Amending survey design to improve statistical inferences: monitoring recruitment of juvenile reef fish in the eastern Gulf of Mexico

Meagan N. Schrandt*, Colin P. Shea, Benjamin N. Kurth, Theodore S. Switzer

Fish and Wildlife Research Institute, Florida Fish and Wildlife Conservation Commission, St. Petersburg, FL 33701, USA

* Corresponding author.
E-mail address: Meagan.Schrandt@MyFWC.com (M.N. Schrandt)

ABSTRACT

Many species of interest to management and conservation remain data-limited, and the data that are available are often unable to produce statistically reliable population trends. We examined 10 years of juvenile reef fish catch data from two gear-specific, fishery-independent surveys in the eastern Gulf of Mexico to assess our ability 1) to characterize population trends for various reef species sampled using haul seines or otter trawls and 2) to amend our survey, as logistics would allow, to improve those characterizations. Species richness, catch-per-unit-effort, and length-frequency distributions were generally similar between gear types, suggesting a single survey (i.e., one gear type) may be sufficient. Simulation-based power analyses for the reef fish species indicated that overall, otter trawl data provided greater power for detecting trends over a 10-year period (41–75% probability to detect a 50% change in abundance) than haul seine data (27–52%). Likewise, otter trawl data provided greater power for detecting trends from one year to the next (22–53% probability to detect a 50% change in abundance in one year) than haul seine data (15–33%). Simplifying data collection from two surveys to a single, trawl-only survey (and approximately doubling the number of trawl tows) was more efficient, more effective, and more powerful (64–91% for the amended design with more tows) in detecting abundance trends. Furthermore, by increasing sample size but retaining all other design elements, the data collected during the trawl survey before and after the change remain comparable; the time series was not interrupted. These changes increase our confidence in estimating population trends, predicting productivity, and informing management and conservation decisions.

Keywords:

Fishery-independent monitoring

Population dynamics

Statistical power

Stock assessment

Trend analysis

1. Introduction

Discerning species' population trends is imperative for critical management and conservation. Since the allocation of limited resources for monitoring, conservation, and management can be influenced by population trends, decision makers should be able to evaluate and trust an apparent trend of a population (Wauchope et al., 2019). This ensures that at-risk species get the attention they need and prevents allocating limited conservation resources to species that are not at risk (Wauchope et al., 2019). Population dynamics can be highly variable, so the collection of trustworthy, statistically powerful data can be resource-intensive and, therefore, limited. For example, White (2019) examined over 800 populations of vertebrate species to assess the number of years needed to estimate a linear population trend and found that to be able to predict a population change within 2% change per year, data sets needed to cover between 5 and 30 years, depending on the species. For most fish populations in White's (2019) analysis, at least 15 years of monitoring data were needed to accurately detect a linear trend; for some species, more than 30 years of data were needed. Not surprisingly, then, more than half (59%) of the federally managed fisheries species in the United States have been considered data-limited for the purpose of setting management targets as mandated by federal law (Newman et al., 2015).

Indices of juvenile abundance are an important component of stock assessments and have also been used to predict recruitment to the fishery for several species (e.g., Stige et al., 2013; Wertheimer et al., 2016). The ability to predict the strength of recruitment to the fishery enables stock managers to better assess management actions and their outcomes (Hansen et al., 2015). Contemporary stock assessments may lack sound estimates and predictions of juvenile recruitment, however, which are essential to managing at-risk stocks and forecasting fishery productivity (Smith, 1993; Koenig and Coleman, 1998; Coleman et al., 1999; Johnson and Koenig, 2005). For example, Switzer et al. (2012) demonstrated significant interannual variation in indices of recruitment for juvenile *Mycteroperca microlepis* (gag) in the eastern Gulf of Mexico using data from a long-term monitoring survey of Florida's estuaries (McMichael, 1991). High levels of variation suggest insufficient statistical power to accurately detect temporal population trends. Therefore, it would be beneficial to periodically evaluate the performance of monitoring programs and assess their ability to collect statistically powerful, long-term monitoring data. Gibbs et al. (1998) suggested conducting power analyses for a monitoring program during its pilot phase. Power analyses could be performed at other critical times as well, such as in reapplications for funding or the search for new funding sources. Re-evaluations also allow researchers and managers to assess the existing study design and potentially improve upon it.

To provide less variable and statistically powerful indices of juvenile abundance for stock assessments for multiple reef fish species in the eastern Gulf of Mexico, the Florida Fish and Wildlife Conservation Commission's Fish and Wildlife Research Institute's Fisheries Independent Monitoring (FIM) program began a multispecies survey of polyhaline seagrass beds in five estuaries along the West Florida Shelf (WFS) (henceforth referred to as the WFS juvenile reef fish survey). The WFS juvenile reef fish survey monitors juvenile recruitment of reef fish species that use nearshore seagrass beds during their early life history (e.g., snappers, groupers) before they transition as adults to offshore reef habitats (e.g., Koenig and Coleman, 1998; Nagelkerken et al., 2001, 2002; review by Gillanders et al., 2003; Casey et al., 2007; Switzer et al., 2012; Lefcheck et al., 2019). Focusing the juvenile reef fish abundance survey on the preferred juvenile habitat (polyhaline seagrass beds) can improve estimates of juvenile abundance. For example, initial analyses for the recruitment of juvenile *Lutjanus griseus* (gray snapper) indicated that sampling polyhaline seagrass beds improved catch rates, increased frequency of occurrence, and reduced the coefficient of variation for catch-per-unit-effort (Flaherty-Walia et al., 2015). Although the WFS juvenile reef fish survey improved recruitment estimates, we are also interested in whether the survey can be further refined to improve the statistical power to detect temporal trends in abundance of juveniles.

Incorporating statistical strength into survey designs not only improves the scientific data upon which management decisions are based, but it can also improve sampling efficiency and reduce costs. Therefore, after 11 years of using two gear-specific surveys (haul seine and otter trawl) to sample juvenile reef fish in the seagrass habitats of the eastern Gulf of Mexico, our primary objective was to evaluate the statistical strength of each gear-specific survey in in detecting temporal trends in abundance. We were interested in the juveniles of reef species captured in the sampled estuaries, namely *Centropristis striata* (black sea bass), *Epinephelus morio* (red grouper), *Haemulon plumierii* (white grunt), *Lachnolaimus maximus* (hogfish), *Lutjanus griseus*, *L. synagris* (lane snapper), and *Mycteroperca microlepis*.

2. Material and methods

2.1. Survey design

The WFS juvenile reef fish survey originally consisted of two gear-specific surveys (haul seine and otter trawl) that provide indices of juvenile abundance for a suite of species. The survey samples seagrass beds with at least 50% cover of submerged aquatic vegetation, in 1.2–2.5 m of water, with a sloping bottom, in polyhaline to euhaline waters (i.e., salinity ≥18). For both gear types, monthly stratified random sampling was conducted in polyhaline seagrass beds located ≥100 m from the shoreline in five estuaries in the eastern Gulf of Mexico (Saint Andrew Bay [SA], Apalachicola Bay [AP], Big Bend region [BB], Tampa Bay [TB], and Charlotte Harbor [CH]; Fig. 1). These estuaries are contiguous with the WFS. Most are shallow, semi-enclosed estuaries separated from the Gulf of Mexico by barrier islands and connected to it via various channels and passes. The exception is the Big Bend region, an open estuary where the low-relief coastline functions as an estuary in which extensive freshwater sheet flow from land mixes directly with marine waters of the Gulf (Geselbracht et al., 2015). The survey began in 2008 and although it has been refined temporally and spatially to improve juvenile reef fish catch and reduce variability in abundance estimates, the underlying protocols and stratified random sampling design have remained unchanged. Therefore, we restricted our analyses to samples from the two gear types during June through November, 2008–2018, in sampling zones [of each estuary] that were continuously sampled during those years (Schrandt et al., 2021).

The number of monthly deployments for each gear type differed by estuary, depending on the coverage of seagrass beds that met our sampling criteria (see section 2.2) and logistical constraints. Fewer monthly haul-seine sets were made than trawl tows (Table 1), and the number of monthly deployments among estuaries ranged from 3 to 7 for haul seines and 6 to 30 for trawls; no haul-seine samples were taken in the BB region. This resulted in N=120 haul seine sets and N=372 trawl tows during a 6-month period (i.e., sampling year) in the eastern Gulf of Mexico.

2.2. Field sampling

The 183-m × 3-m center-bag haul seine with 38-mm stretched nylon mesh netting was used to collect large-bodied (>100 mm SL) fishes associated with shoal seagrass beds (i.e., seagrass beds ≥100 m from shore with sloped bathymetry). Haul seines were set by boat in a rectangular shape along the shoal in ≤0.5 m of water, where wings were located on the shoal in ≤0.5 m of water and the bag was located further down the slope of the shoal in 1.0–2.5 m of water. Haul seines were retrieved by hand by personnel at the end of each wing. The 6.1-m otter trawl with 38-mm mesh and 3.2-mm mesh liner was used to collect individuals of various sizes but generally from deeper seagrass beds than the ones sampled with the haul seine. Trawls were towed in seagrass beds in 1.0–7.6 m of water. For both types of gear, sampling sites were limited to polyhaline (>18) waters with ≥50% bottom coverage of submerged aquatic vegetation, confirmed on-site prior to gear deployment. Sampling was done during daylight hours (one hour after sunrise to one hour before sunset). Detailed descriptions of haul seine and trawl sampling in polyhaline seagrass bed habitats in the eastern Gulf of Mexico can be found in De Angelo et al. (2014) and Schrandt et al. (2018), respectively.

All fish and selected invertebrates were identified to the lowest possible taxon and counted. For a subset of each, standard length (SL) was measured to the nearest mm. Catch-per-unit-effort (CPUE) for each species is presented for each sample as the number of fish per 100 m^2. Effort for haul seines was consistent at approximately 4,120 m^2 because deployment was standardized. Effort for trawls depended on the distance towed and was calculated by multiplying the fishing width of the net by the distance towed (distance between start and end GPS coordinates). A standard tow per our procedures was a 0.1 nautical mile tow at a speed of 1.2 knots (tows lasted ca. 5 minutes), sampling approximately 720 m^2. Tows were considered nonstandard when bycatch (e.g., algae, tunicates) was abundant and prevented safe retrieval of the trawl, in which case the tow was aborted and tried again for 3 minutes. If bycatch remained high, the tow was re-tried for 2 minutes. Regardless of whether a tow was considered standard or nonstandard, effort was calculated from the recorded tow distance.

Water quality parameters were measured (temperature in °C, salinity, and dissolved oxygen in mg/L) at each site from the surface to the bottom in 1.0-m depth intervals. The average value for the water column of each parameter was used for informing the models for power analyses.

2.3. Catch information

To assess the taxa collected with each gear type, we compared the following metrics: 1) the average taxon richness per set; 2) the overall taxon richness by gear type; 3) the number of taxa unique to each gear type; 4) the CPUE of each reef species of interest; and 5) the length-frequency distribution (for the overlapping size range of the two gear types) of each reef species of interest. The first three comparisons allowed us to evaluate the overall effectiveness of each gear type in a multispecies survey; the last two focused on the information provided by the survey specifically for reef species, which were the focus of the subsequent power analyses (section 2.4).

2.4. Power analyses

Fisheries managers have various overarching topics of interest, including long-term trends of managed populations and the effects of perturbations on managed populations; we addressed both with our power analyses. To assess our ability to correctly detect population trends (negative or positive) over a 10-year sampling period using the original sampling design (Design 1 [two gear-specific surveys]; Table 1), we conducted separate power simulations for each combination of reef species of interest and gear type. We did not combine gear types into a single model because the two indices of abundance are not always combined into a single index for Gulf of Mexico stock assessments, and we wished to determine which gear type provided the most statistically significant description of trends detected. Initial assessments of the observed counts (catch data) for all species of interest indicated evidence of overdispersion, which is common for ecological data; therefore, in our simulations, we used negative binomial regression to model our count data (e.g., White and Bennets, 1996; Bolker, 2008; Zuur et al., 2009). Additionally, we accounted for effort differences among trawl samples by including effort as an offset in the model. Assessments of goodness of fit of the negative binomial regression models indicated no evidence of substantial lack of fit based on tests of uniformity of scaled model residuals and tests of zero-inflation; hence, we assumed that predictions based on the negative binomial regression models fit to the 2008–2018 catch data provided a reasonable representation of average CPUEs in each bay for all species over that period.

We separated our 2008–2018 catch data into a separate data set for each combination of species and gear type and fit a negative binomial model to each. The number of fish of each species caught per seine or trawl deployment was the response variable. Predictor variables included bay as a categorical variable, year (as an integer ranging from 1 to 10), and percent coverage of bottom vegetation, salinity, temperature, and dissolved oxygen as continuous predictors. We also included a year × month random effect (i.e., we considered each combination of year and month as a distinct group). With the exception of year, all continuous predictors were standardized to mean 0 and standard deviation of 1, such that the model intercept and coefficients associated with bay and year were interpreted as expected counts under average vegetation cover, salinity, temperature, and dissolved oxygen conditions. For trawls, effort was included as an offset because effort varied with distance towed. In all models, we also included a term for the interaction between year and bay. All models included bay as a predictor variable in the dispersion component of the model, which allowed counts from each bay to exhibit a different degree of variability with respect to the mean. Following model fitting, we extracted the mean (maximum for haul seines due to relatively low counts in haul seines) predicted count among years for each bay, as well as the bay-specific overdispersion parameters; the predicted counts provided the starting values (expected counts in each bay) for the simulations. Finally, we assessed goodness of fit for the model by testing for uniformity of scaled residuals and evidence of zero inflation, using the R package "DHARMa" (Hartig, 2020).

Next, we calculated various levels of annual percent change in abundance (given starting abundances for each bay, estimated above) over a set time period to simulate exponential (linear on the log scale but exponentiated from the log scale) decay (negative trends) or increase (positive trends) in bay-specific populations over a 10-year period. We chose a 10-year period to better represent the available data (11 years) and because it is desirable to be able to detect true population trends as quick as possible for fisheries management decisions. The various levels of percent annual change resulted in year-10 populations ranging from 0.10 to 0.90 of the initial population for negative trends and 10 to 1.10 times the initial population for positive trends. We realize the exponential changes are only realistic to a certain point due to constraints on population growth, but we have constrained our total decay or growth trends to a realistic change. Under the various levels of percent annual change, we simulated sampling (number of nets in each estuary for 6 months over 10 years), adding error to the sampling process by drawing simulated samples from a negative binomial distribution with means (expected counts that decrease or increase, year after year) and overdispersion parameters from the initial negative binomial model fit to observed data. For each simulated data set, we fit a negative binomial model with year as a continuous predictor variable (expressed as an integer ranging from 1 to 10) and extracted the estimated slope associated with year along with its 95% confidence limits. This process was repeated 5000 times for each level of percent annual change. The simulations excluded salinity, dissolved oxygen, temperature, and percent bottom vegetation cover; hence, the simulated populations were representative of those under average salinity, dissolved oxygen, temperature, and percent bottom vegetation cover in each estuary.

Finally, we assessed our ability to estimate the slope associated with the positive and negative trends in abundance. We examined two aspects of the 5000 estimated slopes associated with each level of percent annual change, which we termed coverage and significance. Each was expressed as a binary (0, 1) indicator for each of the 5000 estimated slopes. Coverage was assigned a 1 for the simulation replicates for which the true slope (the known, simulated annual percent increase or decrease) was contained within the 95% confidence interval (CI) of the estimated slope, and a 0 if it was not. Significance was assigned a 1 for the simulation replicates for which the upper 95% CI of the slope estimate was <0 (indicating a negative trend) or the lower 95% CI of the slope estimate was >0 (indicating a positive trend), which provided a measure of how often we detected a statistically significant temporal trend. Power for each replicate was calculated by multiplying the binary variables coverage and significance. Average power of the 5000 replicates indicated how well we were able to correctly detect a temporal trend (in terms of its direction and magnitude).

After assessing our ability to detect long-term abundance trends under Design 1 (two gear-specific surveys), we repeated the power simulations for Design 2, in which sampling was conducted with only the otter trawl and the number of tows was approximately doubled in four of the five estuaries (N = 612, as opposed to 372, trawls per year) (Design 2; Table 1); all other survey design elements remained unchanged. This allowed us to determine whether we could measurably influence our statistical power with a realistic amendment to the survey (i.e., a simplification of the survey to a single gear type with greater sample size).

The power analyses described above were used to assess our ability to detect steady, long-term trends (though varying in magnitude) in the population over a 10-year period; however, fish populations can vary on much shorter time scales. For example, for some species with periodic spawning or recruitment (e.g., *M. microlepis*; Switzer et al., 2012), recruitment can occur in pulses, or individual years of high recruitment. Likewise, negative perturbations to the system may result in pulses of low recruitment. Thus, we were also interested in whether we could detect changes in population abundance for a one-year time step. We did this by simulating three different percent annual changes in CPUE (indicative of abundance) over a one-year interval, resulting in year-2 populations 0.25, 0.5, and 0.67 times the initial population for negative trends and 4.00, 2.00, and 1.50 times the initial population for positive trends. The three levels were a subset of the levels used in the 10-year simulation and were chosen because they were reasonable changes to expect from one year to the next for some of our species of interest. We examined the estimated slope associated with year, which was equivalent to an offset of the intercept (i.e., a change in 1 year). As in the previous analyses, we calculated coverage and significance associated with the slope estimates, and power as the product of the two.

Due to very low CPUE in all estuaries, power simulations were not performed for *E. morio*, as a negative binomial model could not be fit to the catch data. Similarly, for *H. plumierii*, we had to remove two estuaries (SA and AP) because catches were very low there. Finally, because *L. maximus* was almost exclusively captured in the BB region, we restricted the power simulations to a single estuary (BB; and therefore, since haul seines were not deployed there, to a single gear type) but allowed for estimating overdispersion parameters and power for the three sampling zones (BBA, BBB, BBD) that formed the estuary sampling unit.

All data analyses were conducted in R v. 3.6.0 (R Core Team, 2019) using the packages "glmmTMB" (Brooks et al., 2017) for model fitting and "MuMIn" (Barton, 2020) for model selection. Reef fish catch data and example power analysis simulation codes are available online (http://dx.doi.org/10.17632/9bzshm5h46.1) (Schrandt et al., 2021). Example power analysis simulation codes are also provided as supplemental material (Supplemental Material 1, 2).

3. Results

3.1. Catch information

All reef species of interest were collected with both gear types. When we considered all estuaries and years of sampling, average taxon richness differed between gear types ($t = 3.56$; $P = 0.0004$) (Fig. 2A); however, biologically, the number of taxa was similar (14 taxa) because it is not possible to observe a partial taxon. The total number of taxa observed in each gear type from 2008 to 2018 was greater for trawls (205 taxa) than for haul seines (168 taxa) (Fig. 2B). There were 125 taxa caught in both gear types; the haul seine had 43 unique taxa, the trawl 80.

Catch-per-unit-effort for all species of interest was greater in trawls than in haul seines for all sampled estuaries (Fig. 3). The three species with the greatest CPUEs were *C. striata*, *Lutjanus griseus*, and *L. synagris,* which were collected in all five estuaries. *Haemulon plumierii* had an intermediate CPUE compared to that of the other species and was collected mainly in the BB region and in peninsular Florida (TB, CH); *H. plumierii* CPUE was much lower in the panhandle (SA and AP) estuaries. The species with the lowest CPUEs overall were *M. microlepis*, *Lachnolaimus maximus,* and *E. morio. Mycteroperca microlepis* were collected in each estuary (but in very small numbers), *L. maximus* were encountered almost exclusively in the BB region, and *E. morio* were rarely collected.

There was considerable overlap in size range between gear types for all reef species, with more individuals measured from the trawl (Table 2, Fig. 4) for all species except *L. griseus* and *M. microlepis*. Regardless of gear type, most fish were <200 mm SL. The KS tests performed for the size range present in both gear types indicated differences in distribution between the two gear types for each species of interest (Table 2). Generally, more small fish were collected in trawls than in haul seines (minimum and mean SL were smaller in fish caught in trawls than in those caught in haul seines), and the largest fish tended to be collected with haul seines.

3.2. Power analyses

For haul seine power simulations over a 10-year period, classic power curves were obtained: power was greatest for the most extreme trend in population abundance and decreased as the change in population size (i.e., percent annual change) decreased (Fig. 5). The power curves also tended to track relative abundance for each species and estuary, as power was, overall, greatest for species × estuary combinations that had higher abundance of the species of interest. A slight improvement in power was seen for a simulated positive population trend. If we used as a target 0.8 power to detect a halving or doubling of the population (50% change on the exponential scale, and our program's desired power level for population trends), power curves failed to meet the target, except for *L. griseus* in TB and CH. Average power (calculated by averaging the estimated power for each of the estuaries) to detect a 50% change in abundance over a 10-year period ranged from 0.28 to 0.52, depending on species. Average power to detect year-to-year population pulses, positive or negative, ranged from 0.15 to 0.33, depending on species, and was lower than that to detect trends over 10 years. Power reached 0.8 for a few scenarios (species × estuary combinations), typically those in which the population was either ¼ or 4× the size of the original population (Fig. 6).

Trawl power simulations for Design 1 resulted in overall greater power to detect population abundance trends than did haul seine simulations. Power curves for trawls were somewhat higher (>0.6) for less extreme trends than those for haul seines. For example, power decreased below 0.8 for haul seines when the population ratio was ~3–3.5, but for trawls the decline generally did not occur until a ratio of ~2.5 or less was reached (Fig. 7). Average power (pooled across bays and direction of the trend) to detect a 50% change in abundance ranged from 0.41 to 0.75, depending on the species.

With multiple lines of evidence favoring the trawl as a more efficient and statistically powerful gear type, we also conducted power simulations for Design 2 (haul seine deployments were discontinued and additional trawl tows were done). For both positive and negative simulated abundance trends over a 10-year period, the power curve improved under Design 2 (Fig. 7). Power neared or was at 0.8 for a halving or doubling of the population for four of the five species under Design 1 but for all five species under Design 2. For *L. maximus* in BB, power was just below 0.8 for a halving of the population and at 0.8 for a doubling of the population in one of the three zones (BBB), under Design 1 (Fig. 8). Big Bend zones BBA and BBD had very similar power curves and both had consistently lower power curves than BBB (Fig. 8). For each species, power curves were better within each estuary (and zone for *L. maximus*) under Design 2 than in Design 1; more estuaries reached power = 0.8 under Design 2 than under Design 1 (Fig. 7, 8). Average power to detect a halving or doubling of the population (pooled across estuaries) under Design 2 ranged from 0.64 to 0.91, depending on species, half of which had an average power >0.8. For either design, power was generally greater for positive population trends than for negative trends.

The power to detect pulse changes in population abundance was greater for trawls than for haul seines and reached 0.8 for a halving or doubling for more than one species (Fig. 9). Power was greatest for a fourfold change in the population size, although power was 0.8 to detect a halving or doubling from one year to the next for a few species (i.e., *C. striata*, *H. plumierii*, *Lutjanus griseus*). For *Lachnolaimus maximus* in BB, power did not reach 0.8 for either a halving or doubling of the population under either design (Fig. 10). Zones BBA and BBD within BB had very similar power curves, and power was lower than in zone BBB. Power to detect year-to-year population differences was greatest for a fourfold change in the population and did reach 0.8 for at least one zone in the BB under each sampling design. Average power (pooled across sampling zones) to detect a halving or doubling in abundance ranged from 0.22 to 0.53 under Design 1 and from 0.40 to 0.77 under Design 2, depending on the species. On the other hand, average power to detect a fourfold change in abundance over a two-year period ranged from 0.63 to 0.89 under Design 1 and from 0.81 to 0.94 under Design 2, depending on the species.

4. Discussion

Various fishery-independent surveys use multiple gear types to collect information across size classes and life history stages (e.g., McMichael, 1991; Ault et al., 2018; Powers et al., 2018). Some fishery-independent surveys, however, are designed to provide information for a single life stage (e.g., juvenile fish in a specific habitat), and a single gear type may be sufficient. This is the case in the present study: the otter trawl provided similar information about our reef species of interest, was more efficient (i.e., resulted in a higher CPUE), and was more powerful in detecting trends in abundance than the haul seine. Therefore, in 2019, we discontinued the haul seine survey and continued the WFS juvenile reef fish survey as a single gear type survey (otter trawl). To improve our statistical power, we increased the sample size to match Design 2. All other design elements remain unchanged, maintaining continuity of the trawl time series.

We are aware that moving from a multiple-gear to a single-gear survey encompasses many differences, namely effects of sampling a different microhabitat, different size-selectivity of the gear types, and different direct effects on the habitat from the gear types. Although both gear types are useful for sampling fish in polyhaline seagrass beds, and the reef species of interest were present in samples from both gear types, the methodology for the haul seine directly targets shallower, sloping seagrass beds (shoals) while the trawl targets the deeper seagrass beds. The depth and slope of the bottom may affect the catch. Species that rely more on shallow, sloping seagrass beds (i.e., *M. microlepis*; Switzer et al., 2012) are likely to have higher catches with the haul seine than with the trawl. Indeed, *M. microlepis* was one of the few species for which more individuals were captured per haul seine than trawl; but when effort was factored in, CPUE for *M. microlepis* was greater with the trawl. The range of fish lengths overlapped between gear types, but the distribution differed for various species. This may be a result of the combination of sampled microhabitats and r size selectivity of the gear types. The haul seine generally catches individuals >100 mm SL, but the otter trawl, with the small mesh size in the liner, catches both small- and large-bodied individuals. Therefore, otter trawls in our survey collected more smaller individuals than the haul seine. Discontinuing use of the haul seine suggests we will have lost some length-frequency information for larger individuals but overall, this loss is a small portion of the catch data. We do not feel this is detrimental to the users of these data, as most individuals collected in the study, regardless of gear type, were <200 mm SL. For the reef species of interest, juveniles using seagrass beds are generally <200 mm SL (young-of-the-year size <200 mm SL; Hood et al., 1994; Johnson et al., 1995; Burton, 2001; Murie and Parkyn, 2005; McBride and Richardson, 2007; Collins and McBride, 2011), except *M. microlepis*, whose young-of-the-year SL is ca. 300 mm (Hood & Schlieder,1992). Therefore, information from the otter trawl is likely sufficient to provide meaningful length-frequency information for juvenile reef fish using polyhaline seagrass beds as nursery habitat.

The use of otter trawls for scientific surveys has been criticized as disruptive and sometimes destructive of benthic habitats (Trenkel et al., 2019). This is a valid concern, but the otter trawls that we use directly affect only a small percentage of the available seagrass habitat. Also, otter trawls can ride on top of submerged aquatic vegetation, despite the weight of the tickler chain (Leber and Greening, 1986), reducing negative impacts to the seagrass habitat. Furthermore, otter trawls have been used over several decades for research in seagrass beds, and consistent methodologies (over time and space) can aid in understanding temporal trends (Stallings et al., 2014).

Our evaluation of the WFS juvenile reef fish survey suggests that the otter trawl provides not only a better quantity of data (i.e., CPUE, diversity, length-frequency) for use in stock assessments, but also a better quality of data than the haul seine. The overarching goal of the WFS juvenile reef fish survey is to provide statistically strong data for fishery management (e.g., stock assessments, ecosystem models), and the power simulations performed in the present study indicate that the trawl-only survey has the statistical power to detect population changes of various degrees. For trend analyses, statistical power is the probability of detecting trends in population abundance, if a trend exists (Seavy and Reynolds, 2007). Goals for statistical trends have been proposed for surveys of various species. For example, Bart et al. (2004) recommended 80% power to detect a 50% decline over 20 years for land birds, while Hatch (2003) recommended 90% power to detect a 50% decline over 10 years for colonial seabirds. Similarly, Gibbs et al. (1998) used 80% power to detect a 50% change over 10 years as a guideline in their analysis of multiple taxa, including fish species. By extension of such similar statistical goals, the FIM program was originally designed with the goal of being able to achieve 80% power to detect a halving or doubling of the population. The WFS juvenile reef fish survey meets this goal when considering long-term, 10-year trends, for one to three estuaries in four of five reef species under Design 1 with trawls and for one to five estuaries in four of five species for Design 2 with trawls. When the haul seine design is simulated, the goal is met only for one or two species in one or two estuaries. Power is markedly lower for simulating significant population changes using either gear type over a two-year period, but the goal is met for trawl sampling under a few scenarios. For stock assessments, data from all estuaries in the eastern Gulf of Mexico are combined (e.g., *M. microlepis*; SEDAR 2016; *L. griseus*; SEDAR 2018), so meeting power goals for multiple individual estuaries for a given species is promising for improving juvenile recruitment information in stock assessment models. Furthermore, our estuary-specific estimation of power allows for the potential to weight the data from different estuaries according to their strength of statistical power, which could also improve indices of abundance for stock assessments. Average power (for all estuaries combined) was lower than 0.8 because estuaries were included that had relatively lower power. Under Design 2, power was >0.6 for all reef species examined and half the species achieved power >0.8.

The four main contributors to statistical power in hypothesis testing are sample size, variation, effect size, and significance level (Zar, 1999). Sample size and variation likely act in concert to affect our power simulations. Generally, as sample size increases, the variation in the response variable decreases and a better estimate is achieved, resulting in higher statistical power (Zar, 1999). This idea was evident in our power simulations because: 1) overall, trawls achieved greater power than haul seines, and more trawl samples were taken than haul-seine samples; 2) power was greater during the 10-year simulation than during the 2-year simulation, and many more samples were taken over 10 years than over 2 years; and 3) for trawls, power increased from Design 1 to Design 2, where the only difference in the sampling design was nearly doubling the sample size in Design 2 (the sampling methodology, negative binomial model, starting abundances, and variation remained unchanged). Therefore, the discerned trends had greater power under scenarios with larger sample sizes.

Our results also highlight the importance of long-term monitoring for detecting trends in population abundance, as statistical power was much greater for a 10-year sampling period than for a 2-year sampling period, even when the magnitude of change in the population was similar between the two periods. We found that at least 10 years of data was needed for detecting management-relevant trends in juvenile reef fish abundance. This is similar to White's (2019) findings of between 5–30 years of data for small population changes, with most fish populations requiring a data set at least 15 years in duration. Compared with Gibbs et al. (1998), we needed more samples per year and more years to detect trends, but this is likely due to our smaller catches and the negative binomial distribution of our data, vs. larger catches and a linear regression model for Gibbs et al. (1998) work.

The effect size and significance levels were determined *a priori* in our power simulations, and they too resulted in expected trends. Larger changes in population sizes over time resulted in greater power to detect those changes. This is understandable, as large changes or effects are easier to detect than smaller ones (Zar, 1999). Regarding the significance level, we chose the 95% confidence interval to evaluate our ability to detect the appropriate slope in our simulations; therefore, all the power curves that exhibited a plateau exhibited it at ~0.95 power. If we chose, say, 90%, we would expect more of our estimated slopes to fall within the 90% confidence level of the true slope, and power would increase because the proportion of simulations with estimated slopes within the 90% confidence interval would increase.

We also acknowledge the importance of the value of the response variable in power analyses, in this case the number of individuals collected. Enough individuals (non-zero samples) are needed to fit the type of model (i.e., negative binomial) that we used to model fish population trends over time. For example, we did not collect enough *E. morio* during the 10-year period to properly fit the negative binomial model and then to run power simulations. Also, statistical power in modeling the reef fish species that had higher catches overall (e.g., *C. striata*, *L. griseus*, *L. synagris*) was greater, regardless of gear type. This is likely a result of having more individuals from which to estimate the starting size of our population for the simulations; a larger starting population allows for a reasonable number of individuals to be sampled if the population does decrease two- or threefold.

The successfulness of a monitoring survey should be defined not only by the effectiveness and efficiency of the methodology, but also by the information that it provides to users. These data include not only fish abundance, but also life history information for a variety of species. Juvenile length-frequency information from our survey can be used to inform stock assessment models on abundance, growth, and natural mortality parameters (Beverton and Holt, 2012) and, in conjunction with weight information, can be used to estimate biomass for various species in ecosystem models (e.g., Ecopath with Ecosim; Christensen and Walters, 2004). In addition to providing data for stock assessments, multispecies monitoring data are useful in tracking species diversity and ecosystem health (e.g., Whitfield and Elliott, 2002; Mieszkowska et al., 2014; Miloslavich et al., 2018) and detecting spatial distributions, range expansions, and the appearance of nonnative species (e.g., Grüss et al., 2018).

Effective fisheries management relies on the ability to detect abundance trends, and decision makers should be able to have confidence in an apparent trend when it is presented (Wauchope et al. 2019). We have demonstrated that, in our multispecies survey, we can detect statistically sound trends in long-term data sets (e.g., 10 years) and that, by a modification of the survey design, can improve their statistical power. The results of the present study have prompted a realistic change in the FIM WFS juvenile reef fish survey. The haul seine effort was discontinued in 2019 and more trawls are now deployed. Before making this change, we had carefully examined the data obtained for each gear type, the statistical power of each, and the logistical restraints. For example, to obtain the desired power with haul seines, we would have needed to significantly increase sampling effort, which is not feasible, given the small amount of available habitat that meets the sampling requirements. We also considered the time and personnel required for the sampling events, and it did not increase significantly. So, the realistic amendment to the survey design does not require additional funding but does improve the effectiveness of the survey in providing data of a high-enough statistical quality to be useful to fishery managers. It also improves the effectiveness of the multispecies survey. The additional trawl tows under Design 2 resulted in greater power to detect trends for more species, which clearly allows the newer design to provide more information on multiple managed species in the eastern Gulf of Mexico. We are aware that Design 2 does present some limitations for accurately predicting stock productivity for species that regularly exhibit periodic pulses in recruitment (e.g., *M. microlepis*) since the year-to-year power values are lower than the long-term ones. Under Design 2, however, we can reliably detect fourfold and, to a lesser-extent, twofold changes in *M. microlepis* populations from year to year, which are well-within previously reported pulse changes in this species (Switzer et al., 2012).

We have demonstrated that we can provide better-quality data for stock assessments (and other uses) by making realistic changes to survey that had been in use for 11 years. Prior to this study, the design was refined temporally to focus on the months during which the species of interest recruit to polyhaline seagrass habitats. It was also more focused spatially by discontinuing sample in locations that did not meet the desired seagrass habitat characteristics. These previous changes decreased variability in our estimates by removing months and areas in which most samples resulted in zero catch. We recognize that SA Bay has never been a good recruitment area (and so could be dropped from sampling) and increases the number of zeroes in our catch data. Retaining this bay at the western extent of the survey area is important if we wish to continue to monitor for changes in species distributions, such as those related to climate change. The results of this study indicate that retaining SA and using a single gear type allows us a simpler design that is more efficient, effective, and powerful for detecting trends in population abundance, with no increase in cost. The differences in power to detect trends among eastern Gulf of Mexico estuaries suggests it may be possible to weight the estuary-specific data accordingly for stock assessments so that more weight is given to estuaries with greater statistical power to detect trends. This could further improve our confidence in juvenile recruitment estimates, which can lead to more confidence in predicting fishery productivity and informing fisheries management decisions. Ultimately, we have amended the WFS juvenile reef fish survey with plans to continue sampling under the new design into the future, to provide better, stronger, data for multiple species, which provides a step in the direction toward multispecies management.

CRediT author statement

Meagan N. Schrandt: Conceptualization, Methodology, Software, Validation, Formal Analysis, Investigation, Writing – Original Draft, Writing – Review & Editing, Visualization **Colin Shea**: Conceptualization, Methodology, Software, Validation, Formal Analysis, Writing – Original Draft, Writing – Review & Editing, Visualization **Benjamin Kurth:** Software, Investigation, Writing – Review & Editing **Theodore Switzer:** Conceptualization, Methodology, Resources, Writing – Review & Editing, Supervision, Project administration, Funding acquisition

Acknowledgements

We thank the staff of the Florida Fish and Wildlife Conservation Commission’s Fish and Wildlife Research Institute’s Fisheries Independent Monitoring group for their dedication to field sampling and data collection. The data collection and analyses for this project were supported in part by funding from the State of Florida saltwater recreational fishing licenses, from the U.S. Department of Commerce, National Oceanic and Atmospheric Administration, National Marine

Fisheries Service [grant numbers NA08NMF4720645 and NA09NMF4330152], by funding from the U.S. Department of the Interior, U.S. Fish and Wildlife Service, Federal Aid for Sportfish Restoration [grant numbers F14AF00328, F15AF01222, F16AF00898, F17AF00932, and F18AF00665], and by funding from the National Fish and Wildlife Foundation Gulf Environmental Benefit Fund [grants numbers FL 40624, FL 45766, FL 50347, FL 54269, and FL 58101]. The statements, findings, views, conclusions, and recommendations contained in this document are those of the authors and do not necessarily reflect the views of the U.S. Department of the Interior or the U.S. Department of Commerce and should not be interpreted as representing the opinions or policies of the U.S. government. Mention of trade names or commercial products does not constitute their endorsement by the U.S. government.

References

Ault, J.S., Smith, S.G., Richards, B.L., Yau, A.J., Langseth, B.J., O'Malley, J.M., Boggs, C.H., Seki, M.P., DiNardo, G.T., 2018. Towards fishery-independent biomass estimation for Hawaiian Islands deepwater snappers. Fish. Res. 208, 321–328.

Bart, J., Burnham, K.P., Dunn, E.H., Francis, C.M., Ralph, C.J., 2004. Goals and strategies for estimating trends in landbird abundance. J. Wildlife Manage. 68(3), 611–626.

Barton, K., 2020. MuMIn: Multi-model Inference. R package version 1.43.17. https://CRAN.R-project.org/package=MuMIn.

Beverton, R.J., Holt, S.J., 2012. On the dynamics of exploited fish populations (vol. 11). Springer Science & Business Media.

Bolker, B.M., 2008. Ecological Models and Data in R. Princeton University Press, Princeton, New Jersey.

Brooks, M.E., Kristensen, K., van Benthem, K.J., Magnusson, A., Berg, C.W., Nielsen, A., Skaug, H.J., Maechler, M., Bolker, B.M., 2017. glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. R Journal. 9, 378–400. https://doi.org/10.3929/ethz-b-000240890.

Burton, M.L., 2001. Age, growth, and mortality of gray snapper, *Lutjanus griseus*, from the east coast of Florida. Fish. B-NOAA 99, 254–265.

Casey, J.P., Poulakis, G.R., Stevens, P.W., 2007. Habitat use by juvenile gag, *Mycteroperca microlepis* (Pisces: Serranidae), in subtropical Charlotte Harbor, Florida (USA). Gulf Carib. Res. 19(1), 1–9.

Christensen, V., Walters, C.J., 2004. Ecopath with Ecosim: methods, capabilities and limitations. Ecol. Model, 172(2–4), 109–139.

Coleman, F.C., Koenig, C.C., Eklund, A.M., Grimes, C.B., 1999. Management and conservation of temperate reef fishes in the grouper–snapper complex of the southeastern United States, in: Musick, J.A. (Ed.), Life in the Slow Lane: Ecology and Conservation of Long-lived Marine Animals. Symposium 23. American Fisheries Society, Bethesda, Maryland, pp. 233–242.

Collins, A.B., McBride, R.S., 2011. Demographics by depth: spatially explicit life-history dynamics of a protogynous reef fish. Fish. B-NOAA 109(2), 232–242.

De Angelo, J.A., Stevens, P.W., Blewett, D.A., Switzer, T.S., 2014. Fish assemblages of shoal- and shoreline-associated seagrass beds in eastern Gulf of Mexico estuaries. T. Am. Fish. Soc. 143(4), 1037–1048.

Flaherty-Walia, K.E., Switzer, T.S., Winner, B.L., J. Tyler-Jedlund, A., Keenan, S.F., 2015. Improved ability to characterize recruitment of Gray Snapper in three Florida estuaries along the Gulf of Mexico through targeted sampling of polyhaline seagrass beds. T. Am. Fish. Soc. 144(5), 911–926. https://doi.org/10.1080/00028487.2015.1054516.

Geselbracht, L.L., Freeman, K., Birch, A.P., Brenner, J., Gordon, D.R., 2015. Modeled sea level rise impacts on coastal ecosystems at six major estuaries on Florida's gulf coast: Implications for adaptation planning. PLOS ONE 10(7), e0132079.

Gibbs, J.P., Droege, S. and Eagle, P., 1998. Monitoring populations of plants and animals. BioScience. 48(11), 935–940.

Gillanders, B.M., Able, K.W., Brown, J.A., Eggleston, D.B., Sheridan, P.F., 2003. Evidence of connectivity between juvenile and adult habitats for mobile marine fauna: an important component of nurseries. Mar. Ecol. Prog. Ser. 247, 281–295.

Grüss, A., Perryman, H.A., Babcock, E.A., Sagarese, S.R., Thorson, J.T., Ainsworth, C.H., Anderson, E.J., Brennan, K., Campbell, M.D., Christman, M.C., Cross, S., 2018. Monitoring programs of the U.S. Gulf of Mexico: inventory, development and use of a large monitoring database to map fish and invertebrate spatial distributions. Rev. Fish. Biol. Fisher. 28(4), 667–691.

Hansen, J.A., Carpenter, S.R, Gaeta, J.W, Hennessy, J.M., Vander Zander, M.J., 2015. Predicting walleye recruitment as a tool for prioritizing management actions. Can. J. Fish. Aquat. Sci. 72, 661–672. dx.doi.org/10.1139/cjfas-2014-0513

Hartig, F., 2020. DHARMa: Residual Diagnostics for Hierarchical (Multi-level/Mixed) Regression Models. R package version 0.3.2.0. https://CRAN.R-project.org/package=DHARMa

Hatch, S.A., 2003. Statistical power for detecting trends with applications to seabird monitoring. Biol. Conserv. 111(3), 317–329.

Hood, P.B., Godcharles, M.F., Barco, R.S., 1994. Age, growth, reproduction, and the feeding ecology of black sea bass, *Centropristis striata* (Pisces: Serranidae), in the eastern Gulf of Mexico. B. Mar. Sci. 54(1), 24–37.

Hood, P.B., Schlieder, R.A., 1992. Age, growth, and reproduction of gag, *Mycteroperca microlepis* (Pisces: Serranidae), in the eastern Gulf of Mexico. B. Mar. Sci. 51(3), 337–352.

Johnson, A.G., Collins, L.A., Dahl, J., Baker, M.S. Jr., 1995. Age, growth, and mortality of lane snapper from the northern Gulf of Mexico, in: Eversole, A.G.(Ed./s.), Proceedings of the Forty-ninth Annual Conference of the Southeastern Association of Fish and Wildlife Agencies, pp. 178–186.

Johnson, A.G., Koenig, C.C., 2005. Age and size structure of the fishery and juvenile abundance of gag (*Mycteroperca microlepis*), from the northeastern Gulf of Mexico. Proceedings of the 47th Gulf and Caribbean Fisheries Institute. pp. 906–913.

Koenig, C.C., Coleman, F.C., 1998. Absolute abundance and survival of juvenile gags in sea grass beds of the northeastern Gulf of Mexico. T. Am. Fish. Soc. 127(1), 44–55.

Leber, K.M., Greening, H.S., 1986. Community studies in seagrass meadows: a comparison of two methods for sampling macroinvertebrates and fishes. Fish. Bull. 84, 443–457.

Lefcheck, J.S., Hughes, B.B., Johnson, A.J., Pfirrmann, B.W., Rasher, D.B., Smyth, A.R., Williams, B.L., Beck, M. W., Orth, R.J., 2019. Are coastal habitats important nurseries? A meta-analysis. Conserv. Lett. 12(4), e12645.

McBride, R.S., Richardson, A.K., 2007. Evidence of size-selective fishing mortality from an age and growth study of hogfish (Labridae: *Lachnolaimus maximus*), a hermaphroditic reef fish. B. Mar. Sci. 80(2), 401–417.

McMichael, R.H., Jr., 1991. Florida's marine fisheries–independent monitoring program, in: Treat, S.F., Clark, P.A. (Eds.), Proceedings, Tampa Bay Area Scientific Information Symposium 2 (BASIS). Tampa Bay Regional Planning, Tampa, Florida, pp. 255–261.

Mieszkowska, N., Sugden, H., Firth, L.B., Hawkins, S.J., 2014. The role of sustained observations in tracking impacts of environmental change on marine biodiversity and ecosystems. Philos. T. R. Soc. A, 372(2025). http://dx.doi.org/10.1098/rsta.2013.033

Miloslavich, P., Bax, N.J., Simmons, S.E., Klein, E., Appeltans, W., Aburto-Oropeza, O., Andersen Garcia, M., Batten, S.D., Benedetti-Cecchi, L., Checkley, D.M., Jr., Chiba, S., 2018. Essential ocean variables for global sustained observations of biodiversity and ecosystem changes. Glob. Change Biol. 24(6), 2416–2433. DOI: 10.1111/gcb.14108

Murie, D.J., Parkyn, D.C., 2005. Age and growth of white grunt (*Haemulon plumieri*): a comparison of two populations along the west coast of Florida. B. Mar. Sci. 76(1), 73–93.

Nagelkerken, I., Kleijnen, S., Klop, T., van den Brand, R.A.C.J., de la Morinière, E.C., van der Velde, G., 2001. Dependence of Caribbean reef fishes on mangroves and seagrass beds as nursery habitats: a comparison of fish faunas between bays with and without mangroves/seagrass beds. Mar. Ecol. Prog. Ser. 214, 225–235.

Nagelkerken, I., Roberts, C.V., van der Velde, G., Dorenbosch, M., van Riel, M.C., de la Morinière, E.C., Nienhuis, P.H., 2002. How important are mangroves and seagrass beds for coral-reef fish? The nursery hypothesis tested on an island scale. Mar. Ecol. Prog. Ser. 244, 299–305.

Newman, D., Berkson, J., Suatoni, L., 2015. Current methods for setting catch limits for data-limited fish stocks in the United States. Fish. Res. 164, 86–93.

Powers, S.P., Drymon, J.M., Hightower, C.L., Spearman, T., Bosarge, G.S., Jefferson, A., 2018. Distribution and age composition of red snapper across the inner continental shelf of the north-central Gulf of Mexico. T. Am. Fish. Soc. 147(5), 791–805.

R Core Team, 2019. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/.

[dataset] Schrandt, M.N., Shea, C.P., Kurth, B.N., Switzer, T.S., 2021. Fishery-independent catch data of selected juvenile reef fish in the eastern Gulf of Mexico. Mendeley Data, v1. doi:10.17632/9bzshm5h46.1

Schrandt, M.N., Switzer, T.S., Stafford, C.J., Flaherty-Walia, K.E., Paperno, R., Matheson, R.E., Jr., 2018. Similar habitats, different communities: fish and large invertebrate assemblages in eastern Gulf of Mexico polyhaline seagrasses relate more to estuary morphology than latitude. Estuar. Coast. Shelf. S. 213, 217–229.

Seavy, N.E., Reynolds, M.H., 2007. Is statistical power to detect trends a good assessment of population monitoring? Biol. Conserv. 140(1–2), 187–191. https://doi.org/10.1016/j.biocon.2007.08.007

SEDAR, 2016. SEDAR 33 Update – Gulf of Mexico Gag grouper Stock Assessment Report. SEDAR, North Charleston, South Carolina. Available online at: http://sedarweb.org/sedar-projects

SEDAR, 2018. SEDAR 51 – Gulf of Mexico Gray Snapper Stock Assessment Report. SEDAR, North Charleston, South Carolina. Available online at: http://sedarweb.org/sedar-projects

Smith, P.E., 1993. Balancing sampling precision and fisheries management objectives: minimal methods. B. Mar. Sci. 53, 930–935.

Stallings, C.D., Brower, J.P., Heinlein Loch, J.M., Mickle, A., 2014. Catch comparison between otter and rollerframe trawls: implications for sampling in seagrass beds. Fish. Res. 155, 177–184.

Stige, L.C., Hunsicker, M.E., Bailey, K.M., Yaragina, N.A., Hunt Jr, G.L., 2013. Predicting fish recruitment from juvenile abundance and environmental indices. Mar. Ecol. Prog. Ser. 480, 245–261.

Switzer, T.S., MacDonald, T.C., McMichael, R.H., Jr., Keenan, S.F., 2012. Recruitment of juvenile Gags in the eastern Gulf of Mexico and factors contributing to observed spatial and temporal patterns of estuarine occupancy. T. Am. Fish. Soc. 141(3), 707–719.

Trenkel, V.M., Vaz, S., Albouy, C., Brind'Amour, A., Duhamel, E., Laffagrue, P., Romagnan, J.B., Simon, J., Lorance, P., 2019. We can reduce the impact of scientific trawling on marine ecosystems. Mar. Ecol. Prog. Ser. 609, 277–282.

Wauchope, H.S., Amano, T., Sutherland, W.J., Johnston, A., 2019. When can we trust population trends? A method for quantifying the effects of sampling interval and duration. Methods Ecol. Evol. 10(12), 2067–2078.

Wertheimer, A.C., Orsi, J.A., Fergusson, E.C., Gray, A., 2016. Forecasting Pink Salmon Harvest in Southeast Alaska from Juvenile Salmon Abundance and Associated Biophysical Parameters: 2015 Returns and 2016 Forecast. North Pacific Anadromous Fish Commission Doc. NOAA, NMFS, Auke Bay Laboratory, Alaska Fisheries Science Center, http://www.npafc.org

White, E.R., 2019. Minimum time required to detect population trends: the need for long-term monitoring programs. BioScience. 69(1), 40–46.

White and Bennets, 1996;

Whitfield, A.K., Elliott, M., 2002. Fishes as indicators of environmental and ecological changes within estuaries: a review of progress and some suggestions for the future. J. Fish Biol., 61, 229–250. https://doi.org/10.1111/j.1095-8649.2002.tb01773.x

Zar, J.H., 1999. Biostatistical Analysis, fourth ed, Prentice Hall, Upper Saddle River, New Jersey.

Zuur, A.F., Ieno, E.N., Walker, N.J., Saveliev, A.A., Smith, G.M., 2009. Mixed Effects Models and Extensions in Ecology with R. Springer, New York.

Tables

Table 1. Sample summary by estuary for haul seines and trawls for the fishery-independent-monitoring survey of juvenile reef fish recruitment on the West Florida Shelf, June–November 2008–2018 (Design 1) and June–November 2019 (Design 2). The Big Bend (BB) estuary comprises three independent sampling zones; sample sizes for each zone are in gray text to highlight the change in sampling distribution (i.e., the removal of zone BBA in Design 2) between the two designs for BB.

Estuary	Design 1 (2008–2018)				Design 2 (2019)			
	Haul seine		Trawl		Haul seine		Trawl	
	Monthly	Yearly	Monthly	Yearly	Monthly	Yearly	Monthly	Yearly
Saint Andrew (SA)	3	18	6	36	0	0	12	72
Apalachicola Bay (AP)	4	24	8	48	0	0	16	96
Big Bend region (BB)	0	0	30	180	0	0	30	180
Zone BBA	0	0	10	60	0	0	0	0
Zone BBB	0	0	10	60	0	0	15	90
Zone BBD	0	0	10	60	0	0	15	90
Tampa Bay (TB)	7	42	10	60	0	0	24	144
Charlotte Harbor (CH)	6	36	8	48	0	0	20	120
Total	*20*	*120*	*62*	*372*	*0*	*0*	*102*	*612*

Table 2. Summary of length data for seven juvenile reef fish species collected during polyhaline seagrass sampling, 2008–2018, along the West Florida Shelf. The total number of fish measured (Number of Lengths) and the mean standard lengths (SL) are presented for each gear type, followed by the common size range that was used for the two-sided KS test.

Species	Number of Lengths		Mean SL (sd) in mm		Common size range (mm SL)	KS test	
	Haul seine	Trawl	Haul seine	Trawl		D	P-value
Centropristis striata	1,533	22,643	120.29 (26.8)	79.18 (32.6)	33–270	0.54928	2.20E−16
Epinephelus morio	32	124	118.97 (20.0)	111.63 (19.8)	90–172	0.61129	1.79E−10
Haemulon plumierii	1,197	5,098	91.22 (21.8)	51.82 (23.3)	38–196	0.41365	2.20E−16
Lachnolaimus maximus	32	761	107.00 (27.0)	79.40 (29.0)	67–175	0.45402	3.80E−06
Lutjanus griseus	4,772	3,633	150.76 (38.8)	112.22 (51.4)	38–300	0.67598	2.20E−16
Lutjanus synagris	1,547	7,724	93.28 (15.6)	50.37 (24.9)	22–167	0.44794	2.20E−16
Mycteroperca microlepis	1,774	1,042	182.19 (57.5)	165.62 (47.4)	52–365	0.79739	2.20E−16

Figure legends

Fig. 1. Five estuaries on the West Florida Shelf in which juvenile reef fish were sampled for abundance estimation, 2008–2018. They are, from west to east, St. Andrew Bay (SA), Apalachicola Bay (AP), Big Bend region (BB), Tampa Bay (TB), and Charlotte Harbor (CH).

Fig. 2. Comparison of taxon richness values derived using two gear types. Mean (± SE) taxon richness per set for haul seines and trawls (A) and total number of taxa caught using each gear type, 2008–2018 (B). The dashed line in (B) references the number of taxa caught using both gear types; the portion of the gray bar above the dashed line for each gear type represents the number of taxa caught in only that gear type (i.e., unique to that gear type). taxa above the dashed line were unique to each gear type.

Fig. 3. Comparison of mean catch-per-unit-effort (CPUE), by species of interest, of juvenile reef fish between gear types and estuaries. Juveniles were collected during the West Florida Shelf juvenile reef fish survey, 2008–2018. Mean presented as fish/100 m^2 with error bars depicting SE. Estuaries are St. Andrew Bay (SA), Apalachicola Bay (AP), Big Bend region (BB), Tampa Bay (TB), and Charlotte Harbor (CH).

Fig. 4. Length-frequency distributions, by species of interest, of juvenile reef fishes caught in haul seines and trawls, 2008–2018. Juveniles were collected during the West Florida Shelf juvenile reef fish survey. Distributions are shown in gray bars for haul-seine catches and black bars for trawl catches. Size classes are 10-mm standard length (SL) bins.

Fig. 5. Power curves for haul seine sampling simulations for a 10-year population trend for five reef species whose samples in multiple estuaries included enough individuals to allow estimation of statistical power in different estuaries. The top panel displays a simulated negative trend in the population size over ten years and the bottom panel displays a simulated positive trend in the population size over ten years. Dashed vertical gray lines are reference lines for a halving (negative trend) or a doubling (positive trend) of the population. Horizonal dashed gray reference lines denote power = 0.8, which is the desired power for the WFS juvenile reef fish survey.

Fig. 6. Power curves for haul seine sampling simulations for a year-to-year (2-year) population trend for five reef whose samples in multiple estuaries included enough individuals to allow estimation of statistical power in different estuaries. The top panel displays a simulated negative trend in the population size over two years and the bottom panel displays a simulated positive trend in the population size over two years. Dashed vertical gray lines are reference lines for a halving (negative trend) or a doubling (positive trend) of the population. Horizonal dashed gray reference lines denote power = 0.8, which is the desired power for the WFS juvenile reef fish survey.

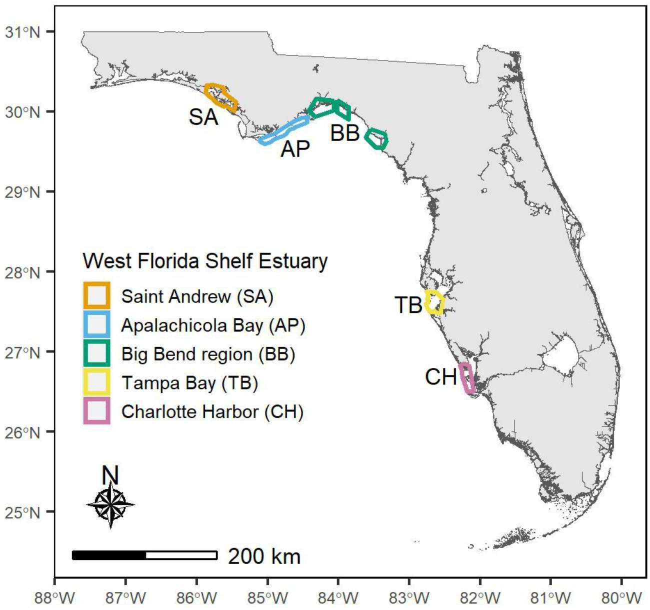
Fig. 7. Power curves for trawl sampling simulations for a 10-year population trend for five reef species whose samples in multiple estuaries included enough individuals to allow estimation of statistical power in different estuaries. The top panel displays a simulated negative trend in the population size over ten years and the bottom panel displays a simulated positive trend in the population size over ten years. Dashed vertical gray lines are reference lines for a halving

(negative trend) or a doubling (positive trend) of the population. Horizonal dashed gray reference lines denote power = 0.8, which is the desired power for the WFS juvenile reef fish survey.

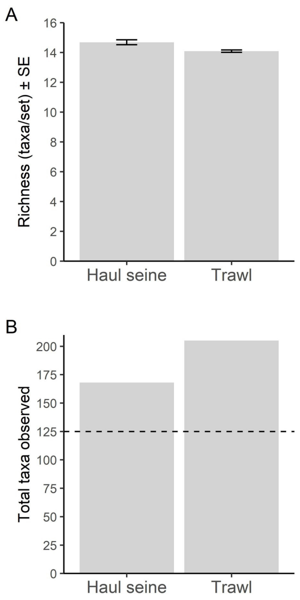
Fig. 8. Power curves for trawl sampling simulations for a 10-year population trend for *Lachnolaimus maximus* in the three sampling zones of the Big Bend estuary. The top two rows display a simulated negative trend in the population size over ten years, under two sampling designs (no change in sample size but a different distribution; Design 1 samples three zones and Design 2 samples two zones). The bottom two rows display a simulated positive trend in the population size over ten years, under the same two sampling designs. Dashed vertical gray lines are reference lines for a halving (negative trend) or a doubling (positive trend) of the population. Horizonal dashed gray reference lines denote power = 0.8, which is the desired power for the WFS juvenile reef fish survey.

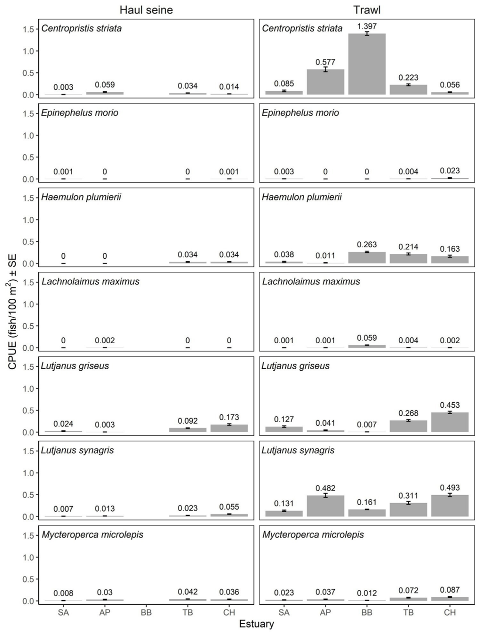
Fig. 9. Power curves for trawl sampling simulations for a year-to-year (2-year) population trend for five reef species whose samples in multiple estuaries included enough individuals to allow estimation of statistical power in different estuaries. The top panel displays a simulated negative trend in the population size over two years and the bottom panel displays a simulated positive trend in the population size over two years. Dashed vertical gray lines are reference lines for a halving (negative trend) or a doubling (positive trend) of the population. Horizonal dashed gray reference lines denote power = 0.8, which is the desired power for the WFS juvenile reef fish survey.

Fig. 10. Power curves for trawl sampling simulations for a year-to-year (2-year) population trend for *Lachnolaimus maximus* in the three sampling zones of the Big Bend estuary. The top two rows display a simulated negative trend in the population size over two years, under two sampling designs (no change in sample size but a different distribution; Design 1 samples three zones and Design 2 samples two zones). The bottom two rows display a simulated positive trend in the population size over two years, under the same two sampling designs. Dashed vertical gray lines are reference lines for a halving (negative trend) or a doubling (positive trend) of the population. Horizonal dashed gray reference lines denote power = 0.8, which is the desired power for the WFS juvenile reef fish survey.

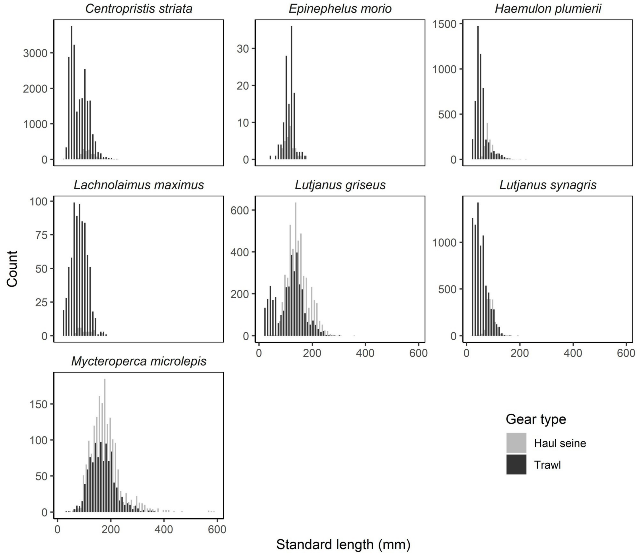


31°N
30°N
29°N
28°N
27°N
26°N
25°N
88°W
87°W
86°W
85°W
84°W
83°W
82°W
81°W
80°W
SA
AP
BB
TB
CH
West Florida Shelf Estuary
Saint Andrew (SA)
Apalachicola Bay (AP)
Big Bend region (BB)
Tampa Bay (TB)
Charlotte Harbor (CH)
N
200 km

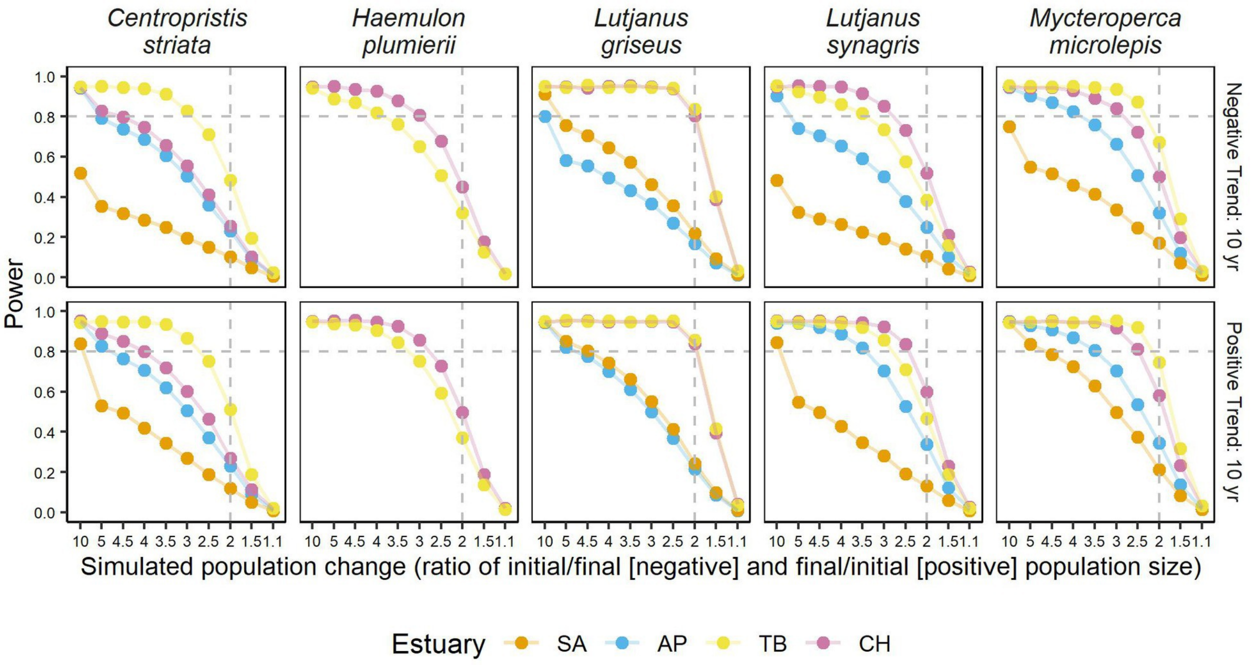

A
Richness (taxa/set) ± SE
0
2
4
6
8
10
12
14
16
Haul seine
Trawl
B
Total taxa observed
0
25
50
75
100
125
150
175
200
Haul seine
Trawl



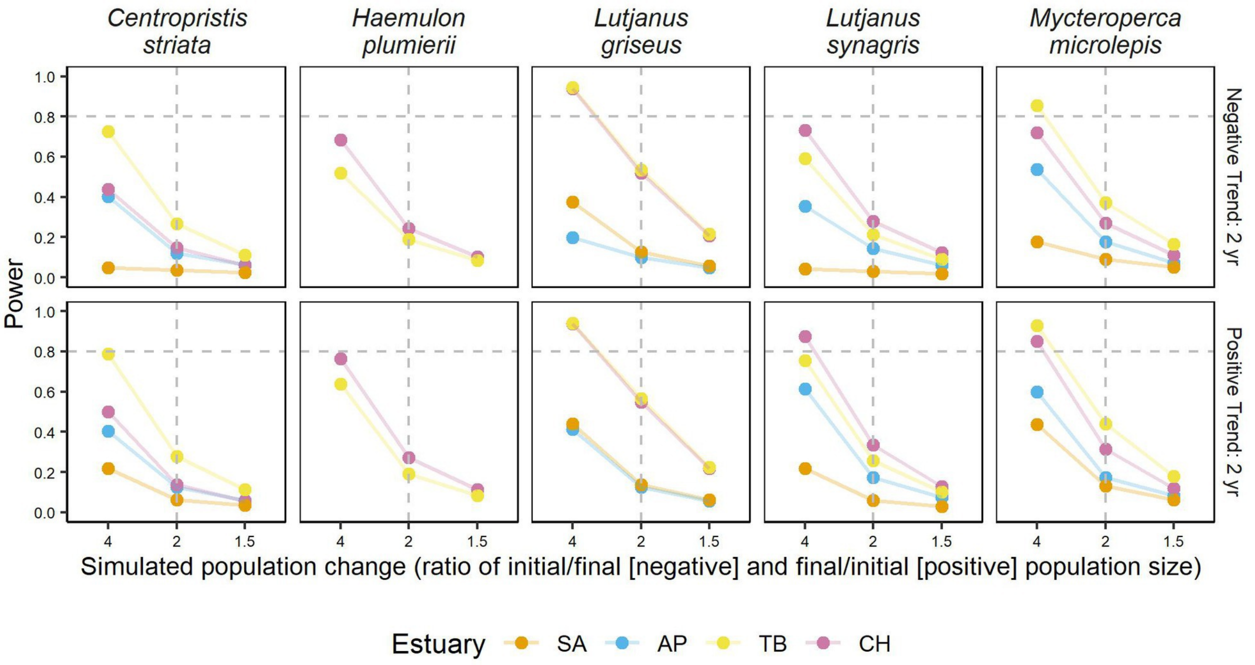
Haul seine
Trawl
Centropristis striata
0.003
0.059
0.034
0.014
0.085
0.577
1.397
0.223
0.056
Epinephelus morio
0.001
0
0
0.001
0.003
0
0
0.004
0.023
Haemulon plumierii
0
0
0.034
0.034
0.038
0.011
0.263
0.214
0.163
Lachnolaimus maximus
0
0.002
0
0
0.001
0.001
0.059
0.004
0.002
Lutjanus griseus
0.024
0.003
0.092
0.173
0.127
0.041
0.007
0.268
0.453
Lutjanus synagris
0.007
0.013
0.023
0.055
0.131
0.482
0.161
0.311
0.493
Mycteroperca microlepis
0.008
0.03
0.042
0.036
0.023
0.037
0.012
0.072
0.087
CPUE (fish/100 m2) ± SE
SA
AP
BB
TB
CH
Estuary

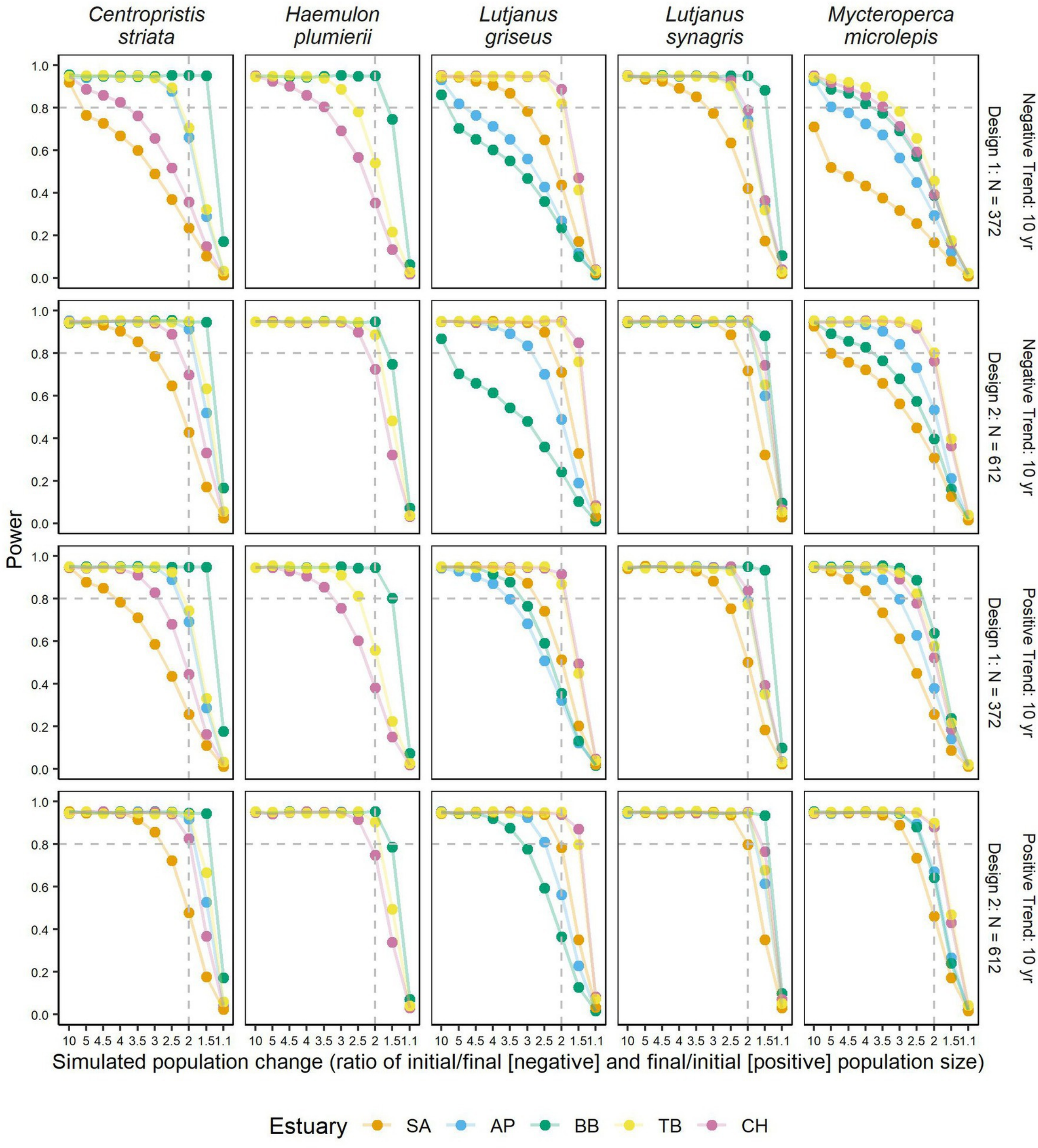


Centropristis striata
Epinephelus morio
Haemulon plumierii
Lachnolaimus maximus
Lutjanus griseus
Lutjanus synagris
Mycteroperca microlepis
Count
Standard length (mm)
Gear type
Haul seine
Trawl
0
200
400
600

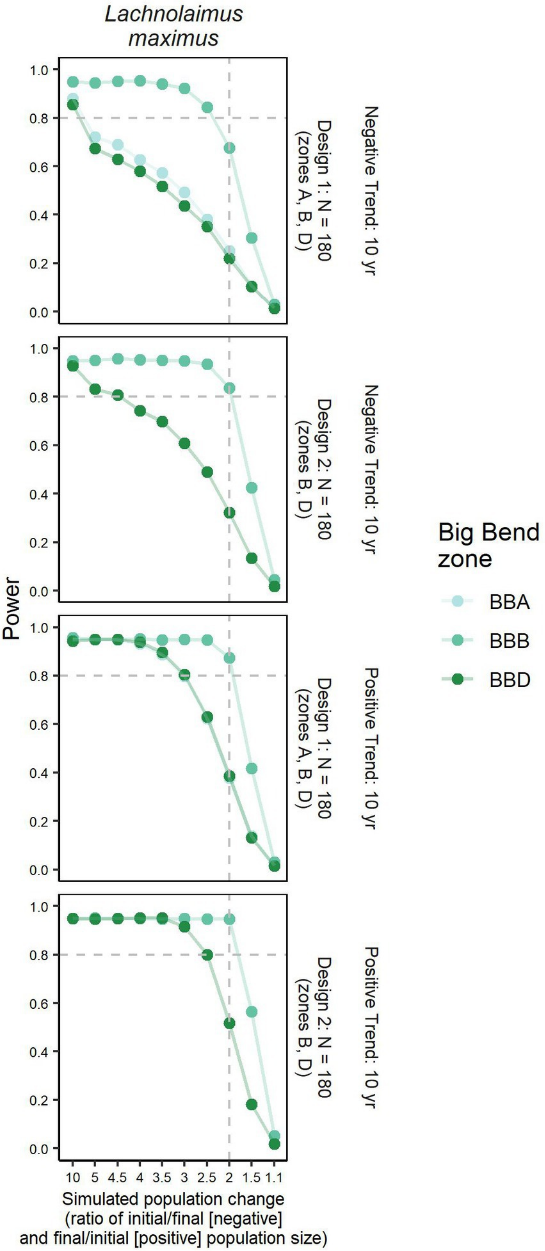


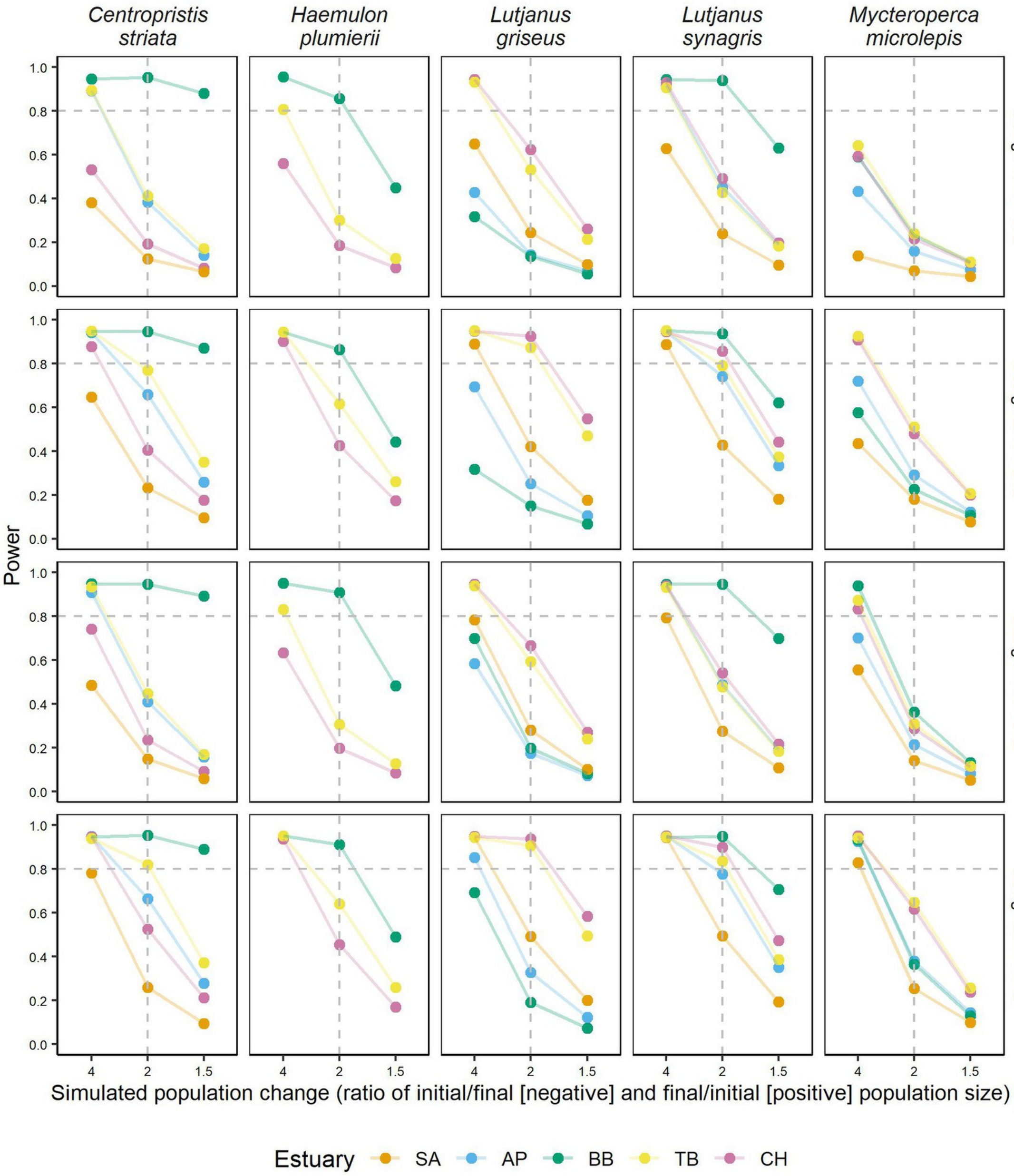
Centropristis striata
Haemulon plumierii
Lutjanus griseus
Lutjanus synagris
Mycteroperca microlepis
Negative Trend: 10 yr
Positive Trend: 10 yr
Power
1.0
0.8
0.6
0.4
0.2
0.0
10 5 4.5 4 3.5 3 2.5 2 1.5 1.1
Simulated population change (ratio of initial/final [negative] and final/initial [positive] population size)
Estuary
SA
AP
TB
CH

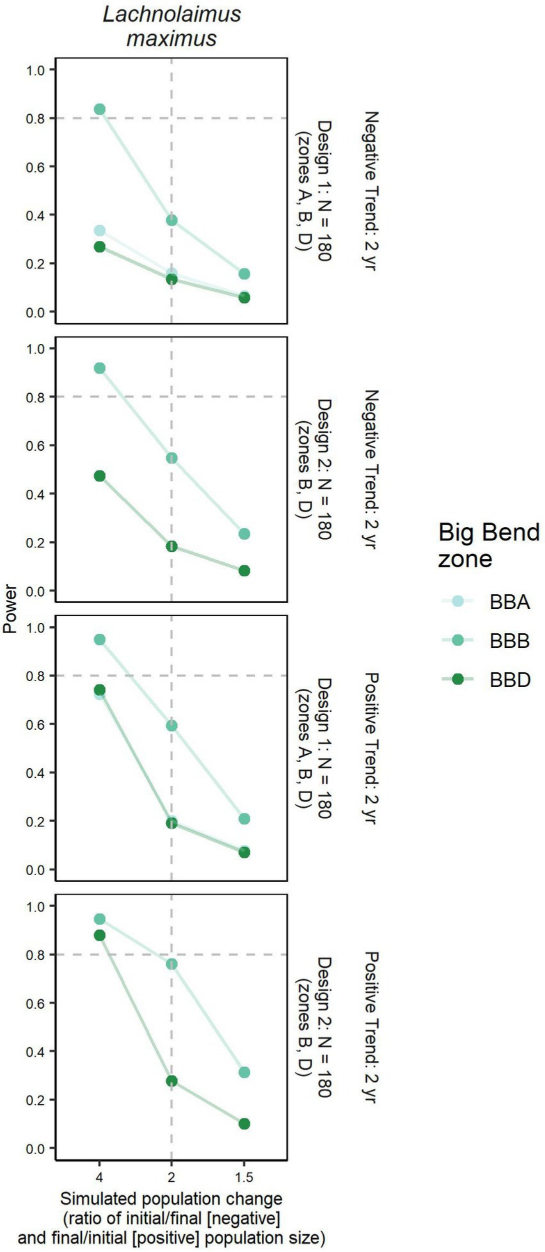

Centropristis striata
Haemulon plumierii
Lutjanus griseus
Lutjanus synagris
Mycteroperca microlepis
Negative Trend: 2 yr
Positive Trend: 2 yr
Power
1.0
0.8
0.6
0.4
0.2
0.0
4
2
1.5
Simulated population change (ratio of initial/final [negative] and final/initial [positive] population size)
Estuary
SA
AP
TB
CH



Centropristis striata
Haemulon plumierii
Lutjanus griseus
Lutjanus synagris
Mycteroperca microlepis
Negative Trend: 10 yr
Design 1: N = 372
Negative Trend: 10 yr
Design 2: N = 612
Positive Trend: 10 yr
Design 1: N = 372
Positive Trend: 10 yr
Design 2: N = 612
Power
Simulated population change (ratio of initial/final [negative] and final/initial [positive] population size)
Estuary
SA
AP
BB
TB
CH

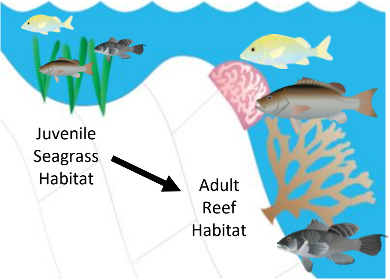

Lachnolaimus maximus
Power
Negative Trend: 10 yr
Design 1: N = 180 (zones A, B, D)
Design 2: N = 180 (zones B, D)
Positive Trend: 10 yr
Big Bend zone
BBA
BBB
BBD
0.0
0.2
0.4
0.6
0.8
1.0
10 5 4.5 4 3.5 3 2.5 2 1.5 1.1
Simulated population change (ratio of initial/final [negative] and final/initial [positive] population size)

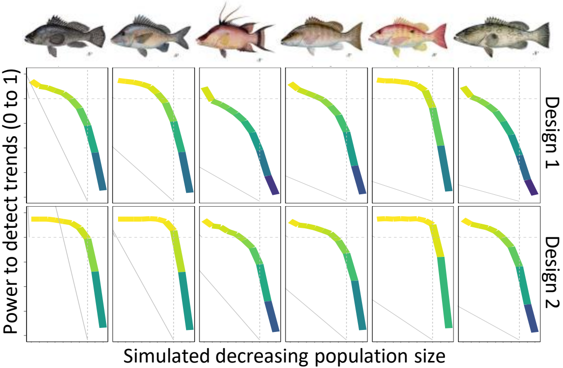

Centropristis striata
Haemulon plumierii
Lutjanus griseus
Lutjanus synagris
Mycteroperca microlepis
Negative Trend: 2 yr Design 1: N = 372
Negative Trend: 2 yr Design 2: N = 612
Positive Trend: 2 yr Design 1: N = 372
Positive Trend: 2 yr Design 2: N = 612
Power
Simulated population change (ratio of initial/final [negative] and final/initial [positive] population size)
Estuary
SA
AP
BB
TB
CH



Lachnolaimus maximus
Power
Negative Trend: 2 yr
Design 1: N = 180 (zones A, B, D)
Design 2: N = 180 (zones B, D)
Positive Trend: 2 yr
Simulated population change (ratio of initial/final [negative] and final/initial [positive] population size)
Big Bend zone
BBA
BBB
BBD

Juvenile reef fish survey design: achieving statistical power targets





Images courtesy of the Integration and Application Network, University of Maryland Center for Environmental Science (ian.umces.edu/symbols/).