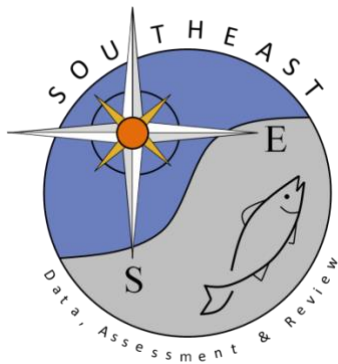


Evaluating Uncertainty in Gulf Red Snapper Estimates: A Preliminary Sensitivity Analysis of Non-Sampling Errors in the Region's Recreational Fishing Surveys

NMFS Office of Science and Technology

SEDAR74-DW-11

15 April 2022



This information is distributed solely for the purpose of pre-dissemination peer review. It does not represent and should not be construed to represent any agency determination or policy.

Please cite this document as:

NMFS Office of Science and Technology. 2022. Evaluating Uncertainty in Gulf Red Snapper Estimates: A Preliminary Sensitivity Analysis of Non-Sampling Errors in the Region's Recreational Fishing Surveys. SEDAR74-DW-11. SEDAR, North Charleston, SC. 39 pp.



NOAA
FISHERIES

Evaluating Uncertainty in Gulf Red Snapper Estimates

A Preliminary Sensitivity Analysis of
Non-Sampling Errors in the Region's Recreational
Fishing Surveys

National Marine Fisheries Service
Office of Science and Technology
Fisheries Statistics Division

Silver Spring, MD | November 2021

The following NOAA Fisheries staff and affiliates contributed to this report: John Foster, Katherine Papacostas, Catherine Krikstan, Rob Andrews, Ryan Kitts-Jensen, Richard Cody, Gordon Colvin and Lauren Dolinger Few.

Executive Summary

Background

There are six different survey programs currently operating in the Gulf of Mexico to monitor the private boat recreational red snapper fishery: NOAA Fisheries' Marine Recreational Information Program (MRIP), which administers the Access Point Angler Intercept Survey (APAIS) and Fishing Effort Survey (FES; which replaced the Coastal Household Telephone Survey, or CHTS) in Mississippi, Alabama, and Florida; the Texas Coastal Creel Survey (CCS); Louisiana's LA Creel; Mississippi Tails n' Scales; Alabama Snapper Check; and the Florida State Reef Fish Survey (SRFS). Where programs overlap, systematic differences exist among estimates of red snapper catch. To date, we cannot definitively state why the estimates are different, other than they likely all suffer from differential levels of non-sampling error, or error that causes estimates to differ from the "true" population values (in this case, "true" red snapper catch). The direction and magnitude of these non-sampling errors are currently unknown. With this study, we begin investigating how non-sampling errors may influence the magnitude of the estimates derived from the different recreational red snapper monitoring programs in the region. This study also motivates and supports a collaborative research initiative in response to the Congressional directive from the 2021 House Committee on Appropriations to conduct an independent assessment of the surveys operating in the Gulf of Mexico, and make recommendations for their improvement.

Study overview

We investigated how two general classes of non-sampling error, **non-response error** and **coverage error**, might drive differences among Gulf red snapper private boat landings estimates.

- To examine **non-response error** effects, we reproduced MRIP estimates under two alternative non-response assumptions and compared the results to the original MRIP estimates and to the state survey program estimates.
- To examine **coverage error** effects, we simulated coverage error in the FES by producing effort estimates for domains that approximate the likely coverage of state survey designs (specifically those of LA Creel, FL SRFS, and TX CCS). We were unable to simulate an MRIP design similar to those of AL Snapper Check and MS Tails n' Scales, but were still able to compare the magnitude of the simulated MRIP estimates with those from these two state programs.

We further examined the comparability of estimates among the state programs by 1) comparing results from simulated applications of several standard designs applied Gulf-wide and 2) calibrating the state program estimates to one another, in order to begin to investigate comparability among individual state program estimates.

Key findings

The results of this analysis indicate:

- Non-sampling errors can cause large systematic differences in estimates of recreational private boat catch and are likely driving the differences between estimates of red snapper landings in the Gulf of Mexico. However, at this preliminary stage, drivers are dependent upon assumptions made regarding the accuracy of the different surveys.
- Simulating common or standardized survey designs across the five Gulf states produces similar patterns in the relative distribution of red snapper landings estimates across each state, strongly suggesting the value of a common regional design.
- Simulating the conduct of each state survey design across the Gulf of Mexico produces different estimates for each state, suggesting individual state program estimates are not directly comparable and reinforcing the need for calibration to ensure comparability of estimates produced from programs using different designs.

These results, while informative, are not definitive. They do not eliminate the possibility of additional unidentified non-sampling errors that could mitigate effects described here. More collaborative research, including empirical work, is needed to better understand how non-sampling errors affect Gulf red snapper estimates.

Introduction

At present, six different state and federal data collection programs collect information on Gulf of Mexico recreational red snapper catch: NOAA Fisheries' Marine Recreational Information Program (MRIP), which administers its Access Point Angler Intercept Survey (APAIS) and Fishing Effort Survey (FES; which replaced the Coastal Household Telephone Survey, or CHTS) in Mississippi, Alabama, and Florida; the Texas Coastal Creel Survey (CCS); Louisiana's LA Creel; Mississippi Tails n' Scales; Alabama

Snapper Check; and the Florida State Reef Fish Survey (SRFS). While the specifics of their survey designs differ, these six programs use three broad survey design approaches, and may be considered “general” or “specialized,” depending on whether they collect information for all or a subset of finfish species encountered. These broad survey design approaches include:

- **A complemented surveys approach**, which pairs an **on-site intercept survey** to collect catch-per-trip data with an **off-site telephone, mail, or email survey** to collect effort data. The MRIP APAIS-FES and previously the MRIP APAIS-CHTS (both general surveys), LA Creel (a general survey), and FL SRFS (a specialized survey) follow this approach.
- An **intercept-only survey**, which collects data from anglers at public fishing access sites. The TX CCS (a general survey) follows this approach.
- A **capture-recapture survey**, which collects data through mandatory electronic trip reports (the “capture” phase) and validates data through an in-person intercept survey at public fishing access sites (the “recapture” phase). MS Tails n’ Scales (a specialized survey) and AL Snapper Check (a specialized survey) follow this approach.

While these data collection programs produce estimates of private boat red snapper catch, design differences among the programs have likely contributed to estimates that differ from one another. It is suspected that these differences are driven by non-sampling errors, or errors that can cause estimates to differ, in a systematic or variable way, from “true” population values. There are many sources of non-sampling errors in surveys, and while their direction and magnitude are challenging to predict, we do know they can vary significantly among survey designs (e.g., Dalenius 1977, Dillman and Christian 2005, Phung et al. 2015,). Two common types of non-sampling errors that may be driving differences among Gulf of Mexico recreational red snapper estimates are:

- **Non-response error**, which occurs when a member of the sample is unable or unwilling to respond to a survey and differs in some key characteristic from those who do respond. For example, those who don’t respond to a fishing activity survey may fish more or less often—or catch more or less fish—than those who do respond.
- **Coverage error**, which occurs when members of a target population are omitted, duplicated, or wrongly included in a sample frame. For example, a landline-based sample frame omits members of the target population who don’t have landline telephones; a fishing license-based sample frame omits those who fish without a

license; and an intercept survey that is only conducted at public fishing access sites omits those who fish from private sites.

Understanding how non-sampling errors may affect each survey design can help us understand what is driving the differences between each set of estimates. It is also an imperative initial step in addressing the Congressional directive NOAA Fisheries has received from the House Committee on Appropriations (2021) “*to contract with a non-governmental entity with expertise in statistics and fisheries-dependent data collection to provide the following: (1) an independent assessment of the accuracy and precision of both the Federal and State recreational catch data programs in the Gulf of Mexico; (2) recommended improvements to be made to the Federal and State recreational catch data programs in the Gulf of Mexico to improve accuracy and precision; and (3) an independent assessment, based on the results of the two prior items, of how best to calibrate the Federal and State recreational catch data programs in the Gulf of Mexico to a common currency.*” As such, this preliminary sensitivity analysis sought to:

- Simulate the presence of non-sampling error in MRIP APAIS-FES estimates, compare the results of these simulations to the estimates produced by each state program, and determine if the non-sampling errors could account for differences between state and federal surveys.
- Use simulations of state and regional survey designs to evaluate the comparability of state program estimates to one another.

While this analysis is an important step toward understanding the sources of differences between estimates of red snapper landings, this research should be considered preliminary. Additional investigation into non-sampling errors will be needed to improve the comparability, accuracy and precision of all available red snapper catch estimates. Further, this report is not NOAA Fisheries’ response to the Congressional directive cited above, but rather a preliminary exploratory analysis to motivate and support a collaborative research initiative that is essential for continued evaluation and improvement of the surveys operating in the Gulf of Mexico.

Methods

Simulating effects of non-sampling errors in MRIP APAIS-FES estimates

Non-response error

We examined possible effects of non-response error on MRIP estimates by producing FES estimates under two alternative non-response assumptions, and then using these revised FES estimates to produce revised MRIP red snapper landings. We then graphically compared the magnitude of these estimates to MRIP's standard red snapper landings estimates, as well as to the estimates of four individual state programs: AL Snapper Check, FL SRFS, LA Creel, and MS Tails n' Scales.

The FES, as currently designed, assumes non-respondents and respondents who share similar characteristics—specifically, proximity to the coast, possession of a fishing license, and registered boat ownership—have similar fishing activity. This assumption is supported by two non-response follow-up studies, which demonstrated no significant differences in fishing activity between those who initially responded to the FES and those who responded to a follow-up questionnaire (Andrews et al. 2014, and Andrews 2021).

Our first non-response error sensitivity scenario assumed that all FES non-respondents have zero fishing activity (i.e., we assumed that the response rate for fishing households was 100%). If this assumption were true, it would mean the current FES estimates suffer from maximum non-response error in the direction of greatly overestimating fishing effort. This assumption, while not supported by empirical research, is a theoretical possibility if the remaining non-respondents from the non-response follow-up studies have zero fishing activity. We generated this maximum non-response error series by down-weighting MRIP estimates by the final response rate from the 2021 non-response follow-up study (42.2%; Andrews 2021).

Our second scenario assumed that only the non-respondents in unlicensed households have zero fishing activity. Since the FES stratifies sampling by licensed and unlicensed households (addresses that do or do not match to at least one state fishing license in the National Saltwater Angler Registry), it is possible to apply this assumption only to unlicensed household strata separate from the licensed strata. If this assumption were true, it would mean that the FES suffers from non-response error and overestimates fishing effort, but to a lesser degree than the first alternative assumption. This series

was produced by calculating effort as we normally do for license strata of the FES sample, but down-weighting the FES unlicensed household strata by their annual response rates, adjusted to approximate an average stratum-level response rate from the non-response follow-up studies (Andrews et al. 2014, Andrews 2021). These adjustments to the annual response rates were done to reflect the higher response rates of the non-response follow-up studies, since neither study showed evidence of systematic non-response error based on the additional households that responded to the follow-up questionnaire.

In most cases, we produced these two alternative MRIP FES estimates in each state, and used these revised effort estimates to produce red snapper landings for 2015-2019 using standard MRIP catch estimation methods (described in Papacostas and Foster 2018). However, in LA, we only have one year (2015) of MRIP estimates, so we simulated MRIP LA estimates for 2016-2019 in order to make multi-year comparisons. We generated these estimates ($\hat{Y}_{MRIP_{yyyy}}$) by multiplying the LA Creel landings estimates ($\hat{Y}_{LC_{yyyy}}$) for those years by the corresponding 2015 ratios of MRIP ($\hat{Y}_{MRIP_{2015}}$) to LA Creel ($\hat{Y}_{LC_{2015}}$) estimates:

$$\hat{Y}_{MRIP_{yyyy}} = \hat{Y}_{LC_{yyyy}} \frac{\hat{Y}_{MRIP_{2015}}}{\hat{Y}_{LC_{2015}}} \quad (1).$$

Similarly, MS Tails n' Scales landings estimates (in numbers of fish) are unavailable for 2015-2016 (Mississippi Department of Marine Resources, 2021). However, using the 2016 Tails n' Scales landings in weight (lbs), we generated a 2016 MS Tails n' Scales landings estimate in numbers of fish ($\hat{Y}_{TnS_{2016}}$) as the ratio of the 2016 Tails n' Scales weight estimate ($\hat{Y}_{TnS,w_{2016}}$) to the MRIP mean weight estimate ($\hat{Y}_{MRIP,w}$) for all years (2015-2019):

$$\hat{Y}_{TnS_{2016}} = \frac{\hat{Y}_{TnS,w_{2016}}}{\hat{Y}_{MRIP,w}} \quad (2).$$

Coverage error

We simulated three potential sources of coverage error using existing MRIP data and compared the magnitude of resulting estimates to the standard MRIP estimates and the

state program estimates. Two sources of coverage error were simulated as estimation domains from within the full FES sample, and the third was simulated by excluding FES data and estimating red snapper landings directly from APAIS data.

Limiting FES Coverage to Households with Only Landline Telephones

Because the MRIP FES asks a question about household phone type and use, we can produce estimates for a landline-only domain that simulate landline-only coverage error. This adjustment, which includes re-weighting of the landline sample to represent the full household population, is meant to make the FES design more similar to the design of the CHTS (the landline-only telephone survey the FES replaced in 2018). CHTS effort estimates began to decline in the later years of the survey, which was more likely due to increasing coverage error caused by the emergence of wireless-only households in the early 2000s rather than a real decline in fishing effort (Andrews et al. 2014). We used these revised effort estimates to estimate red snapper landings for 2016-2017 using standard MRIP catch estimation methods, and compared the results to the standard MRIP APAIS-FES landings, the MRIP APAIS-CHTS landings, AL Snapper Check landings, MS Tails n' Scales landings, and FL SRFS landings.

Limiting FES Coverage to Only Licensed Households

Because the MRIP FES stratifies households based on their match to an entry in the National Saltwater Angler Registry, we can produce estimates for domains that simulate coverage of only licensed households (i.e., non-coverage of unlicensed households). This design approximates LA Creel and SRFS, which sample from angler license/permit databases and do not cover unlicensed anglers (although both surveys use correction factors to adjust for unlicensed fishing activity; see Florida Fish and Wildlife Conservation Commission 2018, Louisiana Department of Wildlife and Fisheries 2017). We used these revised license-only FES effort estimates to produce red snapper landings for 2015-2019, and compared them to the standard MRIP APAIS-FES estimates, LA Creel estimates, and FL SRFS estimates. Since we do not have MRIP LA estimates for 2016-2019, we again simulated the standard MRIP series and the MRIP license-only series for these years by multiplying the LA Creel values in 2016-2019 by the ratios of the corresponding values in 2015 (see equation 1).

A caveat to this comparison is that, for MRIP, we do not have the information needed to remove unlicensed out-of-state anglers from effort calculations, so the MRIP license-only series includes effort from all out-of-state anglers (both licensed or unlicensed). However, historic APAIS data indicate that state resident anglers account for the large majority of private boat effort - typically 85% or higher.

Eliminating FES Coverage and Using Only Public Access Fishing Effort Covered by APAIS

MRIP APAIS-FES estimates are produced through a complemented surveys design meant to account for both public and private access fishing, but we can produce catch estimates using only APAIS to simulate a single intercept-only survey design. This design would make MRIP more similar to TX CCS, which uses an intercept survey design alone to produce catch estimates and does not have an accompanying off-site effort survey. We produced these APAIS-only estimates for AL, FL, LA and MS, and compared them to the standard MRIP APAIS-FES estimates.

We also simulated estimates for Texas using complemented survey designs. The FES was conducted in 2016 in Texas, so we combined TX CCS catch rate data with the FES effort data to produce landings estimates in a similar manner to MRIP APAIS-FES estimates for 2016. We then used the ratio of the simulated 2016 TX CCS-FES estimates to official TX CCS-only estimates to produce estimates for 2015 and 2017-2019:

$$\hat{Y}_{CCS,FES_{yyyy}} = \hat{Y}_{CCS_{yyyy}} \frac{\hat{Y}_{CCS,FES_{2016}}}{\hat{Y}_{CCS_{2016}}} \quad (3).$$

A caveat to this comparison is that we could not effectively simulate the standard coverage adjustment MRIP applies to the FES effort estimates for out-of-state anglers using the APAIS. As such, the TX CCS-FES estimates, although similar, are not entirely comparable to the standard MRIP APAIS-FES estimates. As mentioned earlier regarding other states, it may be that state resident anglers account for the large majority of private boat trips in Texas as well, but that information is not currently available.

Examining the comparability of regional estimate trends using different regional and state designs

Having produced several series of estimates that simulated different scenarios where a common design was used in all states throughout the region, we sought to understand how the relative state-to-state landings of the different designs compared to those of MRIP and the various state programs. We examined these patterns through two general comparisons: (1) the state-to-state differences in the private boat red snapper landings, both at the original scale and indexed (to remove scaling effects), and (2) the proportional contribution of each state estimate to the regional total of red snapper landings. For this comparison, there were three possible common design scenarios: the

intercept-only design (comparable to TX CCS), a complemented design which combined an intercept survey with a license frame effort survey (comparable to FL SRFS and LA Creel), and the MRIP complemented design which combines the APAIS and FES surveys. See Table 1 for descriptions of the estimate series used for each comparison.

Table 1: Estimate series used for comparisons of regional trends resulting from the use of different designs

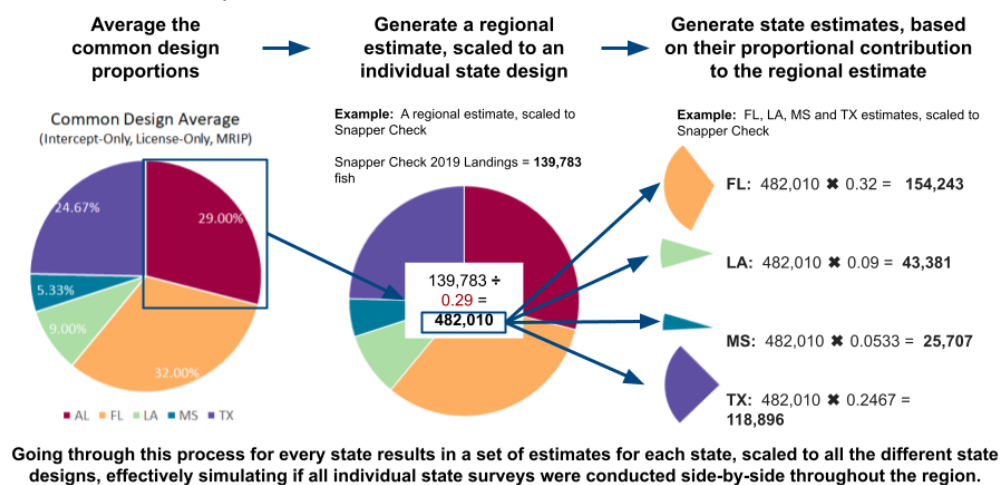
Design Series	Estimate Series Used by State and Year
Intercept-Only (Simulates a common design throughout the region that covers only public access fishing effort)	<p>For Alabama, Florida, Louisiana, and Mississippi:</p> <ul style="list-style-type: none"> 2015-2019: <ul style="list-style-type: none"> MRIP estimates produced using only APAIS data. <p>For Texas:</p> <ul style="list-style-type: none"> 2015-2019: <ul style="list-style-type: none"> TX CCS estimates
License-Only (Simulated common design throughout the region that combines an intercept survey with an effort survey that covers only licensed fishing effort)	<p>For Alabama and Mississippi:</p> <ul style="list-style-type: none"> 2015-2019: <ul style="list-style-type: none"> MRIP estimates produced using FES fishing effort data only from licensed households <p>For Louisiana:</p> <ul style="list-style-type: none"> 2015-2019: <ul style="list-style-type: none"> LA Creel <p>For Florida:</p> <ul style="list-style-type: none"> 2015-2019: <ul style="list-style-type: none"> FL SRFS <p>For Texas:</p> <ul style="list-style-type: none"> 2016: <ul style="list-style-type: none"> TX CCS catch rate estimates combined with 2016 FES pilot survey effort data from licensed households only 2015, 2017-2019: <ul style="list-style-type: none"> TX CCS catch rate estimates multiplied by the ratio of 2016 TX CCS-FES license-only household estimate
MRIP (Common design throughout the region that uses the APAIS and FES surveys, simulated in states/years where MRIP data is unavailable)	<ul style="list-style-type: none"> For Alabama, Florida and Mississippi: <ul style="list-style-type: none"> 2015-2019: <ul style="list-style-type: none"> Standard MRIP APAIS and FES estimates For Louisiana: <ul style="list-style-type: none"> 2015: <ul style="list-style-type: none"> Standard MRIP APAIS and FES estimates 2016-2019: <ul style="list-style-type: none"> LA Creel values in 2016-2019 multiplied by the ratios of MRIP to LA Creel values in 2015 (using equation 1) For Texas: <ul style="list-style-type: none"> 2016: <ul style="list-style-type: none"> TX CCS catch rate estimates combined with 2016 FES pilot survey effort data 2015, 2017-2019: <ul style="list-style-type: none"> TX CCS catch rate estimates multiplied by the ratio of 2016 TX CCS-FES estimate (using equation 3)

State Programs (Different design in each state)	<ul style="list-style-type: none"> • Alabama, Florida, Louisiana, Texas: <ul style="list-style-type: none"> ○ 2015-2019: <ul style="list-style-type: none"> ■ Estimates from the respective state programs (AL Snapper Check, FL SRFS, LA Creel, TX CCS) • Mississippi: <ul style="list-style-type: none"> ○ 2016 <ul style="list-style-type: none"> ■ MS Tails n' Scales estimates produced using equation 2 ○ 2017-2019: <ul style="list-style-type: none"> ■ Estimates from MS Tails n' Scales
---	--

Examining the comparability of state estimates using different state designs

Since MRIP is the only regional survey and state programs are limited to single states, we are typically only able to make direct comparisons between MRIP and state programs where they overlap, while the comparability of the state program designs to each other is unknown. To begin to address this issue, we developed a method to simulate all five state survey designs being run side-by-side throughout the region, effectively calibrating the state estimates to each other. First, we averaged the proportional contribution of each state estimate to the regional red snapper landings estimate for each of the common design scenarios described in Table 1 above (intercept-only, license-only, and MRIP; see Figure 8 and Appendix B for state proportional contributions to the regional estimate under each common design). We then generated five regional estimates by scaling the individual state estimates using a single state program estimate as the control design (see Figure 1 for a schematic of the process with example calculations). The key assumption of this analysis is that we expect the same relative distribution of estimates if any individual state program design was used in all states.

Figure 1: Schematic of methods used to simulate individual state designs across the Gulf, using AL Snapper Check as an example



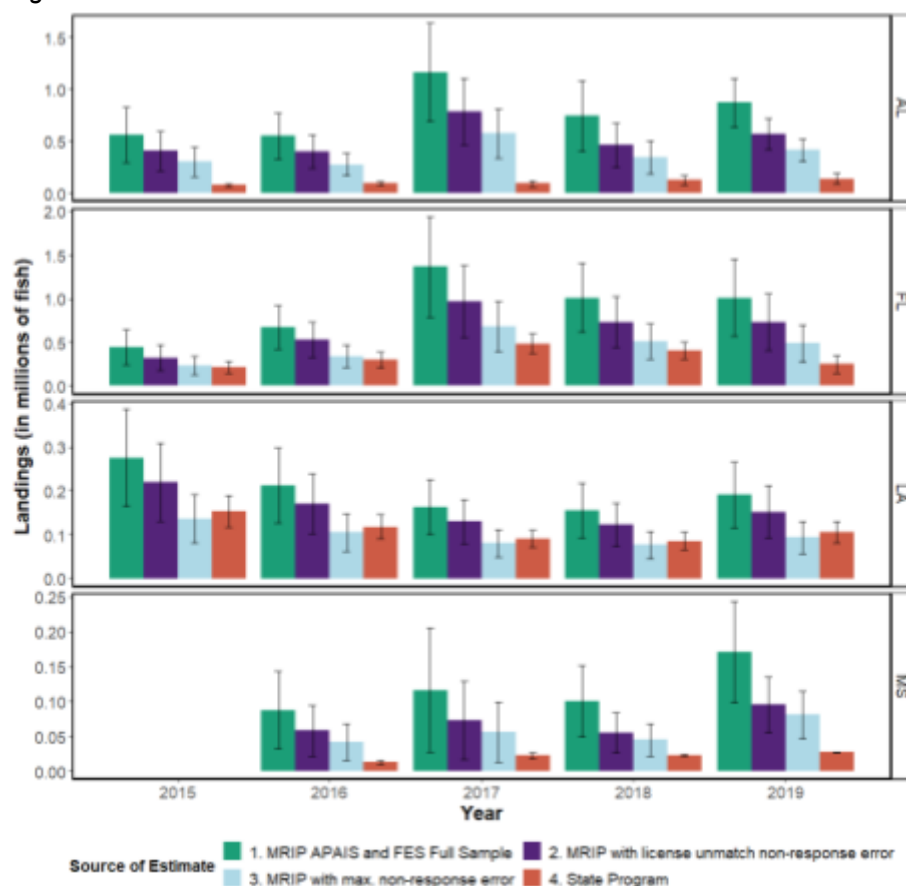
Results

Simulating effects of non-sampling errors in MRIP APAIS-FES estimates

Non-response error

We compared four different 2015-2019 estimate time series: MRIP APAIS-FES estimates, produced using standard MRIP estimation methods (Series 1 in Figure 2); MRIP APAIS-FES estimates downweighted based on the assumption that non-respondents in unlicensed households (i.e., those households that could not be matched to the National Saltwater Angler Registry) have zero fishing activity (Series 2 in Figure 2); MRIP APAIS-FES estimates downweighted based on the assumption that all FES non-respondents have zero fishing activity (Series 3 in Figure 3), and state program estimates (Series 4 in Figure 2—AL Snapper Check in Alabama, FL SRFS in Florida, LA Creel in Louisiana, and MS Tails n' Scales in Mississippi). We found that standard MRIP estimates were consistently highest, the two non-response adjusted scenarios were of intermediate magnitude, and the state program estimates were lowest. When applying the maximum non-response adjustment (Series 3 in Figure 2), MRIP APAIS-FES estimates were similar to LA Creel estimates, but still systematically higher than FL SRFS estimates (albeit not significantly), and much higher than AL Snapper Check and MS Tails n' Scales estimates (Figure 2).

Figure 2: Effects of varying non-response error assumptions on the magnitude of MRIP private boat red snapper landings estimates



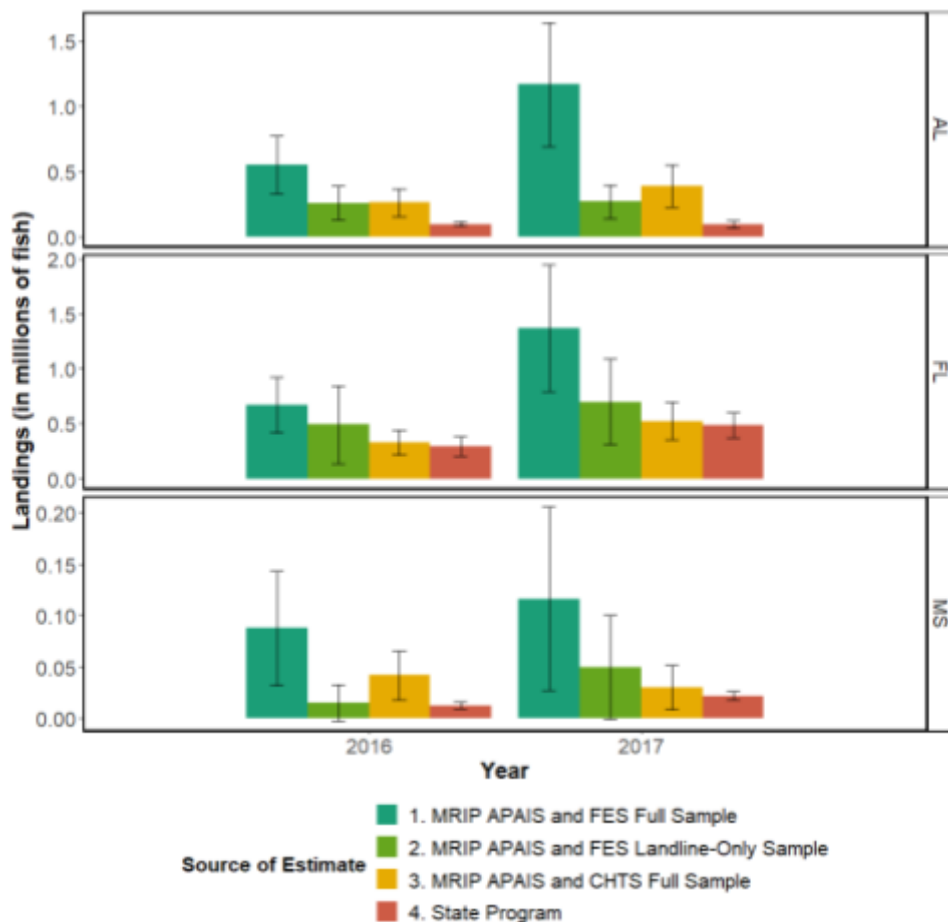
Coverage error

Limiting FES Coverage to Households with Only Landline Telephones

For this comparison, we examined four different 2016-2017 estimate time series: MRIP APAIS-FES estimates, produced using standard MRIP estimation methods (Series 1 in Figure 3); MRIP APAIS-FES estimates for landline-only domains (Series 2 in Figure 3); MRIP APAIS-CHTS estimates (Series 3 in Figure 3); and state program estimates (Series 4 in Figure 3—AL Snapper Check in Alabama, FL SRFS in Florida, and MS Tails n' Scales in Mississippi). MRIP APAIS-FES estimates for landline-only domains resulted in estimates of similar magnitude to those of the MRIP APAIS-CHTS estimates. In comparison to the state programs, the MRIP APAIS-FES landline-only estimate and the MRIP APAIS-CHTS estimates were similar to the FL SRFS estimate, but are systematically higher than the AL Snapper Check estimate and the MS Tails n' Scales estimate (Figure 3). These differences, however, were likely only significant in Alabama,

where there was no overlap in error bars between the MRIP APAIS-FES landline-only estimates and the AL Snapper Check estimates (Figure 3).

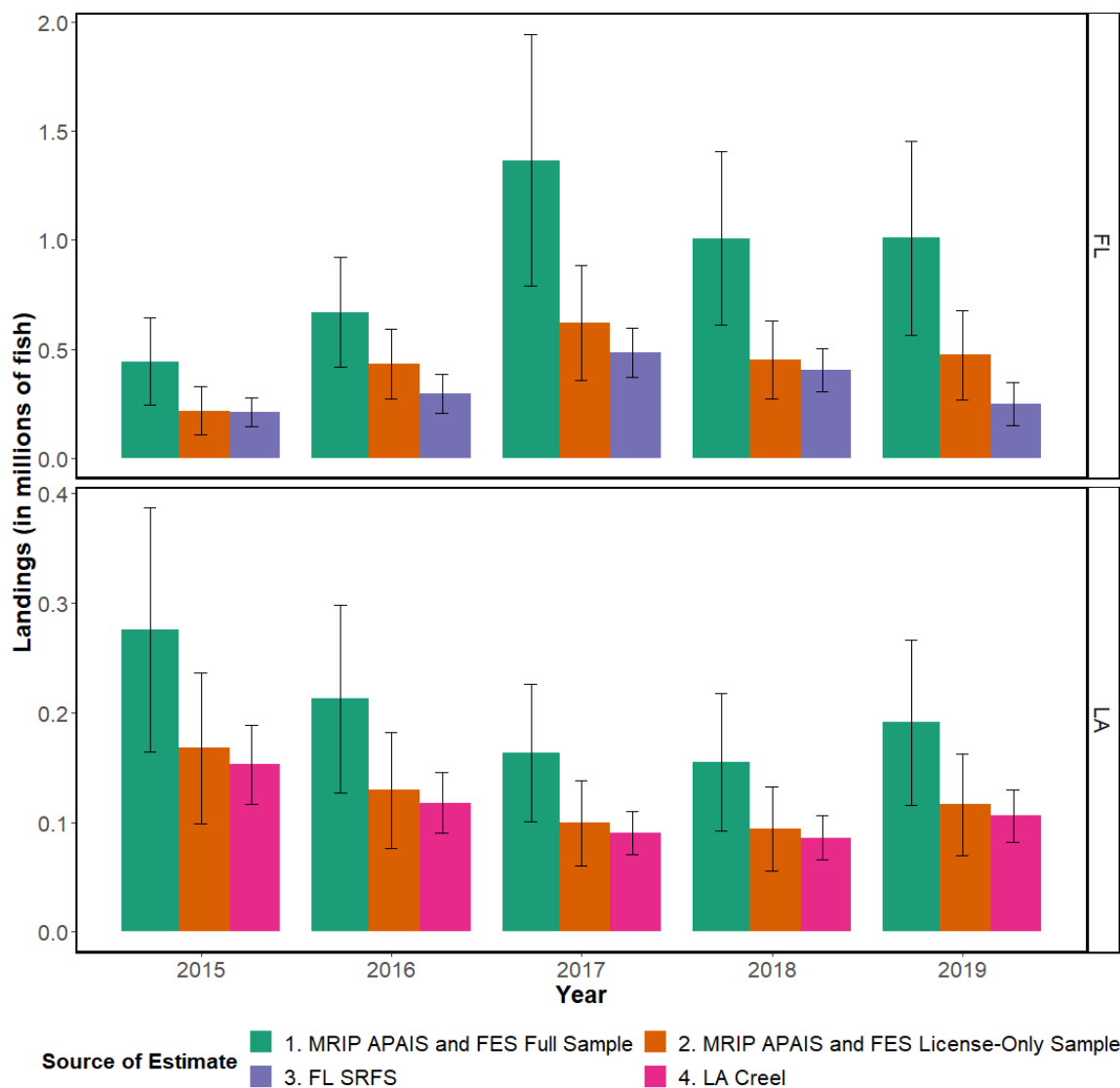
Figure 3: Effects of landline-only coverage on the magnitude of MRIP private boat red snapper landings estimates



Limiting FES Coverage to Only Licensed Households

The MRIP APAIS-FES license-only estimates (Series 2 in Figure 4) were much smaller than the standard MRIP APAIS-FES estimates (Series 1 in Figure 4). MRIP APAIS-FES license-only estimates were still systematically higher than FL SRFS estimates (Series 3 in Figure 4), but much closer to them than standard MRIP APAIS-FES estimates. MRIP APAIS-FES license-only estimates and LA Creel estimates (Series 4 in Figure 4) were of similar magnitude. Since, [as discussed above](#), the MRIP APAIS-FES license-only estimates included both licensed and unlicensed out-of-state-anglers, the “true” MRIP license-only series would likely be lower than our results, potentially resulting in estimates more in-line with those of FL SRFS and somewhat lower than those of LA Creel. Regardless, error bars of the MRIP APAIS-FES license-only series overlapped consistently with those of both LA Creel and FL SRFS (Figure 4).

Figure 4: Effects of license-only coverage on the magnitude of MRIP private boat red snapper landings estimates



Eliminating FES Coverage and Using Only Public Access Fishing Effort Covered by APAIS

The MRIP APAIS-only estimates (Series 2 in Figure 5) were significantly lower than MRIP APAIS-FES estimates (Series 1 in Figure 5) in Alabama, Florida, Louisiana and Mississippi. The differences between standard MRIP APAIS-FES estimates and MRIP APAIS-only estimates in these states were similar in magnitude to the differences in Texas between the TX CCS-FES estimates (Series 3 in Figure 5) and the TX CCS estimates (Series 4 in Figure 5). While not directly comparable in terms of non-sampling errors, AL Snapper Check and MS Tails n' Scales estimates were closest in magnitude to the MRIP APAIS-only series (Figure 6).

Figure 5: Effects of public-access only coverage on the magnitude of MRIP private boat red snapper landings estimates

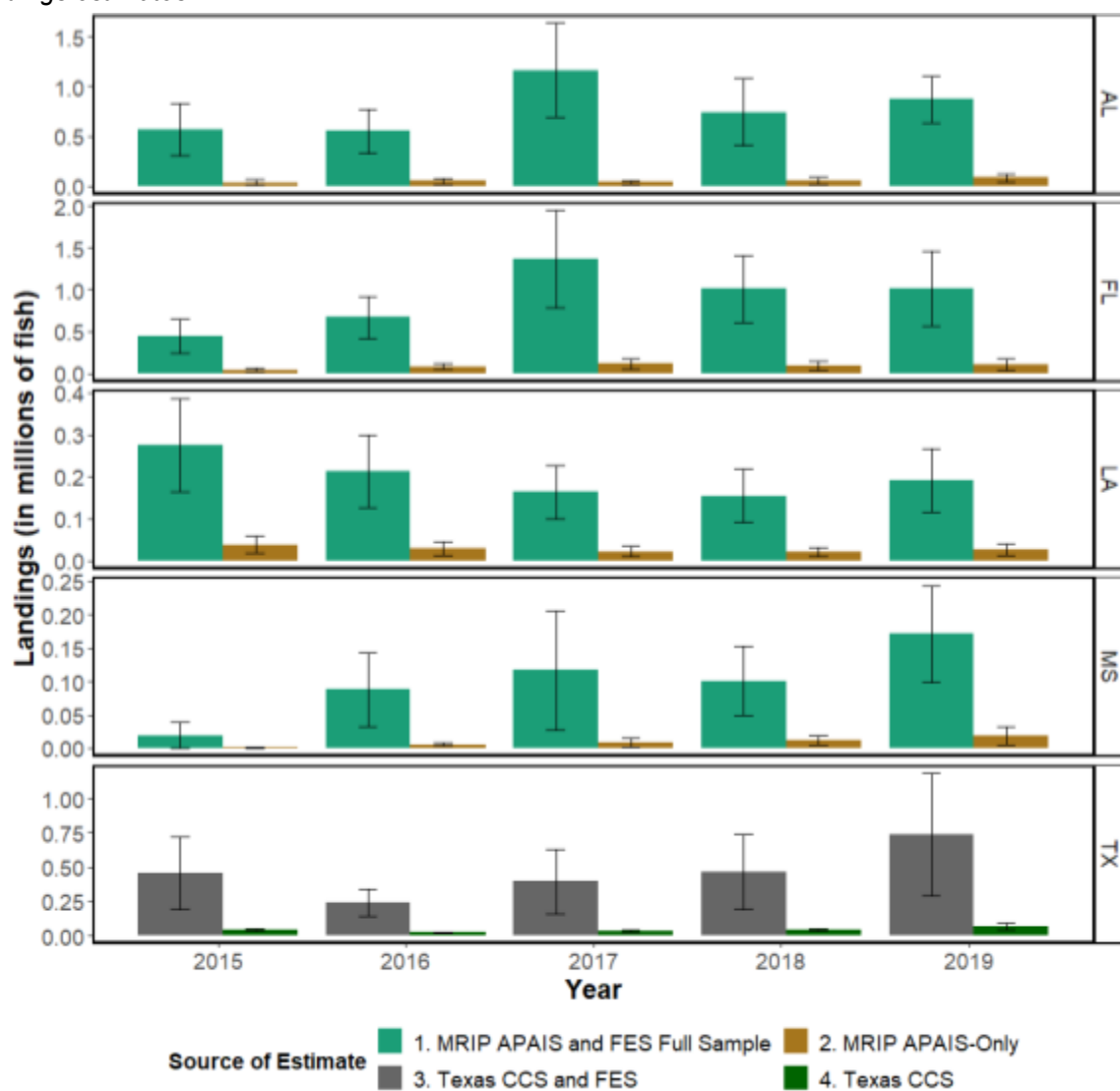
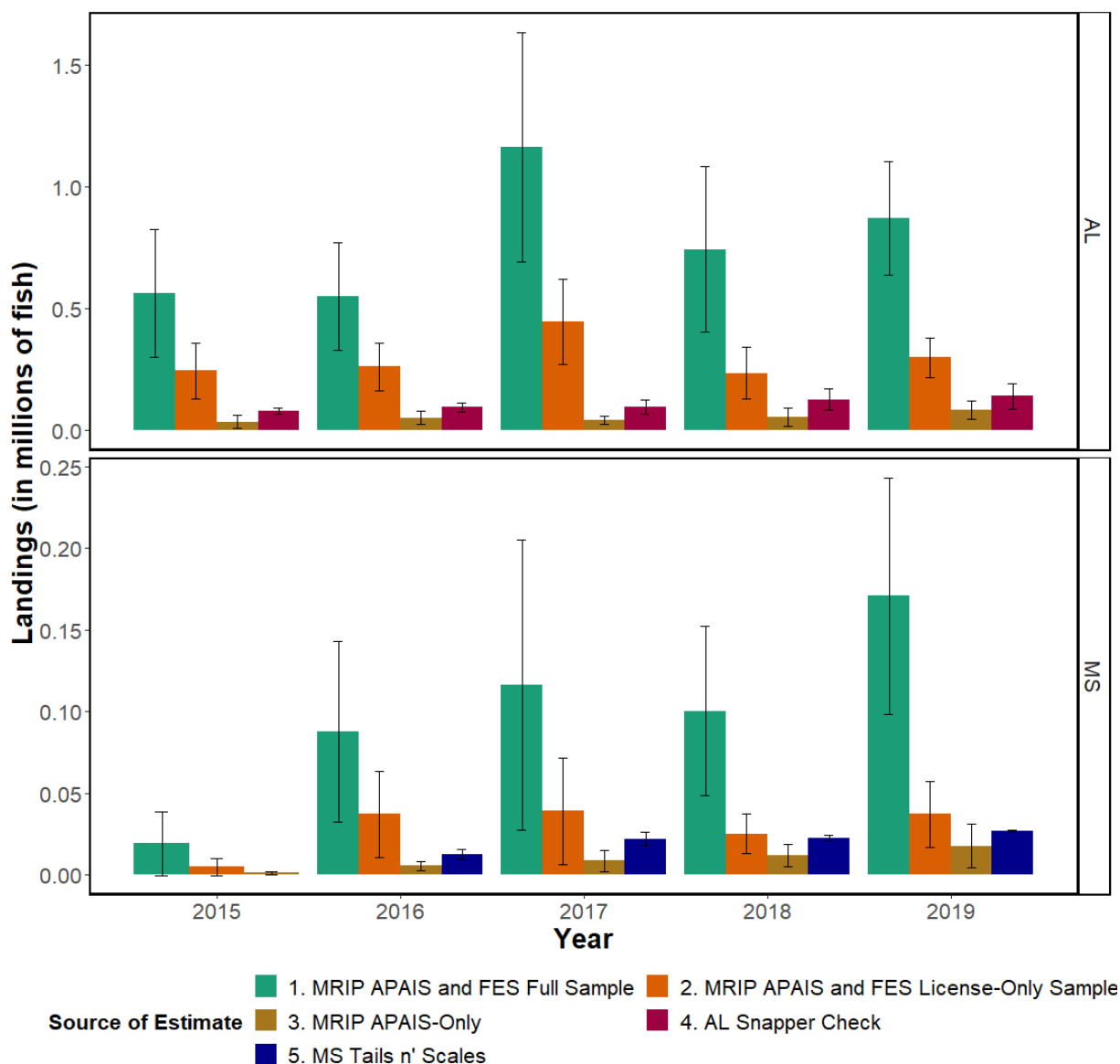


Figure 6: Comparison of private boat red snapper landings produced using the standard MRIP design, the MRIP license-only design, the MRIP APAIS-only design, AL Snapper Check, and MS Tails n' Scales.

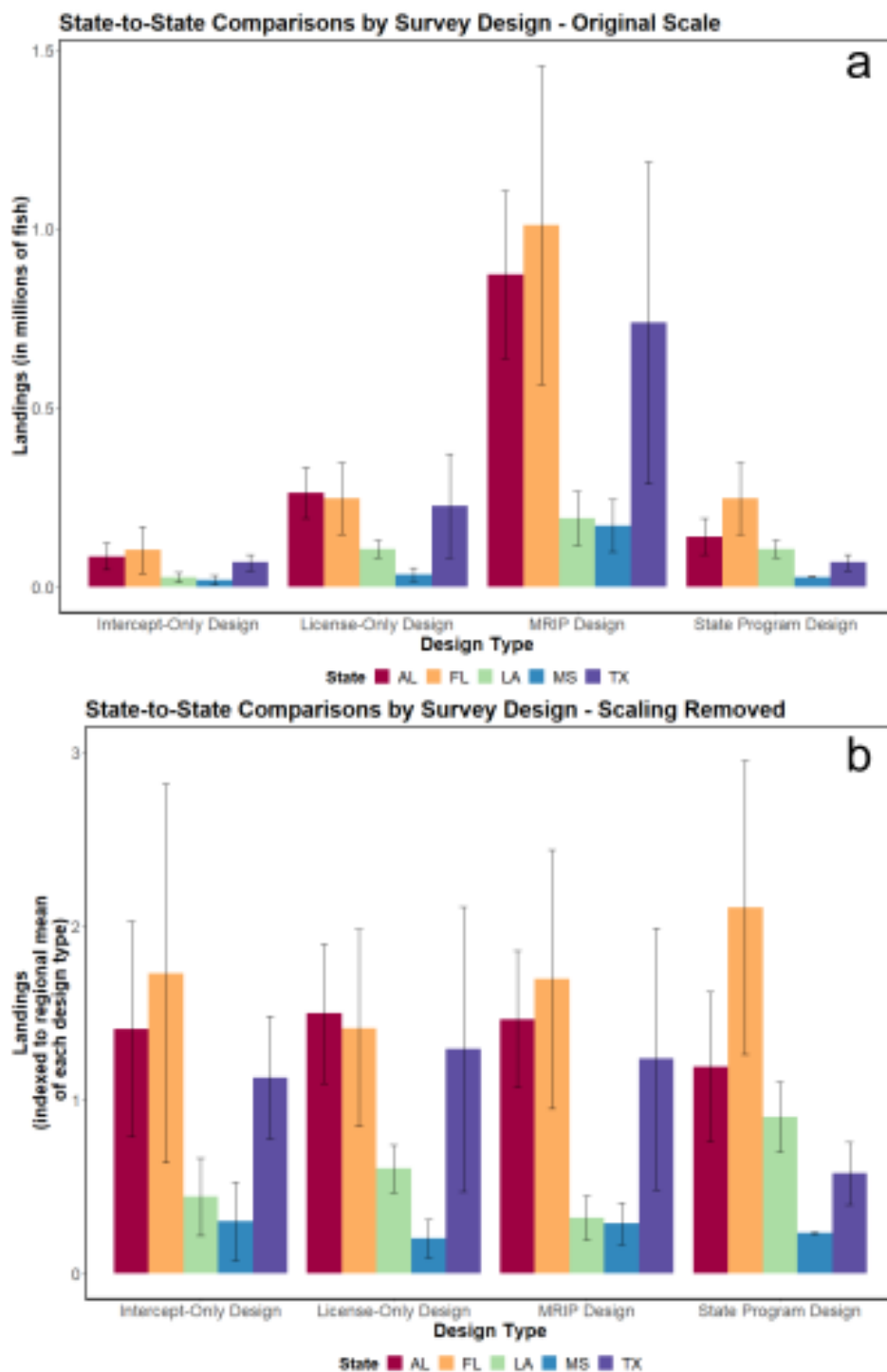


Examining the comparability of regional estimate trends using different regional and state designs

When we compare the estimates produced from the common designs (intercept-only, license-only, and MRIP) across states, the relative state-to-state distributions in red snapper landings are similar (Figure 7). However, the relative state-to-state distributions are noticeably different when every state uses a different design (Figure 7). While

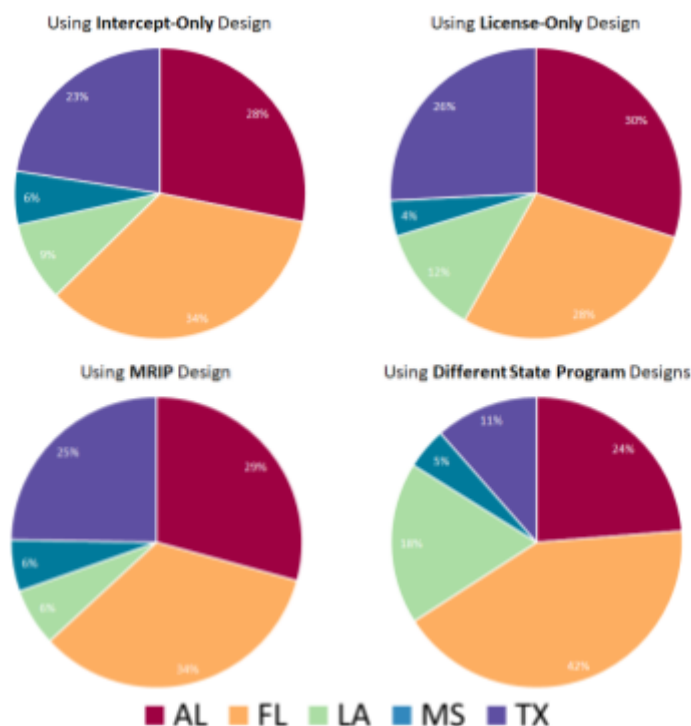
Figure 7 only shows estimates for 2019, this pattern is relatively consistent across years (see Appendix A).

Figure 7: State-to-state differences in 2019 private boat red snapper landings estimates under different survey designs (a) at the original scale and (b) indexed to the regional mean of each design type to remove scaling effects. See Appendix A for 2015-2018 comparisons.



We further found common designs resulted in very similar state proportional contributions to the regional red snapper landings estimate (Figure 8). The different state program designs resulted in state proportions that were systematically different from those of the common designs (Figure 8). This pattern was also relatively consistent across years (see Appendix B).

Figure 8: State contributions to the 2019 regional red snapper landings estimate under different survey designs. See Appendix B for 2015-2018 comparisons.

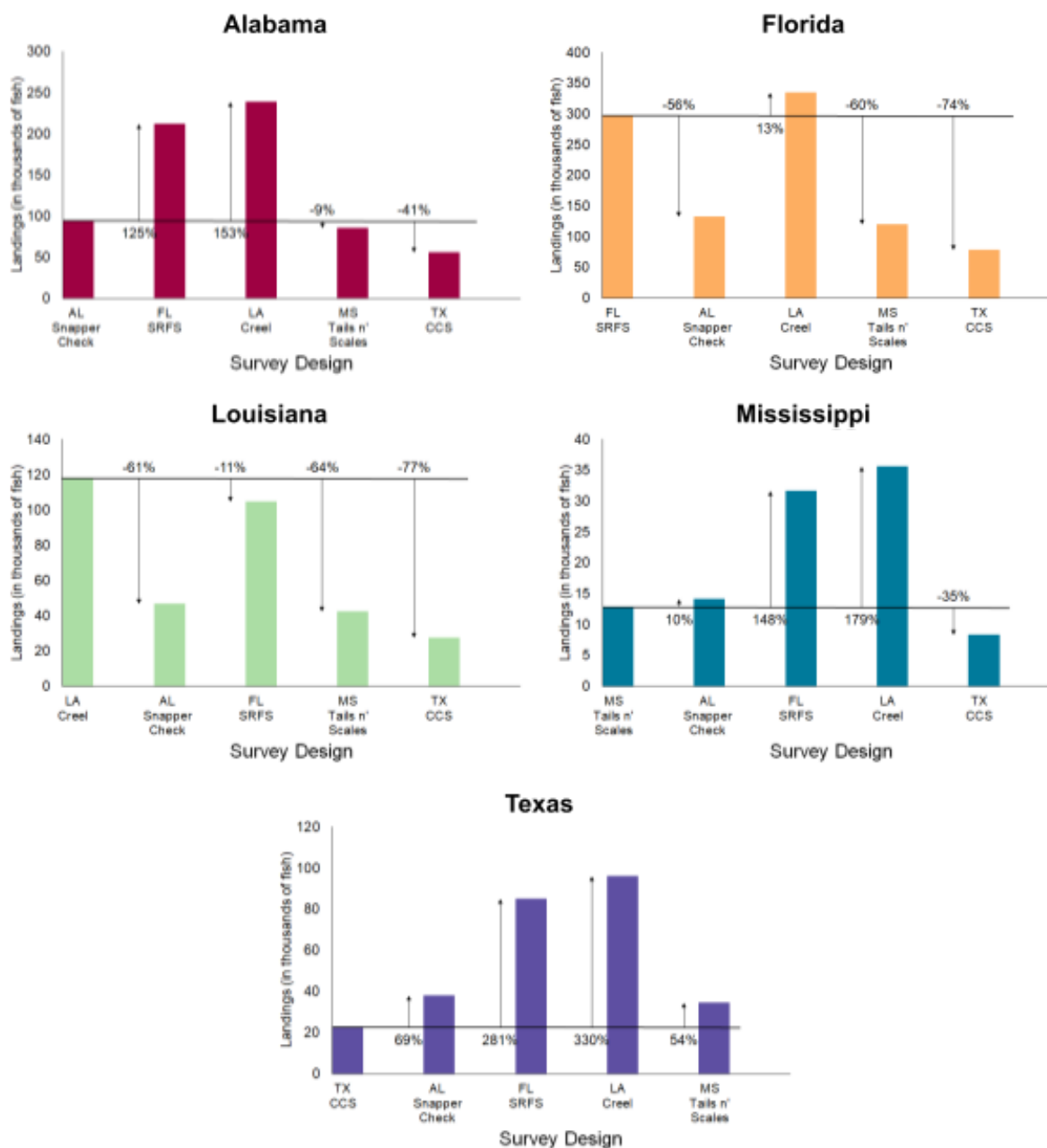


Examining the comparability of state estimates using different state designs

Effectively calibrating all the state program estimates to each other (i.e., simulating all state survey designs conducted side-by-side throughout the region) resulted in five different red snapper estimates for each state (Figure 9). In 2019 in Alabama, AL Snapper Check estimates were over 100% lower than the simulated estimates using FL SRFS and LA Creel designs, and nearly 50% higher than the simulated estimates from the TX CCS design, but were similar to the simulated estimates using the MS Tails n' Scales design (top left panel, Figure 9). In Florida, FL SRFS estimates were over 50% higher than simulated estimates using the AL Snapper Check, MS Tails n' Scales, and TX CCS designs, but were slightly lower than simulated estimates using the LA Creel

design (top right panel, Figure 9). In Louisiana, LA Creel estimates were systematically higher than all simulated estimates using the other state program designs, but were most similar to the simulated FL SRFS estimates (middle left panel, Figure 9). In Mississippi, MS Tails n' Scales estimates were lower than, but similar to, the simulated estimates using the AL Snapper Check design, much lower (>100%) than simulated

Figure 9: State 2019 red snapper landings estimates under different state program designs. See Appendix C for 2016-2018 comparisons.



estimates using the LA Creel and FL SRFS designs, and 35% higher than simulated estimates using the TX CCS design. In Texas, TX CCS estimates were systematically lower than all the other simulated state program estimates by over 50% (bottom panel, Figure 9). Across all years examined, these within-state patterns varied in terms of the exact differences between the original state program's estimates and the simulated estimates from the other designs. However, LA Creel estimates were systematically highest, and TX CCS estimates were systematically lowest (see Appendix C).

Discussion

Can non-response error in the MRIP FES explain the differences between MRIP and state program estimates?

Changing the non-response assumptions of the FES resulted in lower MRIP red snapper estimates, closing at least some of the gap between MRIP and the state program estimates. However, for non-response error in the MRIP FES to be a primary driver of differences between MRIP estimates and those of the state programs, two assumptions would have to be true: (1) that the FES is susceptible to non-response error while the state programs are not, and (2) the FES suffers from maximum non-response error (i.e., that all non-respondent households have zero fishing activity), since that estimate series was closest in magnitude to the state program estimates, albeit still significantly higher in Alabama and Mississippi.

The first assumption is not supported, given that no program has a 100% response rate, nor does the FES have the lowest response rate of all the Gulf effort surveys (Andrews 2020, Bray 2021 *pers. comm.*). Another complicating factor is that two of the state programs, AL Snapper Check and MS Tails n' Scales, estimate compliance rates by design, which are not the same as, nor directly comparable to, the response rates of the other state surveys and the FES. Compliance rates for AL Snapper Check and MS Tails n' Scales are calculated using the validation data collected by the intercept survey component of those programs (Alabama Department of Conservation and Natural Resources/Marine Resource Division 2018, Mississippi Department of Marine Resources 2017). These analyses rely on a key, untested assumption that compliance rates are representative of trips returning to private access sites (i.e., any access point not on their respective intercept survey frames). It is important to further note that while higher response rates decrease the risk of this error, response rates alone cannot

predict the presence, magnitude, or direction of non-response error (American Association for Public Opinion Research 2016).

The second assumption is also not supported. Two FES non-response follow-up studies found households that responded to the FES, after additional contact attempts and a larger incentive, were not significantly different from base respondents with respect to fishing activity (Andrews et al. 2014, Andrews 2021). Further, additional research evaluating a broader range of non-response assumptions for the FES suggests that non-response bias is likely small (Brick et al., 2021). Our hypothetical results of this study therefore suggest **differences between MRIP and state program estimates cannot be fully explained by MRIP FES non-response error**, and that **other sources of non-sampling error are likely to be contributing to these differences**.

Can coverage error in the MRIP CHTS explain the differences between the MRIP APAIS-CHTS and MRIP APAIS-FES estimates?

Simulating coverage error by excluding non-landline telephone households from the FES sample results in estimates similar to MRIP APAIS-CHTS estimates. The similarity of these estimates suggest that the household phone status question of the FES can approximate the CHTS coverage error. This finding is not surprising, as the FES was specifically designed to address known shortcomings of the CHTS—we know that the CHTS suffered from coverage error in its later years, due to the growing prevalence of wireless-only households since the early 2000s. According to estimates from the National Health Interview Survey, 57% of all American homes do not have landline phones, and an additional 15% can be considered wireless-mostly (Blumberg and Luke, 2018), meaning over 70% have effectively excluded themselves from being sampled by landline-based telephone surveys. From 2000 on, the exclusion of wireless-only households from the CHTS sample frame suggests the CHTS samples were becoming less and less representative of the general population. Landline-only households, for example, report older, less active residents and fewer children (Blumberg and Luke, 2018). These demographic differences are correlated with differences in fishing activity: the landline-only households—the only households that could be reached by the CHTS—are much less likely to report fishing than the general population (Andrews et al., *unpublished data*).

Can coverage error in state surveys explain the differences between MRIP FES and LA Creel, SRFS, and TX CCS estimates?

Whether or not coverage error explains differences among MRIP, LA Creel, FL SRFS, and TX CCS estimates depends upon the assumptions made regarding the

accuracy of these different surveys. As such, we approached the results through two different assumptions: (1) MRIP is accurate, and (2) the state programs are accurate. However, there are two important caveats to these assumptions. The first caveat is that all the surveys are likely to be affected by some level of non-sampling error, and therefore none will be 100% accurate. This is because not even the perfect survey, with a scientifically rigorous design based on sound statistical theory, will be error-free when it's implemented in the real world (Biemer 2010, Dalenius 1977, Murthy 1963). The second caveat is that we cannot assume that all the state surveys are equally accurate due to their differences in design, so the second assumption must be made in phases. We may be able to assume LA Creel and FL SRFS have similar levels of accuracy (since they have similar designs), but we cannot simultaneously assume that TX CCS is accurate, since it has a very different design compared to LA Creel and FL SRFS.

If we make the assumption that MRIP is accurate, negative bias in FL SRFS and LA Creel due to incomplete coverage of unlicensed effort explains much of the differences between MRIP and the state programs estimates. It is important to note that LA Creel and FL SRFS do implement coverage adjustments derived from questions in their intercept surveys about license status (Florida Fish and Wildlife Conservation Commission 2018, Louisiana Department of Wildlife and Fisheries, 2017). In theory, these coverage adjustments should have made FL SRFS and LA Creel estimates higher than the simulated MRIP license-only domain estimates. However, survey research suggests sensitive questions (e.g., asking about 'rule-breaking' behavior, such as if an angler is fishing with or without a license) can often lead to bias due to misreporting. The literature on this topic suggests the level of misreporting that occurs is highly situational, depending upon factors like the status of the respondent (e.g., whether they're rule-breaking or not), how sensitive they perceive the question, their sense of anonymity, and/or their level of trust in the interviewer and/or entity administering the survey (see e.g., St John et al. 2010, Tourangeau and Yan 2007). In addition to license status sensitivity, there may also be differences in license status among anglers using public versus private access sites. Any coverage adjustment calculated using intercept survey data would itself be subject to coverage error depending on how different anglers using public access are from those using private sites, and how many individual fishing trips are taken at public versus private sites. This potential source of error in FL SRFS and LA Creel warrants further investigation. **In regards to Texas, negative bias in TX CCS due to a lack of coverage of private access effort (and other off-frame effort such as trips ending at night) would explain the majority of the differences between MRIP and TX CCS estimates.** In some areas, private access sites may outnumber public access sites, and could account for substantial numbers of fishing trips.

If we make the alternative assumption that the state estimates are accurate, MRIP would have to have one or more unidentified non-sampling errors in the MRIP FES (alone or combined with possible, but unlikely, systematic non-response error) creating a large positive bias. Of the state surveys, if FL SRFS and LA Creel are assumed to be most accurate, then this positive bias in MRIP must be so large that it offsets the potential coverage error that we've examined in this study (i.e., the MRIP license-only series, which creates known coverage error by excluding large amounts of data, is actually accurate rather than biased). If TX CCS is assumed to be most accurate, it would mean the FES effort estimates should be much lower (equal to the amount of effort covered by the APAIS), and also might suggest that FL SRFS and LA Creel also overestimate fishing effort, since they also rely on offsite effort surveys. This scenario, while possible, is unlikely, since it would effectively mean that no fishing is occurring at either private sites or any time periods (e.g. night) not covered by the intercept survey. In fact, an inventory of boat access sites conducted in Texas suggested that TX CCS likely underestimates total landings due to the number of non-surveyable, private access sites in the state (Spiller 1987).

Differences in trip characteristics between public and private access sites could also be a source of error in the MRIP estimates, since MRIP only covers private access fishing (effort) through the FES and not through APAIS. MRIP therefore assumes that the fishing activity observed via the APAIS (at public access sites) is representative of the fishing activity reported through the FES (at both public and private access sites), which may or may not be the case. However, LA Creel and FL SRFS make the same general assumption with their intercept surveys and off-site effort surveys, so this error would only drive differences between MRIP and TX CCS. Related to the public access vs. private access issue, another possible bias could exist in MRIP proportions of trips by area fished, which MRIP staff are currently studying and, if discovered, might help explain more of the differences between MRIP and state estimates (results are expected in 2022).

What non-sampling errors might explain differences between MRIP, MS Tails n' Scales, and AL Snapper Check?

Because the AL Snapper Check and MS Tails n' Scales survey designs are very different from MRIP, we couldn't effectively simulate potential sources of error in the MRIP estimates that directly align with design aspects of these state programs. However, if we assume MRIP is accurate, our results suggest that non-sampling errors in MS Tails n' Scales and AL Snapper Check may be of a similar magnitude to that of the APAIS-only series (and TX CCS). If we assume AL Snapper

Check and MS Tails n' Scales are accurate, unidentified non-sampling error in MRIP has a large upward bias that offsets the large downward bias introduced by the APAIS-only series (i.e., the MRIP APAIS-only series, which excludes all FES data, is actually accurate rather than biased). This second scenario is theoretically possible, but may be unlikely given the issues discussed above regarding private access and other possible fishing activity not covered by intercept survey frames.

Possible sources of non-sampling error in the AL Snapper Check and MS Tails n' Scales estimates could include coverage error related to differences in reporting and/or trip characteristics between public and (non-sampled) private fishing access sites—but this, again, is an issue that affects most of the surveys operating in the Gulf. Error in both programs could further result from violations of the assumption of independence between the capture and recapture phases of data collection. Both AL Snapper Check and MS Tails n' Scales use a capture (self-reported trips) and recapture (trips intercepted by samplers meant to validate the self-reports) design, which relies on the assumption that capture is independent of the recapture process. Putting this assumption in other words, whether or not an angler self-reports, and how many fish they report, must be unaffected by whether or not the trip is intercepted in the validation sampling. If this assumption is violated (i.e., if intercepted anglers are more likely to report their fishing activity after being intercepted, or are more likely to report trip and catch information accurately), it can introduce bias (Brick 2018). MS Tails n' Scales requires pre-declaration of trips (Mississippi Department of Marine Resources 2017), which does address independence regarding whether a trip is reported, but may still be susceptible to errors regarding what is reported for the trip. AL Snapper Check could be susceptible to independence violations for both trips being reported and trip information being reported accurately (this topic is being studied by statistical consultants, with results expected in 2022).

Are state-level estimates comparable across a regional design (even when that design has known non-sampling errors?)

While the magnitude or scale of the estimates differ when using different common regional designs (due to differences in non-sampling errors), the individual state-to-state distributions of estimates, and each state's proportional contribution to the regional total estimate, are relatively consistent across common designs. Even the intercept-only design, which in all likelihood suffers from a large downward bias, showed a consistent pattern comparable to the license-only and MRIP designs. **These findings suggest that the state-level estimates are comparable across a regional design, and the consistent pattern in distributions we see when using a common design across**

the region is likely accurate. However, the use of different survey designs, such as individual state programs in every state, appears to introduce differences among the state estimates, potentially masking the “true” regional pattern. If we assume the trends observed when using the common designs are accurate, Florida’s and Louisiana’s proportional contribution to the regional red snapper estimates should be lower than it is at present using individual state monitoring, and Alabama’s and Texas’s proportional contributions to the regional red snapper estimates should be higher. Interestingly, Mississippi’s proportional contribution to the regional estimate is relatively consistent (2-6%) no matter which survey design is used, across all years examined in this study. This pattern may be a function of the Mississippi red snapper fishery being small relative to the other Gulf states.

Are the estimates produced by all the different state survey designs comparable?

In comparing the individual state surveys to each other, results suggest that differences in survey designs likely drive systematic differences in the estimates, and they are therefore not directly comparable to one another. However, these differences were generally smaller when calibrating estimates derived from surveys with similar designs. For instance, calibrating FL SRFS to LA Creel, or vice versa, resulted in smaller changes in estimates (a maximum of a 34% change in 2019, but between 5-14% in all other years), which makes sense since the two surveys use a similar sample frame of licensed anglers to collect effort data. We expected to see similar results calibrating AL Snapper Check to MS Tails n’ Scales, or vice versa, because both programs use mandatory angler apps coupled with a validation survey. However, the results were more variable, showing small changes in some years, but larger changes (up to 85%) in other years, so there may be unknown sources of non-sampling errors driving MS Tails n’ Scales and AL Snapper Check estimates apart. Alternatively, the variable differences between MS Tails n’ Scales and AL Snapper Check may have resulted from increasing angler awareness and compliance or design improvements made in these first few years of the programs. That being said, differences between AL Snapper Check and MS Tails n’ Scales estimates were far less than those among surveys with greater design differences. For example, in most years, Texas landings were very different, on average, from TX CCS when calibrated to any of the other state survey programs.

Conclusions

While non-sampling errors exist in every large-scale, real-world data collection program, the direction and magnitude of the impact of such errors depend on the survey design. The results of this preliminary analysis indicate:

- Non-sampling errors can cause large systematic biases in estimates of recreational catch and may be driving the differences between estimates of red snapper landings in the Gulf of Mexico. However, drivers are dependent upon assumptions made regarding the accuracy of the different surveys.
- Simulating common or standardized survey designs across the five Gulf states produces similar patterns in the relative distribution of red snapper landings estimates across each state, strongly suggesting the value of a common regional design.
- Simulating the side-by-side conduct of each state survey design across the Gulf of Mexico produces different estimates for each state, suggesting individual state program estimates are not directly comparable and reinforcing the need for calibration to ensure comparability of estimates produced from programs using different designs.

Our results do not rule out the possibility of unidentified non-sampling errors in any of the region's data collection programs that, once quantified and understood, could support or refute these findings. Further investigation into potential sources of non-sampling error is needed to improve the accuracy and precision of all available red snapper data, to improve the calibration methodologies that may be used to reconcile differences between available estimates, and to allow for the possible integration of estimates from different survey programs into a single set of comparable estimates for the region. We encourage continued research, collaboration, and information sharing among state and federal partners in working toward our common goal: to produce high-quality estimates to inform the assessment and management of sustainable fish populations.

References

- Alabama Department of Conservation and Natural Resources/Marine Resources Division. (2018). *Alabama Snapper Check: Summary of Estimation Procedures*. Retrieved from <https://media.fisheries.noaa.gov/dam-migration/snapper-check-decision-memo-with-attachments.pdf>
- American Association for Public Opinion Research. (2016). *Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys*. Retrieved from https://www.aapor.org/AAPOR_Main/media/publications/Standard-Definitions20169theditionfinal.pdf
- Andrews, W. R. (2021). *Evaluating Nonresponse Bias in the MRIP Fishing Effort Survey*. Office of Science and Technology. NOAA Fisheries. Silver Spring, MD. Retrieved from https://apps-st.fisheries.noaa.gov/pims/main/public?method=DOWNLOAD_FR_D ATA&record_id=2018
- Andrews, W. R. (2020). *Fishing Effort Survey Annual Report*. NOAA Fisheries. Silver Spring, MD. Retrieved from <https://www.fisheries.noaa.gov/resource/document/annual-report-fishing-effort-survey>
- Andrews, W. R., Brick, J. M., & Mathiowetz, N. A. (2014). *Development and Testing of Recreational Fishing Effort Surveys: Testing a Mail Survey Design*. NOAA Fisheries. Silver Spring, MD. Retrieved from https://www.st.nmfs.noaa.gov/Assets/recreational/pdf/2012-FES_w_review_and_comments_FINAL.pdf
- Biemer, P.P. Total Survey Error: Design, Implementation, and Evaluation, *Public Opinion Quarterly*, Volume 74, Issue 5, 2010, Pages 817–848, <https://doi.org/10.1093/poq/nfq058>
- Blumberg, S. J., & Luke, J. V. (2018). *Wireless Substitution: Early Release of Estimates From the National Health Interview Survey, July–December 2018*. Retrieved from <https://www.cdc.gov/nchs/data/nhis/earlyrelease/wireless201906.pdf>
- Bray, G. (2021). [Gulf State Annual Response/Compliance Rates]. Pers. Comm.
- Brick, J. M., Andrews, & W. R., Foster, J. (2021) *Two Sources of Nonsampling Error in Fishing Effort Surveys*. Manuscript submitted for publication.
- Brick, J. M. (2018). *Electronic Reporting in Survey Research Applied to Estimating Fishing Effort*. Retrieved from https://media.fisheries.noaa.gov/dam-migration/electronic_reporting_in_survey_research_applied_to_estimating_fishing_effort.pdf

- Dalenius, T. (1977). Bibliography on Non-Sampling Errors in Surveys: I (A to G). *International Statistical Review / Revue Internationale de Statistique*, 45(1), 71–89. <http://www.jstor.org/stable/1403005>.
- Dillman D. A., & Christian L. M. Survey Mode as a Source of Instability in Responses across Surveys. *Field Methods*. 2005;17(1):30-52. doi:[10.1177/1525822X04269550](https://doi.org/10.1177/1525822X04269550)
- Florida Fish and Wildlife Conservation Commission. (2018). *Gulf Reef Fish Survey Certification Review Documentation*. Retrieved from https://media.fisheries.noaa.gov/dam-migration/09_gulf-reef-fish-survey-decision-memo-with-attachments.pdf
- House Committee on Appropriations (2021). Joint Explanatory Statement of the Committee of Conference on H.R. 133: Title I, Division B. House Rules Committee on Senate Amendment to the Consolidated Appropriations Act of 2021. *Congressional Record* 166:218 (Book III). <https://www.govinfo.gov/content/pkg/CREC-2020-12-21/pdf/CREC-2020-12-21-house-bk3.pdf>
- Louisiana Department of Wildlife and Fisheries. (2017). *Saltwater Finfish Landing Statistics*. Retrieved from <https://media.fisheries.noaa.gov/dam-migration/la-creel-decision-memo-with-attachments.pdf>
- Mississippi Department of Marine Resources. (2017). *Mandatory Red Snapper Reporting Program: 2016 Methods and Results*. Retrieved from https://media.fisheries.noaa.gov/2020-09/07_Tails-n-Scales-Decision-Memo-with-Attachments.pdf
- Mississippi Department of Marine Resources. (2021). Red Snapper. Retrieved from <https://dmr.ms.gov/snapper/>
- Murthy, M. N. (1963). Assessment and Control of Non-Sampling Errors in Censuses and Surveys. *Sankhyā: The Indian Journal of Statistics, Series B (1960-2002)*, 25(3/4), 263–282. <http://www.jstor.org/stable/25051490>
- Papacostas, K. J., & Foster, J. (2018). *Survey Design and Statistical Methods for Estimation of Recreational Fisheries Catch and Effort*. NOAA Fisheries Marine Recreational Information Program. Retrieved from Silver Spring, MD: <https://www.fisheries.noaa.gov/resource/document/survey-design-and-statistical-methods-estimation-recreational-fisheries-catch-and>
- Phung, T.D., Hardeweg, B., Praneetvatakul, S., & Waibel, H. (2015) Non-Sampling Error and Data Quality: What Can We Learn from Surveys to Collect Data for Vulnerability Measurements? *World Development*, 71, 25-35, <https://doi.org/10.1016/j.worlddev.2013.11.008>.
- Spiller, K. W. (1987). *Inventory of Boat Access Sites on the Texas Coast*. Texas Parks and Wildlife Department, Coastal Fisheries Branch. Austin, TX. Retrieved from

https://tpwd.texas.gov/publications/pwdpubs/media/mds_coastal/Series%201_MDS110.pdf

St. John, F. A. V., Edwards-Jones, G., Gibbons, J. M., & Jones, J. P. G. (2010). Testing novel methods for assessing rule breaking in conservation. *Biological Conservation*, 143(4), 1025-1030.

doi:<https://doi.org/10.1016/j.biocon.2010.01.018>

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859-883. doi:10.1037/0033-2909.133.5.859.

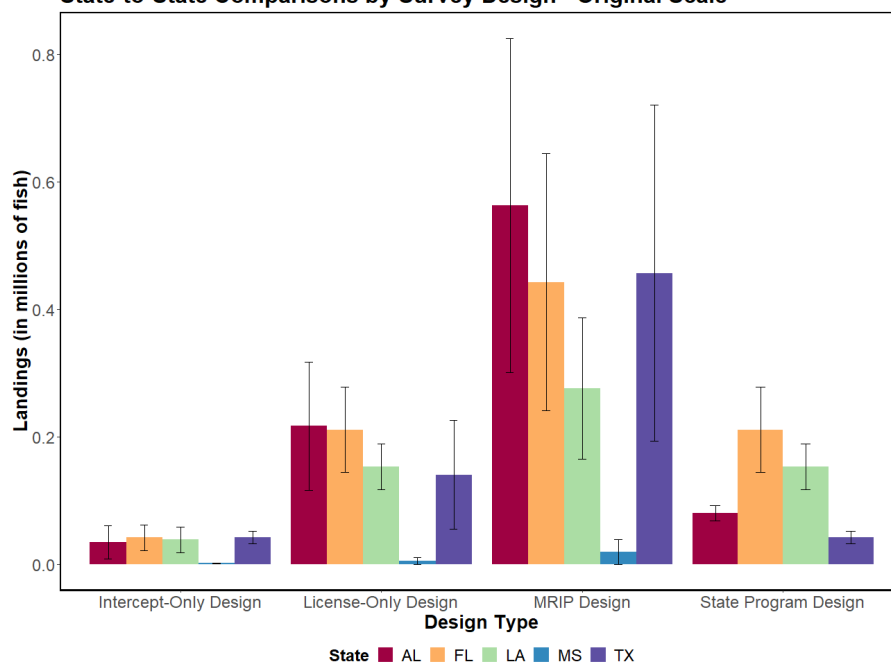
Appendix A

State-to-state differences in red snapper landings estimates under different survey designs (2015-2018)

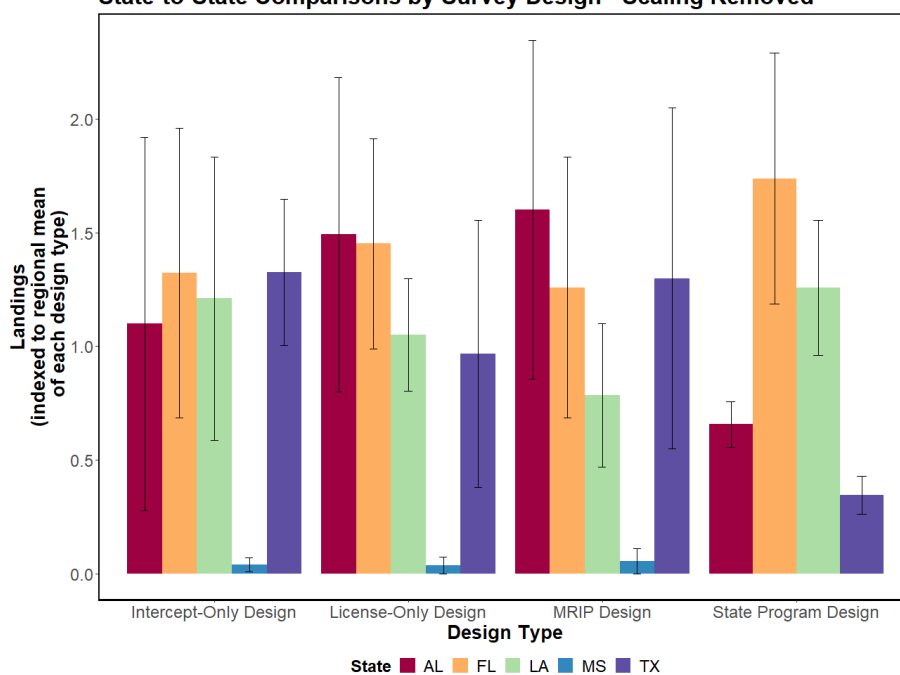
2015:

(Note: MS is missing from “State Program Design” bars because 2015 MS Tails n’ Scales estimates were unavailable)

State-to-State Comparisons by Survey Design - Original Scale

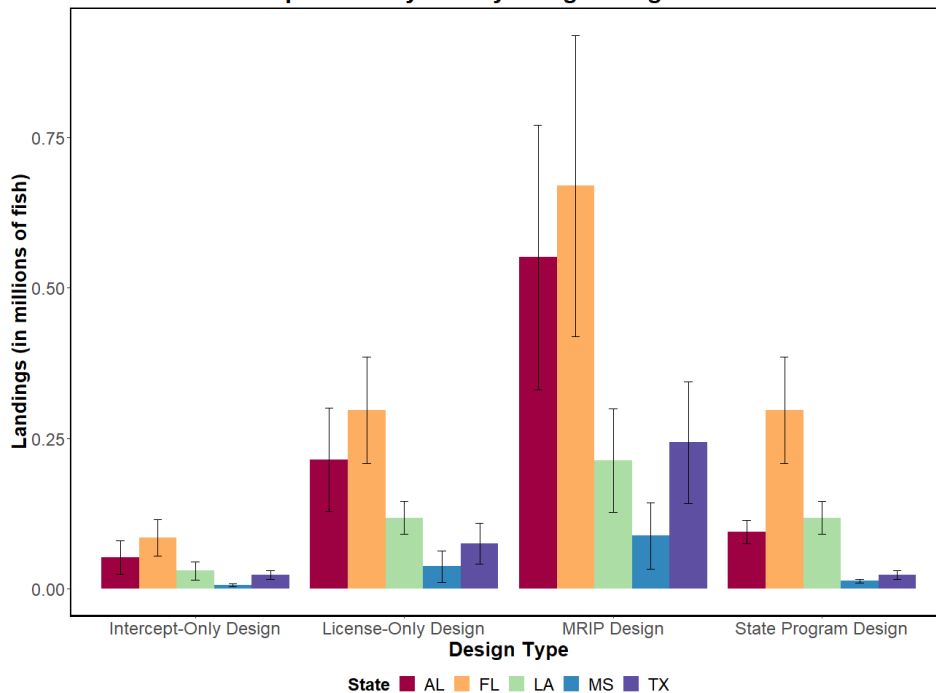


State-to-State Comparisons by Survey Design - Scaling Removed

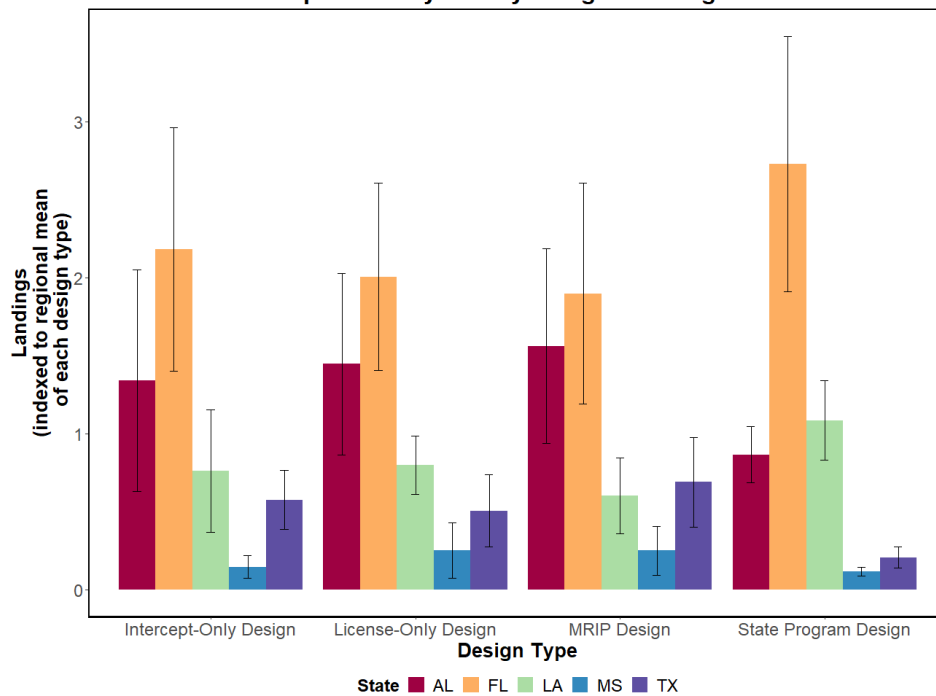


2016:

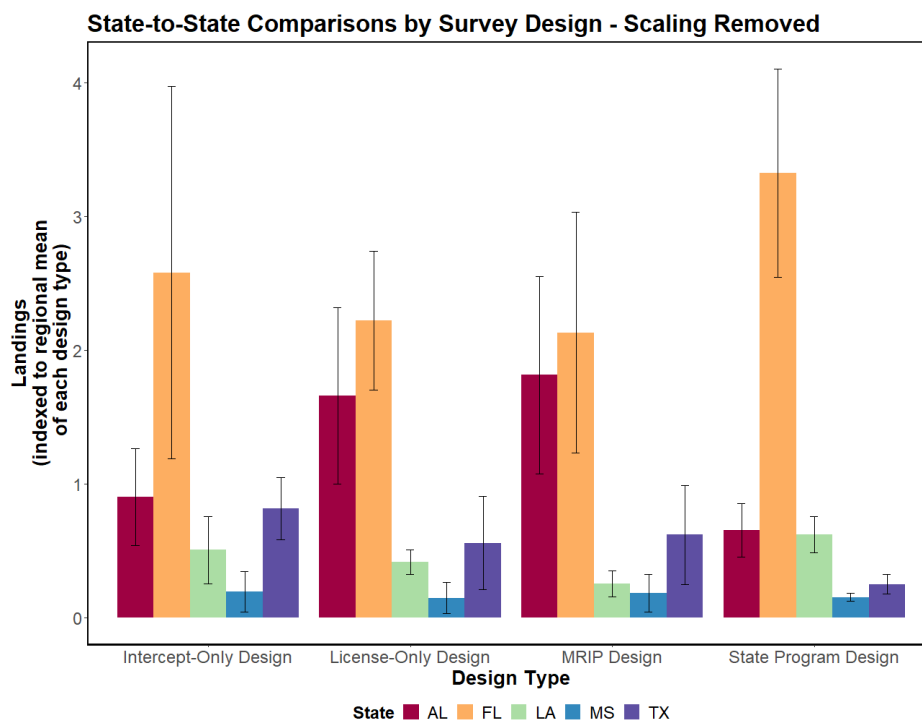
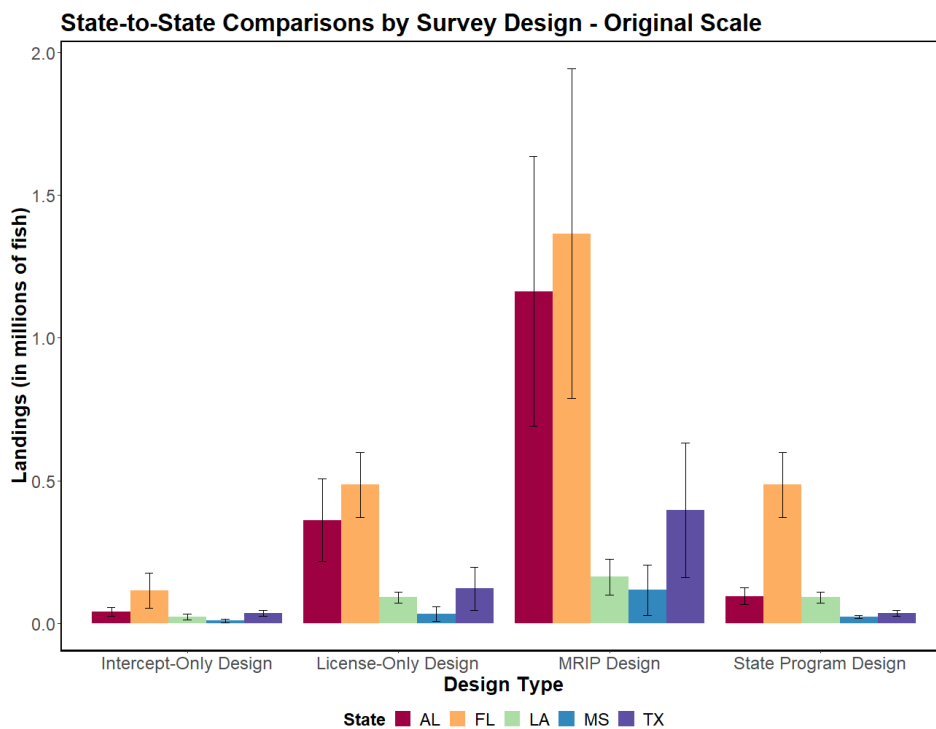
State-to-State Comparisons by Survey Design - Original Scale



State-to-State Comparisons by Survey Design - Scaling Removed

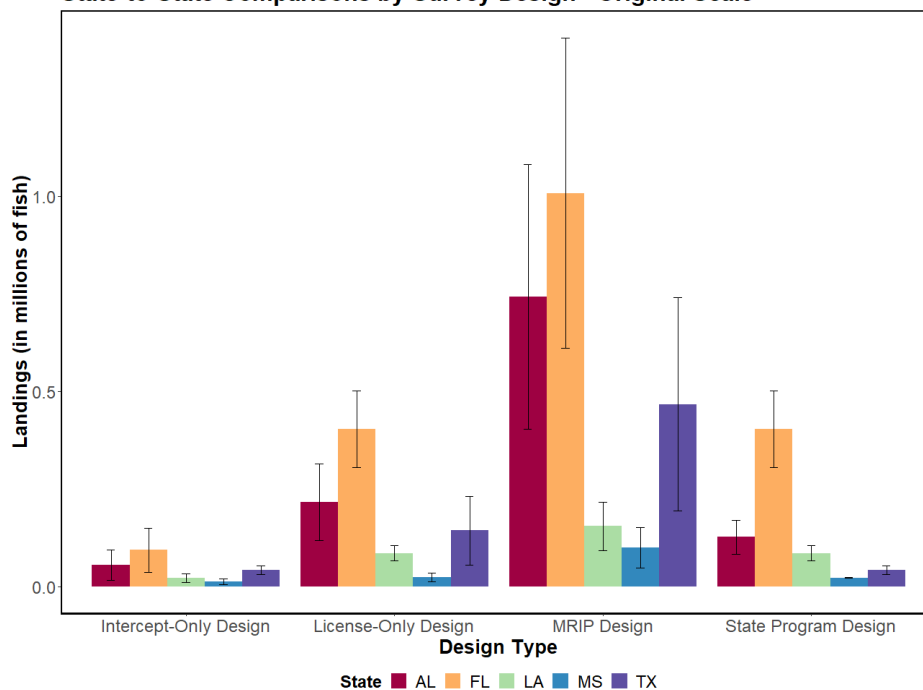


2017:

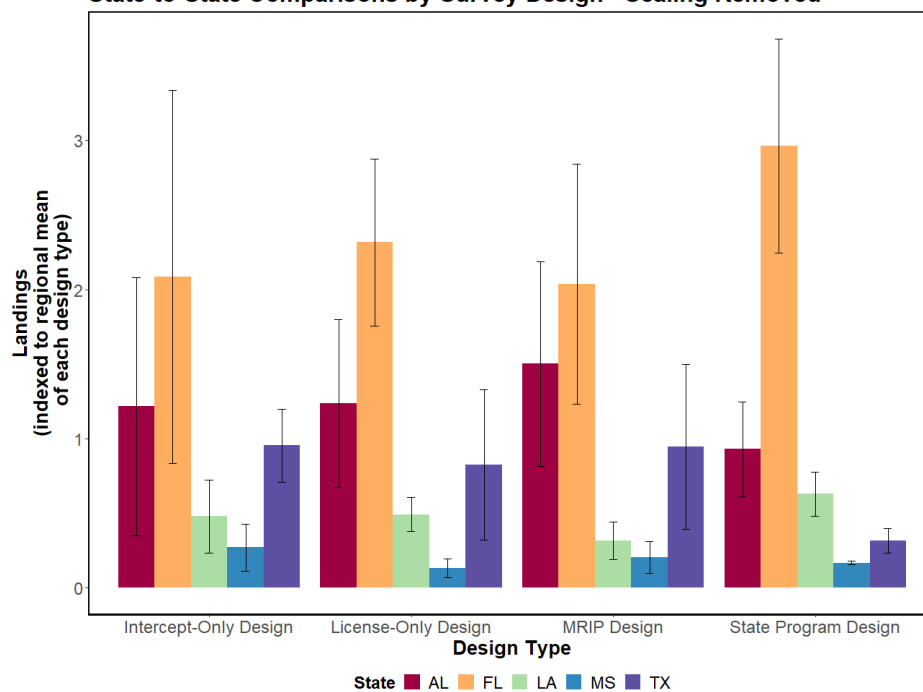


2018:

State-to-State Comparisons by Survey Design - Original Scale



State-to-State Comparisons by Survey Design - Scaling Removed

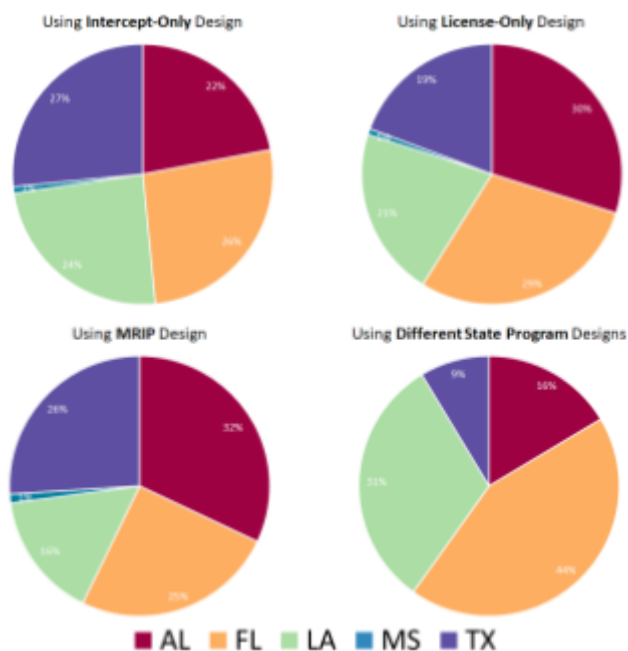


Appendix B

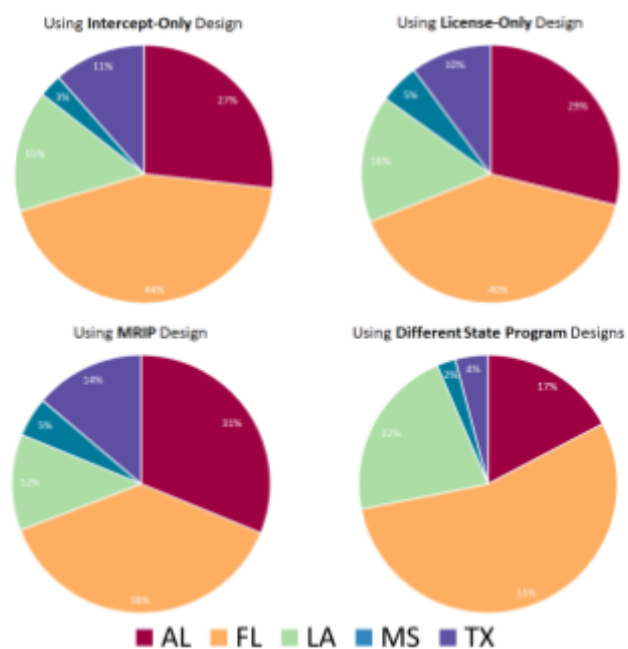
State contributions to the regional red snapper landings estimate under different survey designs (2015-2018)

2015:

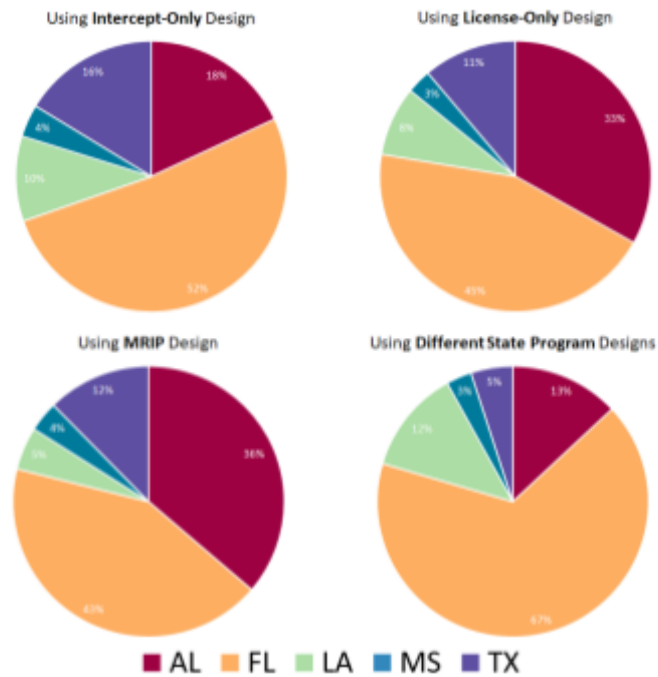
(Note: MS is missing under the "Different state program designs" heading because 2015 MS Tails n' Scales estimates were unavailable)



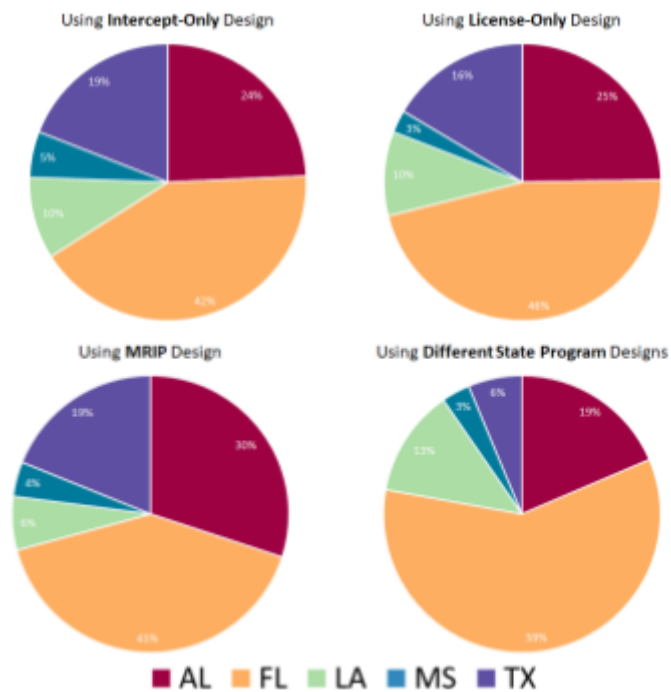
2016:



2017:



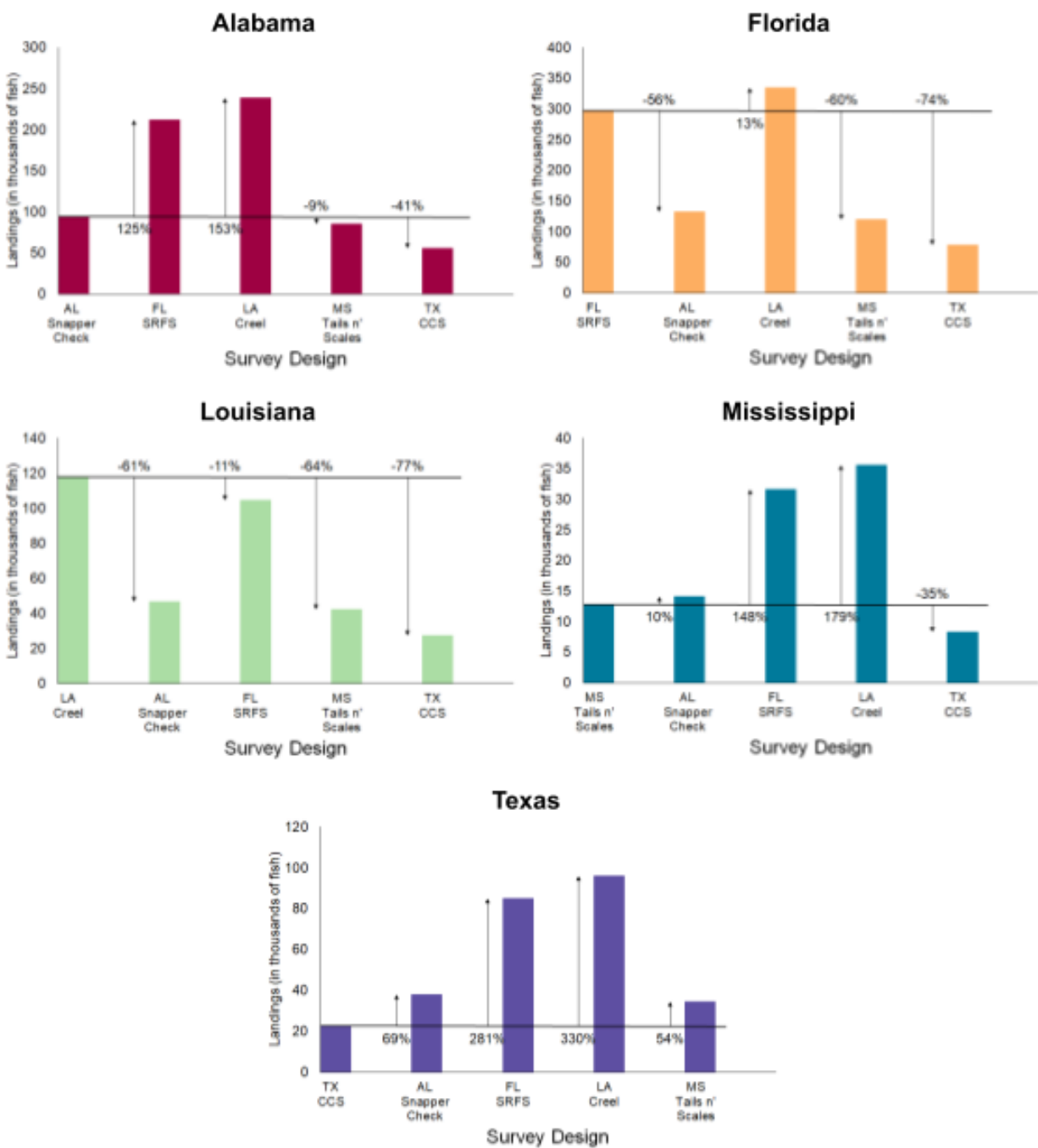
2018:



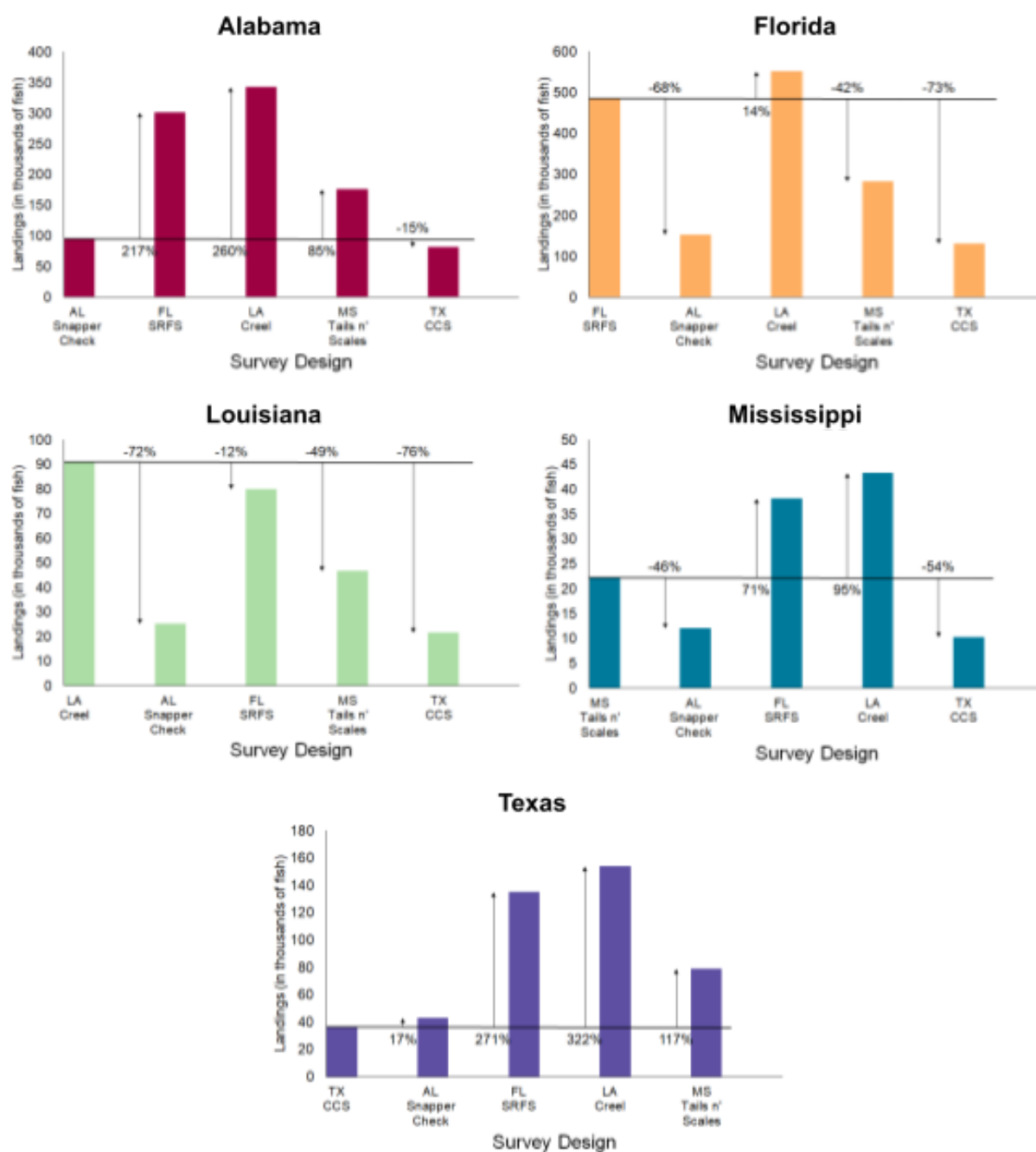
Appendix C

State red snapper landings estimates under different state program designs (2016-2018).

2016:



2017:



2018:

