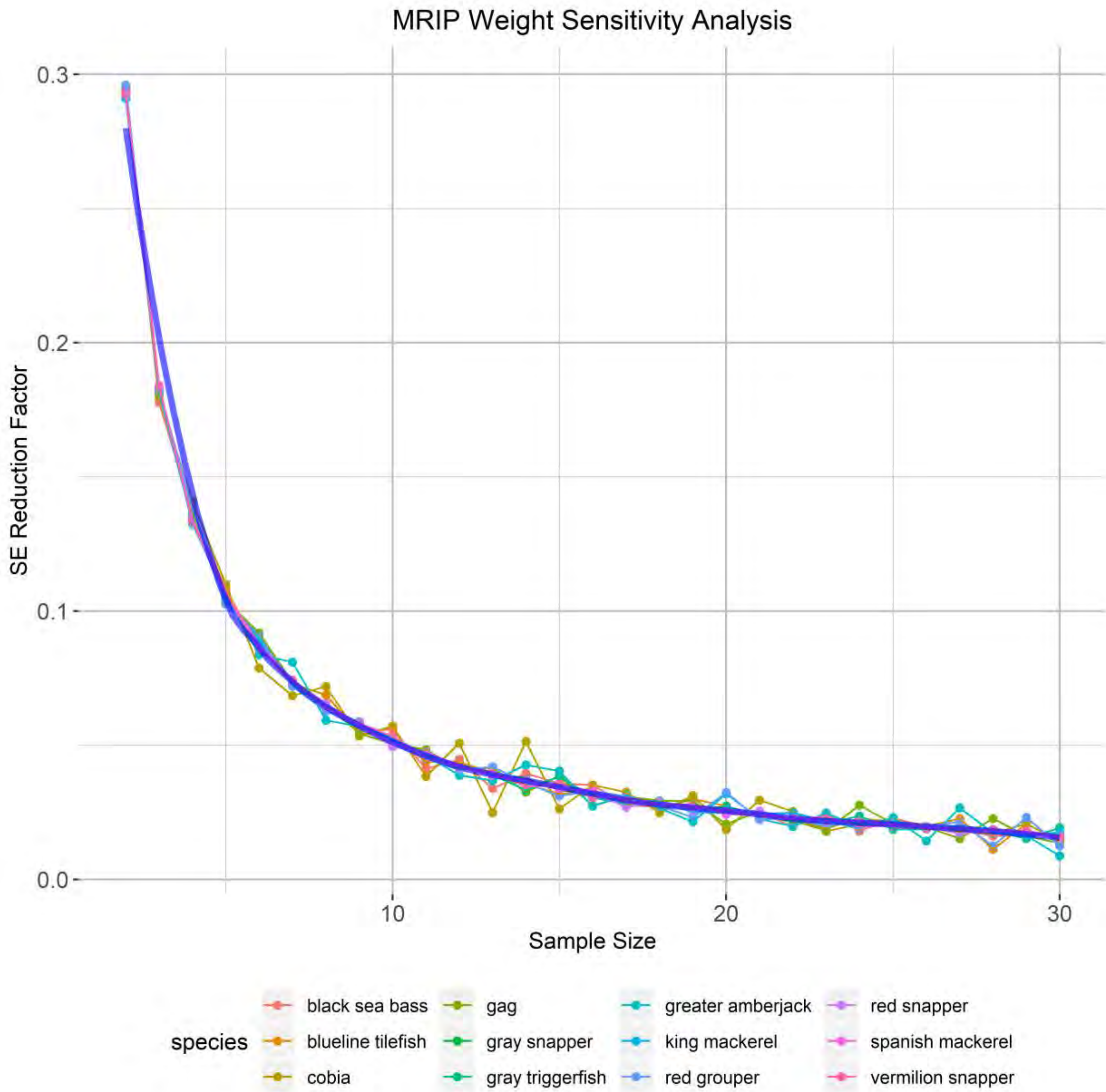
To standardize results across all species, the relative improvement in standard error with each additional sample was calculated by dividing the standard error at sample size n by the standard error at sample size n - 1. The average of these improvements was then calculated at the species level over all strata and plotted to visualize the average error reduction as a function of sample size.

$$SE\ Reduction\ Factor_{species,n} = mean(1 - \frac{SE_n}{SE_{n-1}})_{wave/subregion/state/mode/area}$$

RESULTS AND DISCUSSION

As seen in the Figure 1 below, mean weight estimates show significant gains in stability up to a sample size of about ten. Improvement begins to plateau around a sample size of fifteen, with reductions in standard error consistently below 5%. This suggests that a stratum sample size of not fewer than ten should be required to obtain reliable weight estimates, while samples greater than fifteen are likely not necessary as gains in precision are substantially diminished past this point. This follows what is expected theoretically, with reductions in standard error diminishing according to the square root of the sample size. While this analysis was only be completed for species with sufficient weight information, it is evident this pattern is consistent across multiple species.

Figure 1: Percent standard error reduction in 2018 mean weight estimates as a function of sample size.



Based on this result, it is suggested that a sample size of fifteen be used as a reasonable minimum threshold for calculating mean weight estimates before aggregating samples from coarser strata. This will result in more accurate, finer-resolution estimates, especially for species with fewer than 30 weights available in a stratum.

REFERENCES

Matter, V.M. and A. Rios. 2013. MRFSS to MRIP adjustment ratios and weight estimation procedures for South Atlantic and Gulf of Mexico managed species. Miami, FL. SEDAR 32-DW-02.

Matter, V.M. and S.C. Turner. 2010. Estimated recreational catch in weight: method for filling in missing weight estimates from the recreational surveys with application to yellowedge grouper, tilefish (golden), and blueline tilefish. Sustainable Fisheries Division Contribution No. SFD-2010-003. Miami, FL. SEDAR 22-DW-16.

NOAA Recreational Fishing Data Glossary. NOAA Fisheries. www.fisheries.noaa.gov/recreational-fishing-data/recreational-fishing-data-glossary#weight-data. 14 June 2019.

APPENDIX

Figure 2: Application of minimum sample size thresholds to South Atlantic Golden Tilefish.

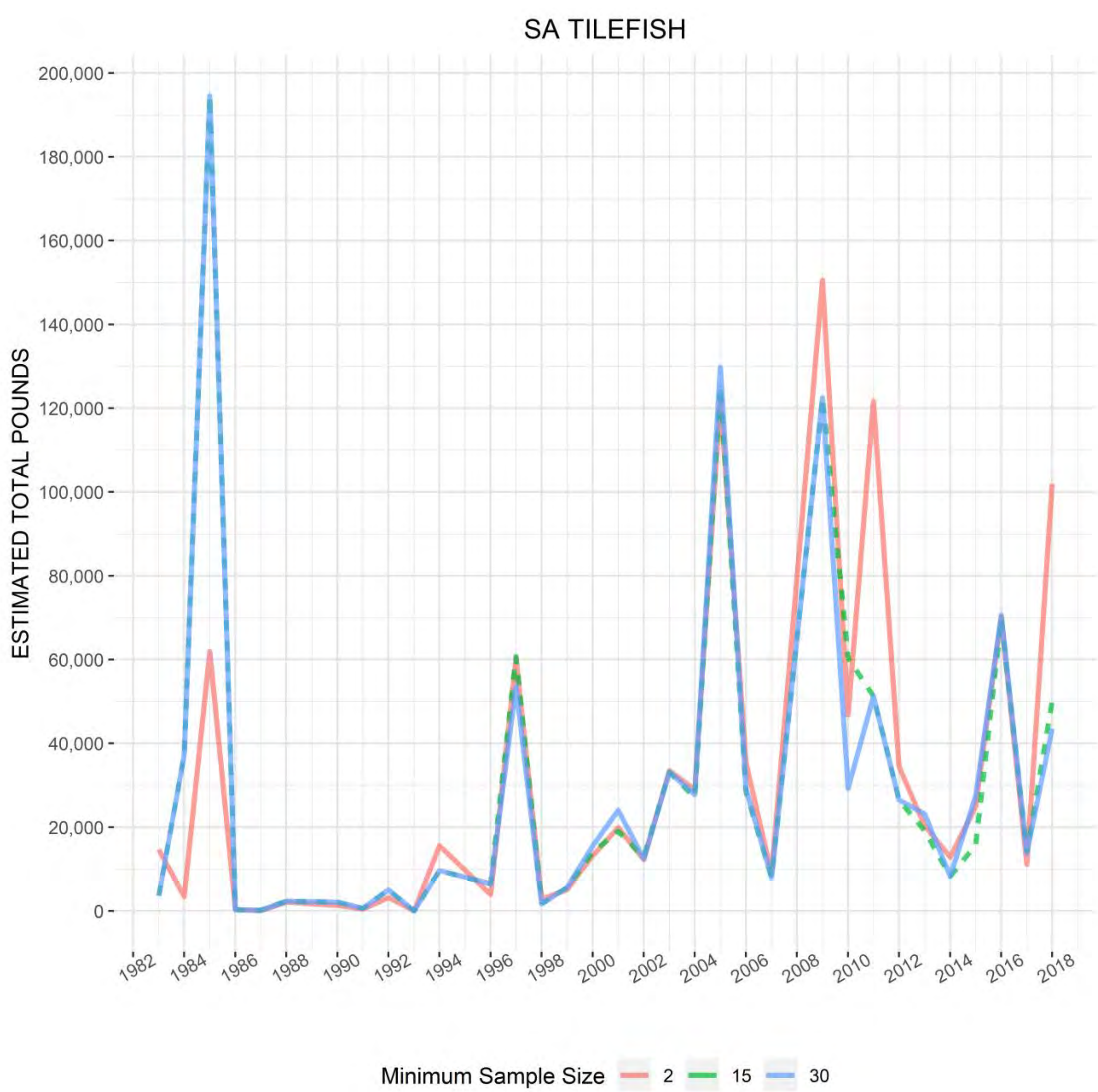


Figure 3: Application of minimum sample size thresholds to Gulf of Mexico Vermilion Snapper.