Prior to the Gulf Council's Scientific and Statistical Committee meeting in September 2021, the Southeast Fishery Science Center (SEFSC) requested our group re-analyze the Red Snapper abundance estimates for Florida by re-incorporating the Random Forest model (RF) based on the sampling regime for that region. During the prior external review in March 2021, an independent team of experts recommended removing the RF model for Florida and analyzing the data based on a stratified random design. As a result, the estimate in the final report, provided to the Mississippi-Alabama Sea Grant Consortium in August 2021, stratified Florida by region and depth, resulting in an estimate of 118 million and 111 million Red Snapper (depending on calculation method). With the most recent request in September 2021 by the SEFSC, our team concluded that the Florida estimate should have included the RF model. To accommodate their request, we re-analyzed the data once again, resulting in an abundance estimate of 97 million and 92 million Red Snapper. We have provided the methods for the abundance calculation and the results of the re-analysis in detail below.

The methodology below has been directly taken from the final report (pp. 81-87) and modified to include how the random forest model (RF) was incorporated in the re-analysis for Florida's UCB/Natural bottom. For additional information, please refer to the full final report available at www.snappercount.org.

C. Final Abundance Estimates

Estimates of age-2+ Red Snapper abundance were produced by region, habitat type, and depth. Where appropriate, population estimates for artificial reefs were made for various categories representing the diversity of artificial structures. In all cases, population estimates were derived by expanded mean densities, with means and variances calculated assuming simple random sampling at the lowest strata level and assuming no error in the individual sample site estimates. Means and variances at higher levels of aggregation (region, total) were calculated following stratified sampling methods. Estimates were performed by two independent groups on the same data to provided cross validation. While the approaches, post-stratification, and application of statistical models differed and were not stipulated a priori, these separate analyses converged with very similar estimates. Overall, this most recent estimate (incorporating the RF model for Florida's UCB/Natural bottom) of absolute abundance was 97 and 92 million age-2+ (percent standard error (PSE) 15%) during late 2019 (dependent on calculation method). While large numbers of fish occurred over well-known habitat features such as artificial reefs and natural hard bottom, we found that the previously uncharacterized bottom habitat (UCB) harbored the majority of Red Snapper. As a result of stratification, by region and depth for the dominant UCB habitat, the estimated PSE for the overall estimate is lower compared to the subcomponents.

What follows is a detailed description of how the team arrived at our final estimate of absolute abundance (Table 1) of Red Snapper by region and habitat type.

1. Abundance Estimates by Region and Habitat Type

Due to the paucity of classified bottom habitat in the Gulf, the majority of habitat fell into the UCB category which was stratified by region and depth. UBC was stratified by state (TX, LA, AL/MS, FL) and depth (10-40 m, 40-100 m, 100-160 m). The FL strata was further subdivided into 3 regions (northwest, mid, south) and a RF model using 3 estimates of the likelihood of Red Snapper occurrence resulted in 27 strata used to determine the weights for the stratified estimates of mean density. For some locations (TX, LA, AL/MS) the areas of well-known large features of hardbottom were removed from the UCB estimates. Where hardbottom habitat was mapped in detail, population estimates were made for the mapped area by region. Population estimates were also made for artificial structures and the subcategory of artificial structure pipelines. The overall population estimate was derived by summing over the individual categories. Estimated densities and numbers per sampled strata (habitat and depth) by region are presented in Table 1.

Uncharacterized bottom

To estimate total population size and uncertainty for each stratum, observed numbers of Red Snapper per 100m^2 within a stratum were treated as simple random samples and population estimates were calculated as the mean density times the number of 100m^2 sampling units in each stratum. Mean density estimate were treated differently depending on the region and sampling method.

In FL, strata were defined by region (northwest, mid, south), with the South region being established post hoc, depth and classification criteria from a RF model (see Siders manuscript, provided below). The RF model used fishery independent and dependent data to determine the probability of Red Snapper presence and categorized them as low, medium, and high. This resulted in 27 strata. In two instances strata samples were not taken and mean and variances were borrowed from similar strata. These substitutions occurred in the Central region where the shallow depth low probability values were used for the mid depth low probability strata and the mid depth high probability values were used for the deep depth high probability strata. Density was estimated from randomly selected ROV point counts where 100% detection was observed at the most basic level (region, depth). Strata specific mean ($\bar{x}_h$) and variance (s_h^2) could be calculated following equations 1 and 2. The number of sampling units in a stratum (N_h) relative to the total number of sampling units (N) are used in the estimation of the stratified mean ($\bar{x}$) following equation 3 where K is the number of stratum and $\frac{N_h}{N}$ is the stratum weight. The variance of the random stratified mean ($s_{\bar{x}}^2$) is a function of the stratum weight, the number of observations in a stratum (n_h), the stratum variance, and the finite population correction and was calculated using equation 4. To estimate total population size (T), the random stratified mean is expanded by the total number of sampling units (N).

(1) $\bar{x}_h = \sum_{i=1}^{n} \frac{1}{n} x_i$

(2) $s_{x_h}^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x}_h)\text{^}2}{n-1}$

(3) $\bar{x} = \sum_{h=1}^{K} \frac{N_h}{N} \bar{x}_h$

(4) $s_{\bar{x}}^2 = \sum_{h=1}^{K} \left[\left(\frac{N_h}{N} \right)^2 \left(1 - \frac{n_h}{N_h} \right) \frac{s_h^2}{n_h} \right]$

In TX, for the 2 shallowest stratum, where acoustic counts were taken, the total number of fish encountered on a transect over the total area covered by the acoustic gear was used as the density estimate. Transect were post hoc stratified by region (South, Central, and North) to accommodate region specific estimates of the proportion of Red Snapper in an acoustic estimate of numbers of fish. Transects were assumed to be selected randomly within strata with mean and variance calculated following equations 1 and 2. To account for region specific estimates of the proportion of Red Snapper in a sample and the uncertainty associated with this estimate (Table 1) the standard equation for the variance of the product of two independent variables was used. For each region the mean density of fish ($\bar{x}_h$) was multiplied by the mean proportion of Red

Snapper ($\bar{y}_h$) estimated to be in the sample from visual surveys. The resulting variance was calculated using equation 5.

(5) $$s_{xy}^2 = s_{x_h}^2 * s_{y_h}^2 + s_{x_h}^2 * \bar{x}_h + s_{y_h}^2 * \bar{y}_h$$

For the deepest TX stratum as well as UCB estimates for LA, estimates of the total number of snapper along a transect over the area surveyed from CBASS camera tows were used for density estimate. These estimates assumed 100% detection of Red Snapper within the area surveyed. Each transect was randomly selected and equations 1 and 2 were used to estimate the mean density and variance within a region. Due to the low number of transects in LA and MS/AL waters, density estimates were not made for each depth stratum.

To estimate stratum specific population size (T_h) mean density per 100m^2 ($\bar{x}_h$) was multiplied by the number of 100m^2 units within a given stratum (equation 6) with the associated variance calculated using equation 7. To estimate regional total population sizes as well as the overall Gulf of Mexico population over UCB, strata specific mean density and associated variances were combined using equations 3 and 4 with the stratum weight based on the area of each stratum (note for LA and MS/AL no depth stratification was used).

(6) $$T_h = N_h * \bar{x}_h$$

(7) $$s_{T_h}^2 = N_h^2 * s_{\bar{x}_h}^2$$

The resulting estimates of mean Red Snapper density and variance were for each region and depth stratum were combined into a single depth specific mean following equations 3 and 4 for stratified sampling.

Natural Hard Bottom

Population estimates for natural hard bottom were calculated as expanded mean densities assuming the data were collected from a simple random sample. Mean and variances were calculated using equations 1 and 2 and expanded to the mapped area of hard bottom for TX, LA, and AL/MS using equations 6 and 7. For LA, samples from TX were substituted.

Artificial structures

Artificial structures in TX were categorized as small and extra-large. In the small category, it was assumed that 3.25 pyramids comprised a small structure given the nature of the sampling conducted. Structures were also categorized by depth strata. Within each category simple random sampling was conducted and mean numbers and associated variance per structure were estimated using equations 1 and 2 from total fish counts converted to Red Snapper numbers from site specific estimates of the proportion of red snapper. For each site, the proportion Red Snapper was assumed know without error. Total population estimates were calculated from expanded mean numbers per structure expanded by the assumed known number of structures (equation 6). To estimate total number per artificial structure category, mean density per structure was calculated using depth strata following equations 3 and 4 with total numbers

estimated from expanded mean numbers per structure tie the assumed know number of structures.

Population estimates for LA were estimated from data for TX. All structures in Louisiana were assumed to be extra-large and the number of structures was assumed known without error. Depth specific TX data was substituted to estimate depth specific mean densities and calculations for each stratum and the combined estimates were calculated similar to the TX data.

For AL the number of artificial structures per depth strata was estimated (see detail in the methods section). As a result, for each depth strata the total variance in the estimate was calculated by combining the variance in the estimated mean numbers per structure times the estimate of the variance in the number of structures (equation 5). Within each category (depth and authorization zone) simple random sampling was conducted and the mean and variance in numbers per structure were calculated using equations 1 and 2. Samples were stratified by authorization zone category to obtain estimated numbers in a given depth category. Means and variances were calculated using equations 3 and 4. Total numbers were estimated from expanded mean numbers per structure and the estimated number of structures (see methods for greater detail). For MS, estimates of number per structure were calculated from simple random samples and expanded assuming the number of structures was known without error.

No difference in mean densities and variance were apparent for structures in FL and all samples were combined to get a single mean number per structure assuming simple random sampling (equations 1 and 2). The number of artificial structures by depth was assumed known without error and total population estimated by depth and region were estimated as expanded mean numbers (equation 6 and 7).

Pipelines

The population estimate of Red Snapper on pipelines was estimated from an expanded mean densities per meter of pipeline times the total extent of pipeline in the Gulf, calculated from georeferenced polyline data of operational pipeline from the BOEM (Bureau of Safety and Environmental Enforcement, Office of Technical Data Management, Data Administration Unit 2020-12-01, Pipelines vector digital data). Red Snapper density per meter of pipeline was estimated from total video counts per transect, assuming 100% detection of Red Snapper, over the length of pipeline surveyed. Mean density and the associated variance were calculated assuming transects were selected at random out of the available pipeline units in the BOEM database.

Table 1. Re-analysis of the Florida natural/unconsolidated bottom-type data to include the random forest design stratification, resulting in a decrease of approximately 21 million fish from the previous estimate of 118 million Red Snapper.

State/Region	Habitat Type	Total Area (km^2) or Structures	Number of Samples (*n*)	Area Sampled (km^2)	Mean Density ($100m^2$) or by Structure	Number	SE	CV (%)
TX	Natural	1,570	36	6.13	0.45	7,037,443	2,537,014	36
	Artificial	4,348	49			417,761	88,469	21
	Large	*941*	*45*		*362*	*340,905*	*79,287*	*23*
	Small	*3,460*	*4*		*22*	*76,855*	*39,246*	*51*
	Uncharacterized Bottom	57,535	140	6.26	0.03	14,569,830	6,663,776	46
	Total		225			22,025,035	7,130,931	32
LA	Natural	821	22	*n/a*	0.47	3,852,652	1,671,470	43
	Artificial	1,771	42		2174	3,849,325	576,234	15
	Uncharacterized Bottom	53,052	87	3.61	0.02	9,729,387	5,699,448	59
	Total		151			17,431,364	5,967,375	34
AL&MS	Natural	211	32	0.013	1.78	3,751,988	752,467	20
	Artificial	9,410	128		160	1,509,625	167,506	11
	Uncharacterized Bottom	18,500	3	0.74	0.02	3,199,472	1,625,263	51
	Total		163			8,461,085	1,798,817	21
FL	Natural & Uncharacterized	143,538	748	0.61	0.03	48,124,414	10,437,839	22
	Artificial	7,763	79		16	127,560	21,088	17
	Total		832			48,251,974	10,437,861	22
Pipelines (Gulf-wide)		26,686 linear km	27	0.49	0.02	507,661	218,961	43
Gulf of Mexico						**96,677,119**	**13,969,084**	**14**

2. Validation Analysis for Abundance Estimate

We performed a separate independent analysis to validate our primary estimate of absolute abundance on the same data set to provide validation. The results of this secondary analysis are shown in Table 2. While the approaches, post-stratification, and application of statistical methods differed somewhat and were not stipulated *a priori*, these independent analyses produced similar estimates (i.e., within 4.7%; 4.6 million Red Snapper difference from each estimate).

While these two analyses were performed independently using the same data, guidance was not given in terms of a preferred statistical approach, post-stratification, and various other small nuances regarding how these data were treated. Total abundance estimates were made for 4 regions: Texas, Louisiana, Alabama/Mississippi, and Florida. The primary abundance estimation method for artificial reefs and pipelines is based on a model in which expected abundance in each site is assumed proportional to its area for all sites in the stratum (i.e., it used the average of ratio estimator). The validation method presented here used the standard ratio estimator for abundance, which does not require adherence to a model for consistency. Only small differences in the estimates from the two methods were observed, so the implicit model assumption for the primary estimation method was deemed adequate. Within each region, total abundance was estimated by habitat: artificial reefs (ART), natural banks (NAT), and uncharacterized bottoms (UCB). This section details the different methods used for estimating Red Snapper abundance, data pre-processing, and the mathematical expressions for the different estimators used for estimating total abundance in the various Red Snapper habitats. The rest of the analytical description is organized as follows: the different estimators used to estimate Red Snapper abundance in the different regions/habitats are defined, and the resulting total abundance estimates for each of the regions and habitats as well as their associated estimators are summarized.

Estimators

Within each stratum and post-stratum, a separate estimate of total Red Snapper Abundance was made. Then the estimates of total abundance were summed to achieve a Gulf-wide estimate of abundance. In each stratum or post-stratum, either a Mean-per-unit estimator (if sampling units were the same size or there was no size measure beyond a classification, as for artificial reefs) or a Ratio estimator (if sampling units varied in size, such as varying size transects) was used.

Mean per-unit ($\hat{t}_{y,mpu}$)

In strata in which the sampling unit was artificial structure or grid with fixed size, total abundance was estimated by multiplying the number of artificial structures or grids in the population by the average Red Snapper count per structure or grid (mean per-unit). Let N_h denote the number of units in the stratum *h* universe (e.g., number of large structures in a region) and n_h denote the number sampled, and let y_{hi} denote the abundance of Red Snapper observed (or estimated) in the i^{th} sampled unit of the h^{th} stratum. Then, the total abundance estimate for the stratum is given by:

$$\hat{t}_{hy,mpu} = N_h \times \bar{y}_h, \quad (24)$$

where $\bar{y}_h = average\ abundance\ observed\ (or\ estimated)\ per\ structure\ or\ grid\ \left(\bar{y}_h = \frac{\sum_{i=1}^{n} y_{hi}}{n}\right)$.

The variance of the MPU estimator for the h^{th} stratum (Eqn. 24) was estimated as

$$v(\hat{t}_{hy,mpu}) = N_h^2 \times (s_h^2/n_h)\left(1 - \frac{n_h}{N_h}\right). \quad (25)$$

Ratio Estimator ($\hat{t}_{y,r}$)

In strata in which the sampling units were areal and varied in size (e.g., transects), total abundance was estimated with a standard ratio estimator. Let x_{hi} denote the area of the i^{th} sampled unit of the h^{th} stratum and let t_{hx} denote the total area of the stratum. Then, the total abundance estimate for the stratum is given by:

$$\hat{t}_{hy,r} = t_{hx} \times \frac{\sum_{i=1}^{n_h} y_{hi}}{\sum_{i=1}^{n} x_{hi}} = t_{hx} \times \hat{d}. \quad (26)$$

The variance of the ratio estimator for the h^{th} stratum (Eqn. 26) was estimated using the Taylor Series approximate variance:

$$v(\hat{t}_{hy,r}) = t_{hx}^2 \times (s_d^2/n_h)\left(1 - \frac{n_h}{\widehat{N}_h}\right), \quad (27)$$

where s_d^2 is the sample variance of the residuals $d_{hi} = y_{hi} - \hat{d}x_{hi}$ and the estimated number of transects in the population is $\widehat{N}_h = t_{hx}/\bar{x}_h$, where $\bar{x}_h$ is the average area of the sampled units.

Pyramid Structure strata ($\hat{t}_{y,r(pyr)}$)

Artificial structures in TX were classified into two strata, pyramid-like and non-pyramid, because structures are typically large artificial reef (e.g., oil and gas platforms) or smaller artificial reef pyramids (e.g., small discrete structures). These two types required different approaches for estimating abundance. Though abundance of Red Snapper on large structures in Texas were estimated with the mean-per-unit estimator shown in (24), the total abundance on the pyramid structures was estimated by a ratio estimator, as shown in (26). The regions where the pyramids appear was gridded into equal size grid cells. Then a sample of n_h grids cell was selected. However, rather than using the area as the auxiliary variable, the number of pyramids in each grid cell was used. That is, $x_{hi} =$ the number of pyramids in grid unit *i* in the stratum and the total number of pyramids in the stratum is denoted by t_{hx}. Then total abundance was estimated using the ratio estimator as in (26). Note though that the density estimate $\hat{d}_{h,}$ is now the density of Red Snapper per pyramid in the sampled grids. The variance of this estimator is as shown in (27).

Substitution ($\hat{t}_{y,sub}$)

In regions in which samples were not available or missing, total abundance was estimated by substituting the missing samples with samples from similar/nearby areas. The total abundance estimate is:

$$\hat{t}_{hy,sub} = t_{hx} \times \hat{d}_{h,sub}$$

where $\hat{d}_{h,sub}$ is the abundance density for the area where sample is available (the substitute area).

Alabama/Mississippi Estimates

The one exception to the method just described was for Alabama/Mississippi estimates. The AL/MS team produced estimates and their standard errors directly, which are reported in Section **Error! Reference source not found.** (Alabama/Mississippi Region). The validation estimation team incorporated their estimates into the Gulf-wide total Red Snapper estimate and its variance, using the method we describe subsequently.

Adjustment for calibration variance

The estimated variance expressions in (25) and (27) do not account for uncertainty in the measurement of RS abundance y_{hi}. The so-called "observed" values of Red Snapper count in expressions (24) – (26) are in some cases approximated rather than directly observed. One method for approximating Red Snapper was as a fraction of total fish abundance, which was directly observable by using visual sampling methods (e.g., ROV or TCA). This fraction, called a calibration factor, was itself estimated from experimental data in which fish and Red Snapper abundance could both be measured accurately in a sample of transects. From these data, a proportion of Red Snapper was noted for each of a sample of transects. Then the proportions

were averaged to obtain the calibration factor for a specific region. This calibration factor was then multiplied by total fish observed in transects in which the counting technology does not allow species identification, thereby producing an approximate value of Red Snapper counted for the transect. This is the method that was used for the Mid and Shallow depths of the UCB stratum in Texas. A separate calibration factor was estimated by region (Central, North, and South), defining post-strata.

Calibration adds variability to the final estimate beyond what is shown in (25) and (27). To see how much, we must examine the expression for the estimator and calculate an estimate of its variance. Let u_{hi} denote the fish abundance in post-stratum *h* and transect *i* and $\hat{p}_h$ denote the calibration factor for post-stratum *h*, and $\hat{y}_{hi} = \hat{p}_h u_{hi}$ denote the calibrated measure of Red Snapper abundance in transect *i*. Then the calibrated ratio estimator of Red Snapper in the UCB post-strata, is

$$(28) \qquad \hat{t}_{hy,r} = t_{hx} \times \frac{\sum_{i=1}^{n_h} \hat{p}_h u_{hi}}{\sum_{i=1}^{n} x_{hi}} = t_{hx} \frac{\sum_{i=1}^{n_h} u_{hi}}{\sum_{i=1}^{n} x_{hi}} \times \hat{p}_h = \hat{t}_{hu,r} \times \hat{p}_h.$$

From (28) we see that the calibrated estimator can be written as a product of two random variables, one in the form of the original estimator (except it is an estimate of total fish abundance rather than Red Snapper abundance) and the calibration factor. Since the calibration data was independently collected from the fish abundance data, the two terms of the product are independent. The variance for a product of two independent estimators that are both approximately unbiased (so that $E(\hat{t}_{hu,r}) \approx t_{hu}$ and $E(\hat{p}_h) \approx p_{h,}$ the true calibration factor, if it could be observed) can be estimated (Goodman 1962) as

$$(29) \qquad v(\hat{t}_{hy,r}) = V(\hat{t}_{hu,r}\hat{p}_h) = \hat{p}_h^2 V(\hat{t}_{hu,r}) + V(\hat{p}_h)[\hat{t}_{hu,r}^2 - V(\hat{t}_{hu,r}))].$$

Since $\hat{t}_{hy} = \hat{p}_h \hat{t}_{hu,r}$, the first term of (29) can be thought of as an estimate of the variance of the uncalibrated estimator in (28). The second term of (29) is therefore an estimate of the increase in variance due to calibration. When the calibration factor is a sample mean (of proportions) as it is in this case, then $v(\hat{p}_h) = s_{hp}^2/m$, where s_{hp}^2 is the sample variance of the calibration proportions and *m* is their sample size. This is the method we used to determine the SE's for the estimates of total Red Snapper in Table 7 for the Mid and Shallow UCB strata. (This method was not used for the Deep UCB stratum because calibration was not used, but rather direct counts of Red Snapper were used for estimation where the CBASS gear was used).

Note that the AL/MS estimation team incorporated an adjustment for the uncertainty in the number of artificial reefs in their state, as was described in Section **Error! Reference source not found.** (Alabama/Mississippi Region). Since the number of artificial reefs was unknown, the expression in (24) also required a product of two random variables for their estimator. As a result, they also used the variance estimate shown in (29), as shown in Section **Error! Reference source not found.**.

Besides the Texas UCB, approximation of Red Snapper count in transects of the natural habitats and artificial reefs in Texas also used calibration methods. The analytical methods needed for this calibration are most likely not possible with the current data and analytical methods

available. Thus, no additional variance estimate for this calibration factor was calculated. As a result, we cannot directly assess the effect on the standard error and CV of the estimates of Red Snapper abundance in these strata. Nevertheless, to understand the impact that this calibration might have on the uncertainty of Red Snapper abundance in Texas and the Gulf as a whole, we undertook a conservative "worst case scenario" approach to examine this issue. We estimated the multiplicative increase in variance of Red Snapper abundance due to calibration for each of the post-strata of the UCB in Texas. This quantity, known as design effect or efficiency when comparing sample designs or estimators, ranged from a low of 1.01 (Central region, mid depth of Texas UCB) to a high of 2.77 (South region, mid depth of Texas UCB). The latter value means that the variance of the estimator of Red Snapper abundance in that post-stratum is 2.77 times larger than it would have been if Red Snapper count could have been observed directly, or without uncertainty due to calibration. To examine the impact that calibration might have in the other strata of Texas that used it, we multiplied each variance estimate by 2.77, to determine a "worst-case scenario" for the effect of calibration on variance. Then these conservative estimates of variance were used to determine a CV for Texas, and for its impact on the estimate of total for the Gulf. Our findings, as shown in the last column of Table 7, are that the estimated CV of Red Snapper abundance for Texas increased from 22% to 25% by applying this factor to all the additional strata of Texas that used calibration. Since LA also used Texas data, we carried out this exercise for LA Red Snapper abundance estimate as well. The CV of Red Snapper abundance in Louisiana increased from 23% to 39% by applying the factor to all its strata, also shown in Table 2.

Total Abundance Estimates

To obtain estimates of total abundance for state areas and Gulf-wide, the estimates in the strata and post-strata (which we refer to collectively as sub-areas) making up those areas were added. The estimated variance of the aggregated estimate was calculated as the sum of the variances for the component sub-areas, and its standard error was estimated as the square root of the aggregate. That is, if we denote the set of sub-areas using MPU estimators as H_{mpu} and the set of sub-areas using ratio estimators as H_r, then we can represent the estimator of abundance for any aggregated area A made of entire sub-areas and its standard error as

$$\hat{t}_A = \sum_{h\in(A\cap H_{mpu})} \hat{t}_{h,mpu} + \sum_{h\in(A\cap H_r)} \hat{t}_{h,r}$$

and

$$SE(\hat{t}_A) = \sqrt{\sum_{h\in(A\cap H_{mpu})} v(\hat{t}_{h,mpu}) + \sum_{h\in(A\cap H_r)} v(\hat{t}_{h,r})}\,.$$

The results of this validation estimation process are shown in Table 2 for various aggregation levels, from individual strata to regions to Gulf-wide estimates. It also includes sample sizes and estimates for individual strata, standard errors of the estimates, their coefficient of variation (standard error divided by the estimated abundance), and a conservative (worst-case) CV, based on assuming a large value (2.77) for the design effect for Red Snapper abundance estimates based on calibration. (We do not combine LA and the rest of the Gulf since LA re-uses data from Texas in its estimate. Thus, combining variances as shown in (30) misrepresents the combined uncertainty.)

Table 2. Re-analysis of the Florida natural/unconsolidated bottom-type data to include the random forest design stratification, resulting in a decrease of approximately 19 million fish from the previous estimate of 111 million Red Snapper. This estimate was provided as a separate "validation estimate" in which the same data as provided above was analyzed by an independent group to ensure accuracy in estimate calculations.

State/Region	Habitat Type	Area (km^2) or Structures	Number of Samples (n)	Area Sampled (km^2)	Mean Density ($100m^2$) or by Structure	Number	SE	CV (%)	Conservative CV(%)	Estimator
TX	Natural	1,570	36	6.13		5,218,915	1,390,733	27	44	
	Deep	*209*	*11*		*0.09*	*178,682*	*70,111*	*39*	*65*	$\hat{t}_{y,r}$
	Mid	*953*	*22*		*0.35*	*3,381,753*	*955,545*	*28*	*47*	$\hat{t}_{y,r}$
	Shallow	*409*	*3*		*0.41*	*1,658,480*	*1,008,046*	*61*	*101*	$\hat{t}_{y,r}$
	Artificial	12,010	31			706,327	191,728	27	45	$\hat{t}_{y,r}$
	Pyramids	*10,902*	*13*		*11*	*125,300*	*80,777*	*64*	*107*	$\hat{t}_{y,r(pyr)}$
	Non-Pyramids	*1,108*	*18*		*524*	*581,027*	*173,881*	*30*	*50*	$\hat{t}_{y,r(pyr)}$
	Uncharacterized Bottom	57,535	140	6.22		10,332,018	3,449,733	33	33	$\hat{t}_{y,mpu}$
	Deep	*4,034*	*4*	*1.35*	*0.002*	*71,460*	*38,584*	*54*	*90*	$\hat{t}_{y,r}$
	Mid-North	*8,765*	*39*	*1.75*	*0.015*	*747,705*	*512,361*	*69*	*N/A*	$\hat{t}_{y,r}$
	Mid-Central	*6,450*	*22*	*1.05*	*0.033*	*2,159,374*	*2,014,526*	*93*	*N/A*	$\hat{t}_{y,r}$
	Mid-South	*6,503*	*16*	*0.92*	*0.005*	*340,824*	*205,910*	*60*	*N/A*	$\hat{t}_{y,r}$
	Shallow- North	*17,036*	*36*	*0.51*	*0.014*	*2,335,968*	*1,426,726*	*61*	*N/A*	$\hat{t}_{y,r}$
	Shallow- Central	*8,951*	*15*	*0.38*	*0.038*	*3,367,881*	*2,183,282*	*65*	*N/A*	$\hat{t}_{y,r}$
	Shallow- South	*5,797*	*8*	*0.25*	*0.023*	*1,308,806*	*856,547*	*65*	*N/A*	$\hat{t}_{y,r}$
	Total		198			16,257,260	3,724,454	23	26	$\hat{t}_{y,r}$
LA	Natural	821	22	N/A		3,683,745	958,570	26	43	
	Deep	*105*	*6*		*0.14*	*151,361*	*51,731*	*34*	*57*	$\hat{t}_{y,sub}$
	Mid & Shallow	*716*	*16*		*0.49*	*3,532,384*	*957,173*	*27*	*45*	$\hat{t}_{y,sub}$
	Artificial	1,771	42			3,849,325	1,341,617	35	58	$\hat{t}_{y,sub}$
	Deep	*93*	*7*		*710*	*66,046*	*38,272*	*58*	*96*	$\hat{t}_{y,mpu}$
	Mid	*602*	*29*		*1,399*	*842,219*	*363,261*	*43*	*72*	$\hat{t}_{y,mpu}$
	Shallow	*1,076*	*6*		*2,733*	*2,941,060*	1,290,935	*44*	*73*	$\hat{t}_{y,mpu}$
	Uncharacterized Bottom	53,052	65	2.42		11,043,973	4,024,820	36	61	$\hat{t}_{y,mpu}$
	Deep	*5,348*	*3*	*0.68*	*0.01*	*406,320*	*387,513*	*95*	*159*	$\hat{t}_{y,sub}$
	Mid	*19,077*	*11*	*0.85*	*0.02*	*3,756,598*	*2,715,533*	*72*	*120*	$\hat{t}_{y,sub}$
	Shallow	*28,627*	*51*	*0.89*	*0.02*	*6,881,055*	*2,945,317*	*43*	*71*	$\hat{t}_{y,sub}$
	Total		129			18,577,043	4,349,479	23	39	$\hat{t}_{y,sub}$
AL/MS	Natural	211	32	0.01	1.78	3,751,988	752,467	20	N/A	
	Artificial	9,410	128		160	1,509,625	167,506	11	N/A	
	Uncharacterized Bottom	18,500	3	0.74	0.02	4,425,687	1,730,961	39	N/A	$\hat{t}_{y,r}$
	Total		163			9,687,300	1,894,859	20	N/A	$\hat{t}_{y,r}$
FL	Natural & Uncharacterized	143,538	748	0.61		46,921,038	10,300,890	37	N/A	
	Red Snapper low probability	*92,616*				*14,653,325*	*5,462,227*		*N/A*	
	NW Region- Deep	*1,557*	*13*	*0.009*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	NW Region- Mid	*1,148*	*17*	*0.014*	*0.007*	*81,238*	*82,058*	*101*	*N/A*	$\hat{t}_{y,r}$
	NW Region- Shallow	*2,009*	*23*	*0.024*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	Mid Region- Deep	*3,295*	*2*	*0.001*	*0.000*	*0*	*0*		*N/A*	$\hat{t}_{y,r}$
	Mid Region-Mid	*3,013*	*0*	-	-	*0*			*N/A*	$\hat{t}_{y,r}$
	Mid Region- Shallow	*19,460*	*77*	*0.052*	*0.271*	*5,265,679*	*2,616,464*	*50*	*N/A*	$\hat{t}_{y,r}$
	Southern Region- Deep	*9,871*	*15*	*0.010*	*0.000*	*0*	*0*		*N/A*	$\hat{t}_{y,r}$
	Southern Region- Mid	*18,358*	*13*	*0.013*	*0.315*	*5,786,192*	*3,859,150*	*67*	*N/A*	$\hat{t}_{y,r}$
	Southern Region- Shallow	*33,905*	*53*	*0.048*	*0.104*	*3,520,216*	*2,844,339*	*81*	*N/A*	$\hat{t}_{y,r}$
	Red Snapper probable	*28,065*				*15,454,698*	*5,838,704*		*N/A*	$\hat{t}_{y,mpu}$
	NW Region- Deep	*98*	*7*	*0.005*	*0.211*	*20,614*	*20,410*	*99*	*N/A*	$\hat{t}_{y,r}$
	NW Region- Mid	*693*	*7*	*0.006*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	NW Region- Shallow	*1,145*	*11*	*0.008*	*1.847*	*2,115,089*	*2,118,505*	*100*	*N/A*	$\hat{t}_{y,r}$
	Mid Region- Deep	*419*	*2*	*0.001*	*0.000*	*0*	*0*		*N/A*	$\hat{t}_{y,r}$
	Mid Region-Mid	*4,026*	*10*	*0.009*	*1.057*	*4,256,027*	*3,042,427*	*71*	*N/A*	$\hat{t}_{y,r}$
	Mid Region- Shallow	*8,030*	*138*	*0.107*	*1.021*	*8,199,695*	*4,479,071*	*55*	*N/A*	$\hat{t}_{y,r}$
	Southern Region- Deep	*1,928*	*6*	*0.004*	*0.000*	*0*	*0*		*N/A*	$\hat{t}_{y,r}$
	Southern Region- Mid	*9,383*	*10*	*0.016*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	Southern Region- Shallow	*2,343*	*49*	*0.038*	*0.368*	*863,273*	*532,486*	*62*	*N/A*	$\hat{t}_{y,r}$
	Red Snapper high probability	*22,858*				*16,813,015*	*6,494,764*		*N/A*	
	NW Region- Deep	*8*	*6*	*0.004*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	NW Region- Mid	*220*	*5*	*0.004*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	NW Region- Shallow	*399*	*18*	*0.016*	*0.635*	*253,470*	*227,876*	*90*	*N/A*	$\hat{t}_{y,r}$
	Mid Region- Deep	*45*	*0*	-	-	*0*	*0*		*N/A*	$\hat{t}_{y,r}$
	Mid Region-Mid	*5,074*	*10*	*0.011*	*1.418*	*7,195,848*	*5,984,849*		*N/A*	$\hat{t}_{y,r}$
	Mid Region- Shallow	*6,487*	*210*	*0.174*	*1.424*	*9,236,065*	*2,510,522*	*27*	*N/A*	$\hat{t}_{y,r}$
	Southern Region- Deep	*390*	*4*	*0.003*	*0.000*	*0*	*0*		*N/A*	$\hat{t}_{y,r}$
	Southern Region- Mid	*9,301*	*14*	*0.014*	*0.000*	*0*			*N/A*	$\hat{t}_{y,r}$
	Southern Region- Shallow	*932*	*28*	*0.022*	*0.137*	*127,631*	*94,323*	*74*	*N/A*	$\hat{t}_{y,r}$
	Artificial	7,763	84		16	123,377	20,125	16	N/A	$\hat{t}_{y,mpu}$
	Total		832			47,044,415	10,300,910	22	N/A	
Pipelines (Gulf-wide)		26,686 linear km	27	0.49	0.021	546,988	358,761	64	N/A	$\hat{t}_{y,r}$
Gulf of Mexico						**92,113,006**				
TX, MS, AL, FL						73,535,963	13,942,031	15	15	
Louisiana*						18,577,043	4,349,479	23	39	

REPORT ON RANDOM FORESTS APPLICATION TO GULF OF MEXICO RED SNAPPER

Zachary A. Siders
Fisheries and Aquatic Sciences Program, School of Forestry, Fisheries, and Geomatic Sciences, University of Florida, Gainesville, FL

METHODS

Data Integration

We collated 14 presence-absence datasets in the Gulf of Mexico (Table 1)(Figure 1). These broke down by gear into the following categories: 1) camera collected presences (three datasets), 2) vertical line (five datasets), 3) bottom longline (four datasets), 4) bottom trawl (one dataset), and 5) mixed (one dataset). All but two of the 14 datasets were collected using fishery independent surveys with the remainder collected by observers onboard commercial vessels through the NOAA NMFS National Observer Program. One dataset was accessed from the Global Biodiversity Information Facility (GBIF), which is collated from various sources including museum collections and citizen scientists.

Environmental Covariates

We compiled nine environmental covariates: bathymetry (seafloor depth) (Figure 2), distance to shore (Figure 3), distance to submersed aquatic vegetation (SAV) (e.g. seagrass) (Figure 4), distance to hardbottom (rock substrate)(Figure 5), distance to oil and natural gas pipelines (Figure 6), distance to artificial structures (Figure 7), bottom temperature (Figure 9), bottom salinity (Figure 10), and the catch-per-unit effort (CPUE) estimated for the vertical line commercial fishery by vessel monitoring systems (Ducharme-Barth & Ahrens, 2017)(Figure 10). Details on the processing of each variable are provided below and each variable was projected to a 3 x 3 arc-second (roughly 90 x 90 m) grid from roughly 98 – 78°W and 23 – 31°N on a GRS80 ellipsoid with a NAD83 datum, we refer to this grid as the spatial frame hereafter. We chose this resolution as a compromise of computational resource use (the grid has 230,400,000 cells), the size of most artificial structures in the Gulf of Mexico, and the approximate area reasonably surveyed by visual census (Patterson pers. comm.).

Seafloor bathymetry was cropped from the STRM30+ (Version 6) digital elevation model distributed from the Gulf of Mexico Coastal Observing System (http://gcoos.org/products/topography/SRTM30PLUS.html) at a 30 arc-second resolution. Distance to shore was interpolated for the spatial frame by determining the minimum Euclidean distance for a given grid cell centroid to the shoreline extracted at an intermediate resolution (Level 1) from the Global Self-consistent, Hierarchical, Highresolution Geography Database (https://www.ngdc.noaa.gov/mgg/shorelines/data/gshhg/latest/). Distance to SAV was determined as the minimum Euclidean distance for a given grid cell centroid to the edge of any spatial polygon from the Gulfwide Submersed Aquatic Vegetation database (https://gis.ngdc.noaa.gov/arcgis/rest/services/GulfDataAtlas/SAV_Gulfwide/MapServer). Distance to hardbottom was determined as the minimum Euclidean distance for a given grid cell centroid to the edge of any spatial polygon of the rock sediment layer extracted from the dbSEABED database

(http://instaar.colorado.edu/~jenkinsc/dbseabed/). Distance to oil and natural gas pipelines was determined by extracting the Bureau of Ocean Energy Management's (BOEM) Gulf of Mexico Outer Continental Shelf oil and natural gas pipeline dataset (https://www.data.boem.gov/Main/HtmlPage.aspx?page=gomrpipelines) and subsetting to the existing pipelines (removing those that had been removed or proposed). The minimum Euclidean distance between a grid cell centroid and oil and natural gas pipelines was determined.

Distance to artificial structure was determined by first compiling four databases on artificial structures in the Gulf of Mexico and classifying the individual structures into small, medium, large, and extra-large artificial structures: 1) National Oceanographic and Atmospheric Administration (NOAA) Office of Costal Management's Artificial Reef database (AR) (ftp://ftp.coast.noaa.gov/pub/MSP/ArtificialReefs.zip); 2) NOAA Office of Coast Survey's Automated Wreck and Obstruction Information System (AWOIS) (https://nauticalcharts.noaa.gov/data/wrecks-and-obstructions.html); 3) BOEM's Outer Continental Shelf Oil and Natural Gas Platforms – Gulf of Mexico Region dataset (ONGP)(https://data.doi.gov/dataset/outer-continental-shelf-oil-and-natural-gas-platforms-gulf-of-mexico-region-nad-27); and 4) Steinhatchee Fisheries Management Area Artificial Reefs database (SFMAAR). Briefly, the NOAA AR database was parsed by the reef designation into groups (small designed ARs by design type, Ships and ship-like Objects, Sherman WWII Tanks, Convair F-106 Delta Dart Airplanes, Oil and Gas Platforms, Boxcars, Automobiles, Culverts, Bridges, Miscellaneous concrete, and Tires) and using provided weight, weight proxies, or imputation were classified into one of the four artificial structure classes. Vessels that were cross-listed in the NOAA AR and the NOAA AWOIS databases were removed from the latter. Additionally, several oil and natural gas platforms had been converted into artificial reefs under the Bureau of Safety and Environmental Enforcement's Rigs to Reefs program (https://www.bsee.gov/what-we-do/environmental-focuses/rigs-to-reefs) and were removed from the BOEM's ONGP dataset. For artificial structures in the NOAA AWOIS and BOEM ONGP databases, the artificial structure class was assumed to be large. Artificial structure locations were provided confidentially for the "fisheries conservation reefs" in the Steinhatchee Fisheries Management Area in the Big Bend area of Florida (Lindberg, pers. comm.). Reefs were assumed to be consistent with the "Lindberg" reef type in the NOAA AR database and belong to the small artificial structure class.

Bottom temperature and salinity were derived from the HYbrid Coordinate Ocean Model (HYCOM) experiment 31.0 and 32.5 (https://www.hycom.org/dataserver/gom-analysis) for the Gulf of Mexico (representing model predictions from 2009 to 2017). The average across all years was taken, masked to the study region, and depth-matched to the benthos using the STRM30+ bathymetry dataset. The *raster* (Hijmans, 2018), *rgeos* (Bivand & Rundel, 2018), *rgdal* (Bivand, Keitt & Rowlingson, 2018), and *maptools* (Bivand & Lewin-Koh, 2018) packages were used to read, extract, and process the various spatial datasets, the *spatstat* (Baddeley, Rubak & Turner, 2015) package was used for fast Euclidean distance calculations, as well as the *doParallel* (Microsoft Corporation & Weston, 2018) and *foreach* (Microsoft & Weston, 2017) packages were used for parallelization in program *R* (R Core Team 2018).

Random Forest

Random Forest is a supervised machine-learning algorithm based on

classification / decision trees in which classified data is passed through the tree and covariates are used to separate the data into classified groups (Breiman, 2001). In the composite Red Snapper occurrence dataset, samples (location where a gear was deployed or sighting occurred) with an occurrence are dummy-coded as 1's and those without are dummy-coded as 0's. Categorical and continuous covariates are then used (cutoffs applied to continuous datasets) to make decisions at every node in individual classification trees to attempt to best separate the occurrence classes (0's and 1's). A random subset of data is provided for each tree (the in-the-bag dataset) and a random subset of covariates are tried at each node in each tree. Each classification tree votes on whether a datum's class (0's and 1's) and the proportion of trees that votes correctly equal the probability of an occurrence occurring at that sample locale.

Implementation

The base Random Forest procedure, implemented using the *randomForest* package in R (Liaw & Wiener, 2002), was modified to incorporate modeling uncertainty into the final estimates called Ensemble Random Forest (Siders et al. *in prep*). The procedure is roughly equivalent to a k-folds crossvalidation procedure where the composite Red Snapper occurrence dataset is divided into *k* training–test subdatasets. For each subdataset, a Random Forest is trained and its performance evaluated on the test dataset. From this ensemble of random forests, the distribution of model performance metrics and a distribution of model predictions can be determined. In the former, the mean performance across models in the ensemble is calculated. In the latter, the model predictions are used to generate uncertainty in the sample locale predictions and the variable importance. Uncertainty across the study's spatial frame is generated by predicting for each 3 arc-second cell across the northern Gulf of Mexico with each Random Forest in the ensemble.

As we mixed sample surveys with presence-absence and presence-only data, adjustments to the Random Forest algorithm were required. Samples with a Red Snapper occurrence were downsampled so that each tree received a random subset of the same size for samples with and without occurrences, referred to as balancing in machine learning literature. Each Random Forest received a training set containing 90% of the data and a test set of 10%, where each subset received a representative sample (i.e. the proportion of each interaction class was equal to their proportion in the whole dataset). The test set was used to measure model performance. Each Random Forest model had 1000 decision trees and 4 covariates were randomly drawn for each node in each tree. The number of decision trees to fit and the number covariates to try at each node were internally optimized using a single Random Forest model.

Performance metrics

Performance of classification models is typically measured by metrics based on the number of true positives (TP), the number of true negatives (TN), the number of false positives (FP; Type I error), and the number of false negatives (FN; Type II error). Classic metrics are sensitivity, specificity, precision or positive predictive value (PPV), and negative predictive value (NPV). Plotting the 1-Specificity against the Sensitivity creates the Receiver Operator Characteristic Curve (ROC). We calculated the ROC metrics using the *ROCR* (Sing *et al.*, 2005) package in *R*. The area under the ROC curve (AUC) ranges between 0 and 1 with values less than 0.5 performing worse than random, 0.5 equaling random, and typically models with AUCs > 0.7 are deemed useful

(Phillips, Anderson & Schapire, 2006), and models with AUCs greater than 0.95 deemed performing exceptionally well.

Threshold selection

Based on Receiver Operator Characteristic curves, we calculated two thresholds: one splitting high and medium probability of presence classes; and a second splitting medium and low probability of presence classes. We based the threshold choice on previous studies, choosing the maximum sensitivity and specificity metric (Liu, White & Newell, 2013). These initial thresholds were used for designing the sampling strata and locations only.

Validation

We conducted a comparison between the FL Red Snapper counts and the Random Forest predictions to validate the initial thresholds set by the model. To do such, we classified the positive counts of Red Snapper in FL waters using ROV into three bins, using *kmeans* clustering. We assumed that zero counts were assigned to the low threshold. We then conducted a Bayesian Ordinal Logistic Regression using the *arm* package in R (Gelman & Su, 2020) to identify the cutoff values of Random Forest predictions that corresponded to the clusters. These cutoffs values were back-transformed out of logistic space and used to reclassify the gulf-wide predictions of the Random Forest model into low, medium, and high classes. The reclassified Random Forest model predictions were used for subsequent stratified sampling calculations and summations.

RESULTS

The Random Forest model had high performance with exceptionally high AUC values (Figure 11). The predictions from this model resulted into thresholds of 0.23 and 0.77 and were predicted gulf-wide to design sampling strata (Figure 12). A total of 699 sampling locations from the FL shelf were used in the validation (Figure 13). The validation resulted in the thresholds shifting to 0.923 and 0.989, respectively.

REFERENCES

Baddeley A., Rubak E. & Turner R. (2015). *Spatial Point Patterns: Methodology and Applications with R*. Chapman and Hall/CRC, London.

Bivand R., Keitt T. & Rowlingson B. (2018). *rgdal: Bindings for the "Geospatial" Data Abstraction Library*.

Bivand R. & Lewin-Koh N. (2018). *maptools: Tools for Handling Spatial Objects*.

Bivand R. & Rundel C. (2018). *rgeos: Interface to Geometry Engine - Open Source ('GEOS')*.

Breiman L. (2001). Random forests. *Machine learning* **45**, 5–32

Ducharme-Barth N.D. & Ahrens R.N. (2017). Classification and analysis of VMS data in vertical line fisheries: incorporating uncertainty into spatial distributions. *Canadian Journal of Fisheries and Aquatic Sciences* **74**, 1749–1764

Gelman A. & Su Y.-S. (2020). *arm: Data Analysis Using Regression and Multilevel/Hierarchical Models*.

Hijmans R.J. (2018). *raster: Geographic Data Analysis and Modeling*.

Liaw A. & Wiener M. (2002). Classification and Regression by randomForest. *R News* **2**, 18–22

Liu C., White M. & Newell G. (2013). Selecting thresholds for the prediction of species occurrence with presence-only data. *Journal of biogeography* **40**, 778–789

Microsoft Corporation & Weston S. (2018). *doParallel: Foreach Parallel Adaptor for the "parallel" Package*.

Microsoft & Weston S. (2017). *foreach: Provides Foreach Looping Construct for R*.

Phillips S.J., Anderson R.P. & Schapire R.E. (2006). Maximum entropy modeling of species geographic distributions. *Ecological Modelling* **190**, 231–259. https://doi.org/10.1016/j.ecolmodel.2005.03.026

Sing T., Sander O., Beerenwinkel N. & Lengauer T. (2005). ROCR: visualizing classifier performance in R. *Bioinformatics* **21**, 7881

TABLE 1
Summary of datasets used to build the Random Forests model. Code is the dataset identifier, nsamp is the number of samples, nRS is the number of Red Snapper, nnoRS is the number of samples without Red Snapper, nwRS is the number of samples with Red Snapper, propwRS is the proportion of samples with Red Snapper, and data_type is the type of dataset records.

code	nsamp	nRS	nnoRS	nwRS	propwRS	data_type
CAfwri	2377	2715	1751	626	0.26	Presence-Absence
CAms	2147	1355	1716	431	0.2	Presence-Absence
CApc	1294	3482	580	714	0.55	Presence-Absence
GOgbif	834	834	0	834	1	Presence-only
LLcssp	820	709	635	185	0.23	Presence-Absence
LLnmfs	2469	712	2252	217	0.09	Presence-Absence
LLseamap	198	80	190	8	0.04	Presence-Absence
OBll	5462	25161	0	5462	1	Presence-only
OBvl	13556	145619	0	13556	1	Presence-only
TRseamap	37780	96573	28894	8886	0.24	Presence-Absence
VLart	134	509	30	104	0.78	Presence-Absence
VLcssp	1888	1709	1573	315	0.17	Presence-Absence
VLplat	121	509	37	84	0.69	Presence-Absence
VLsea	1022	820	814	208	0.2	Presence-Absence

FIGURE 1

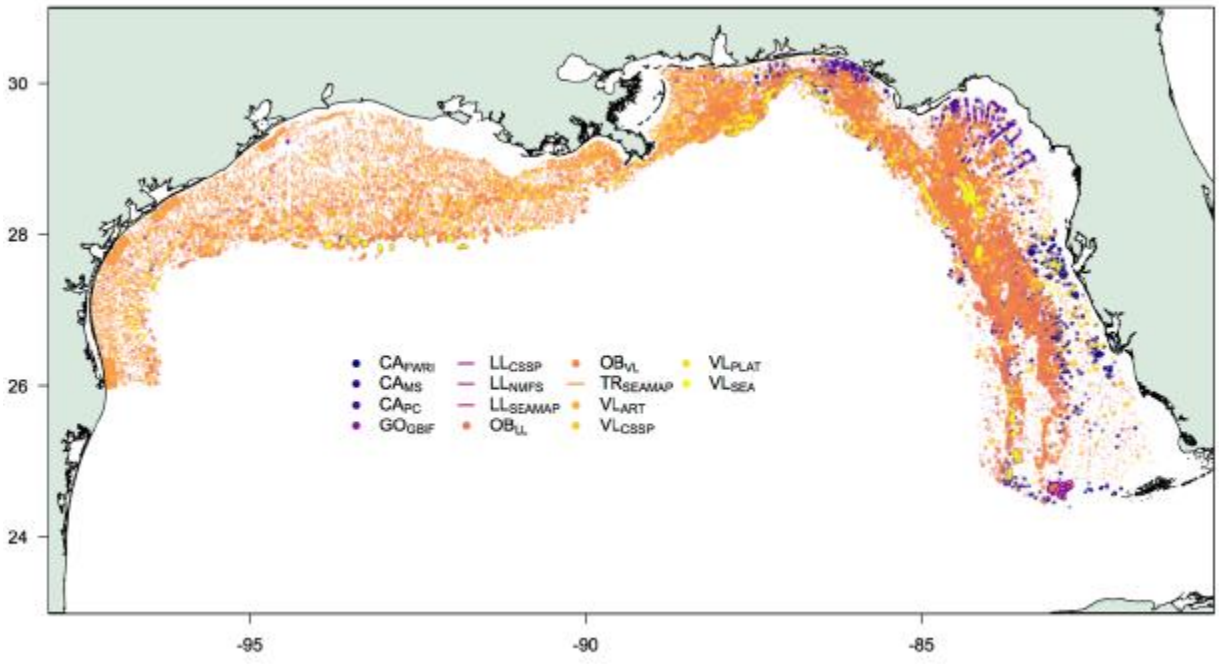


FIGURE 2

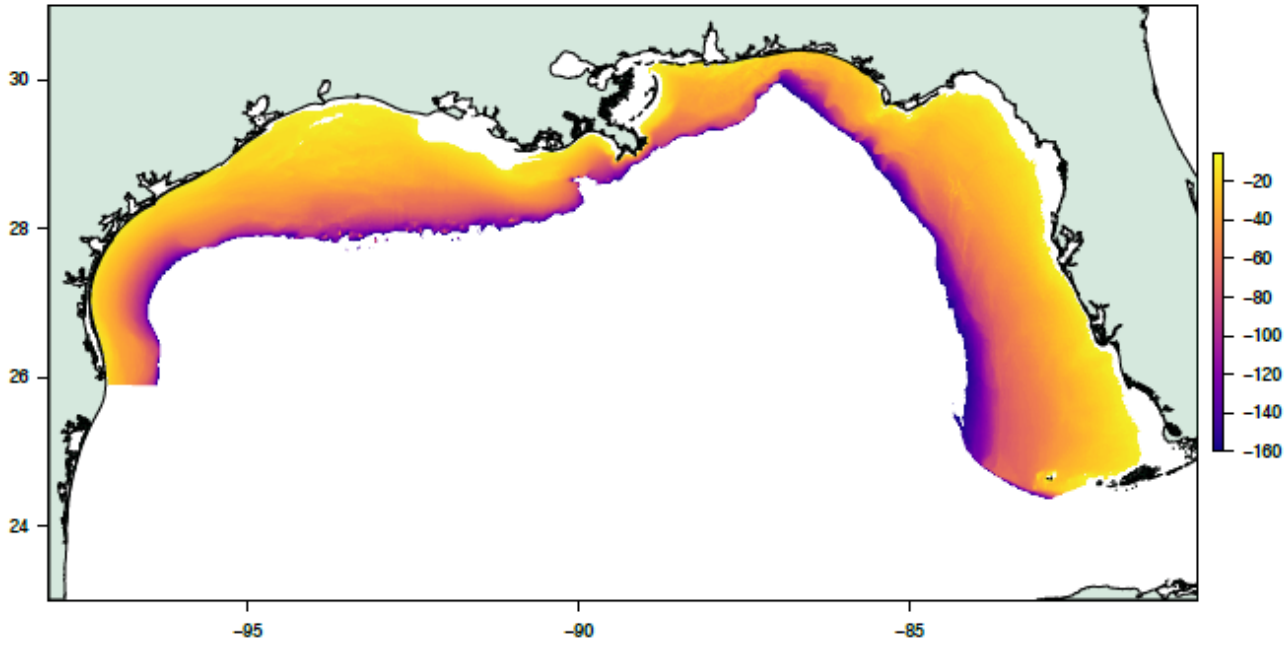


FIGURE 3

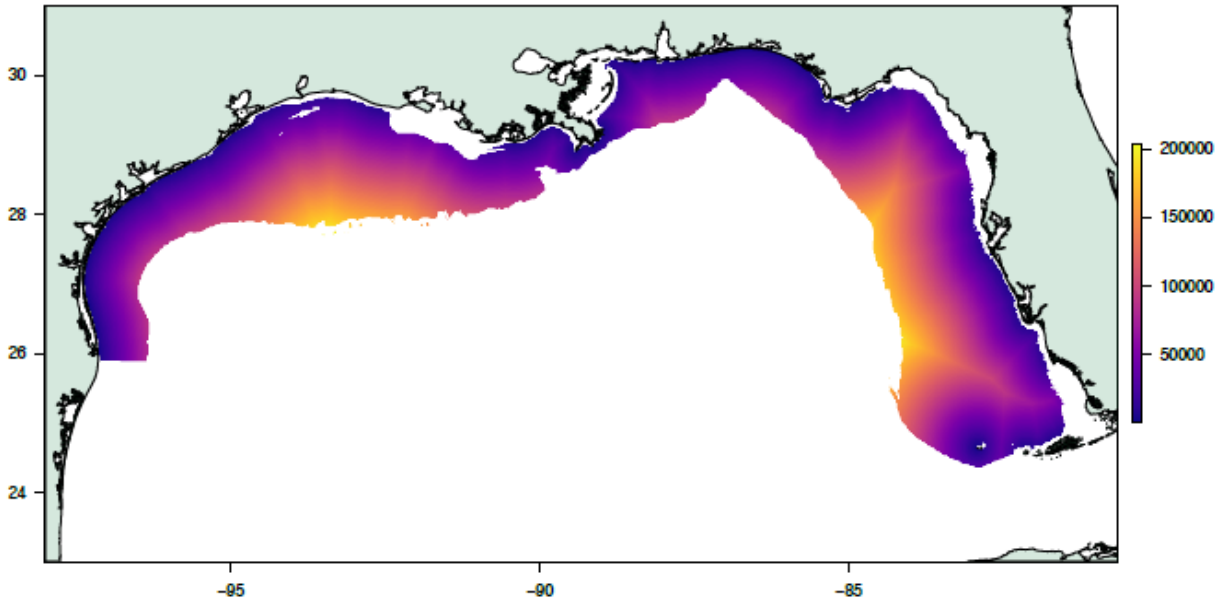


FIGURE 4

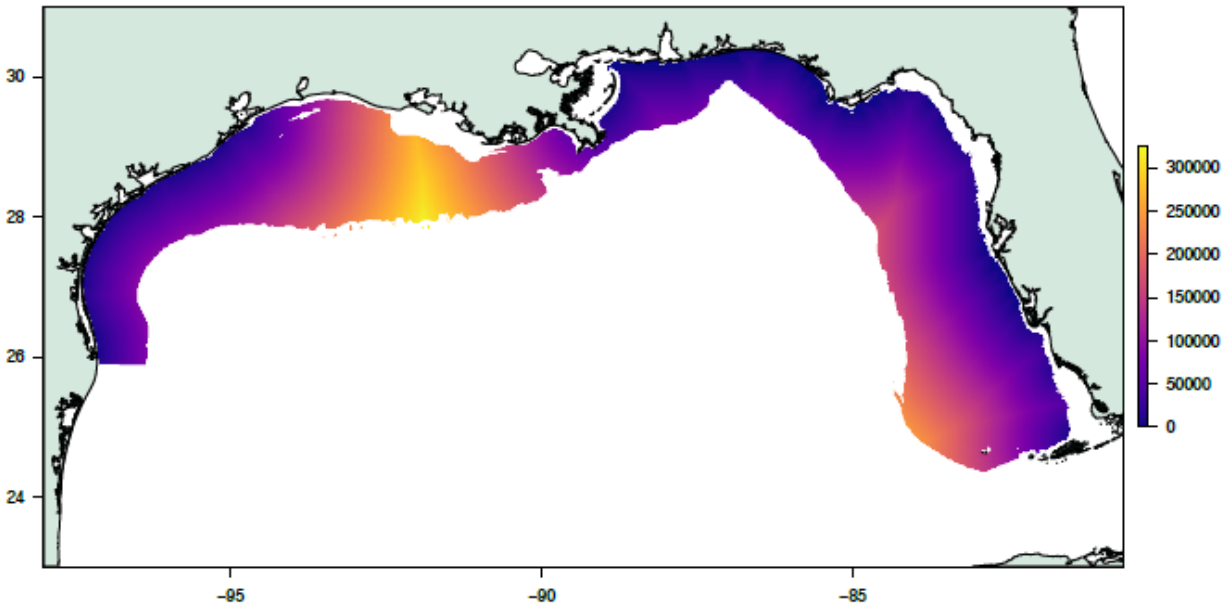


FIGURE 5

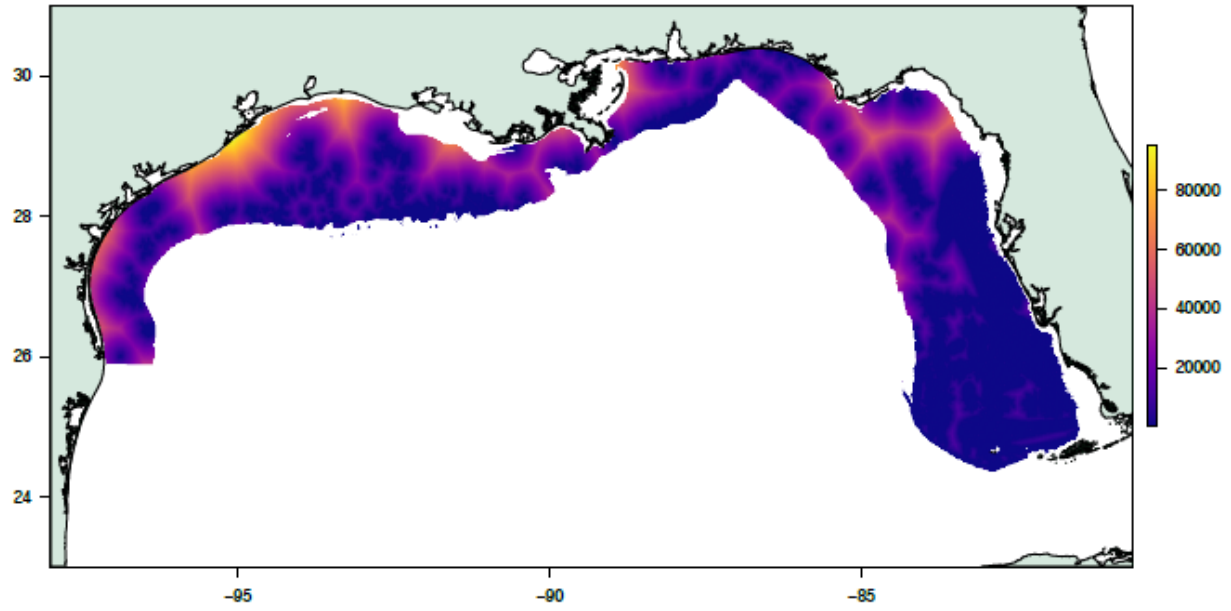


FIGURE 6

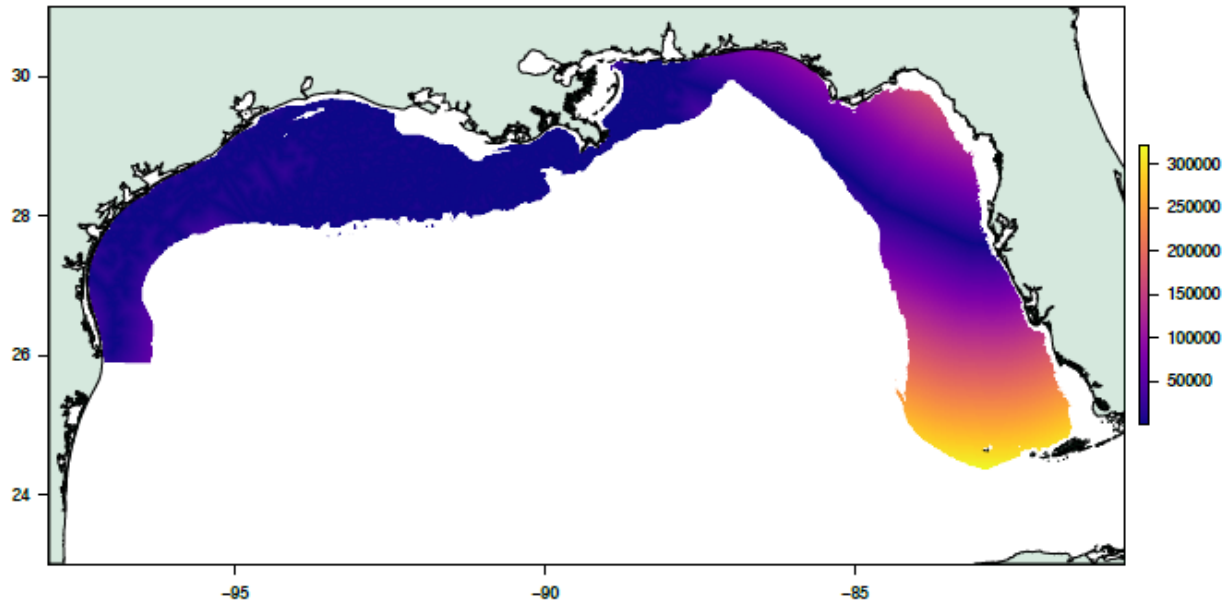


FIGURE 7

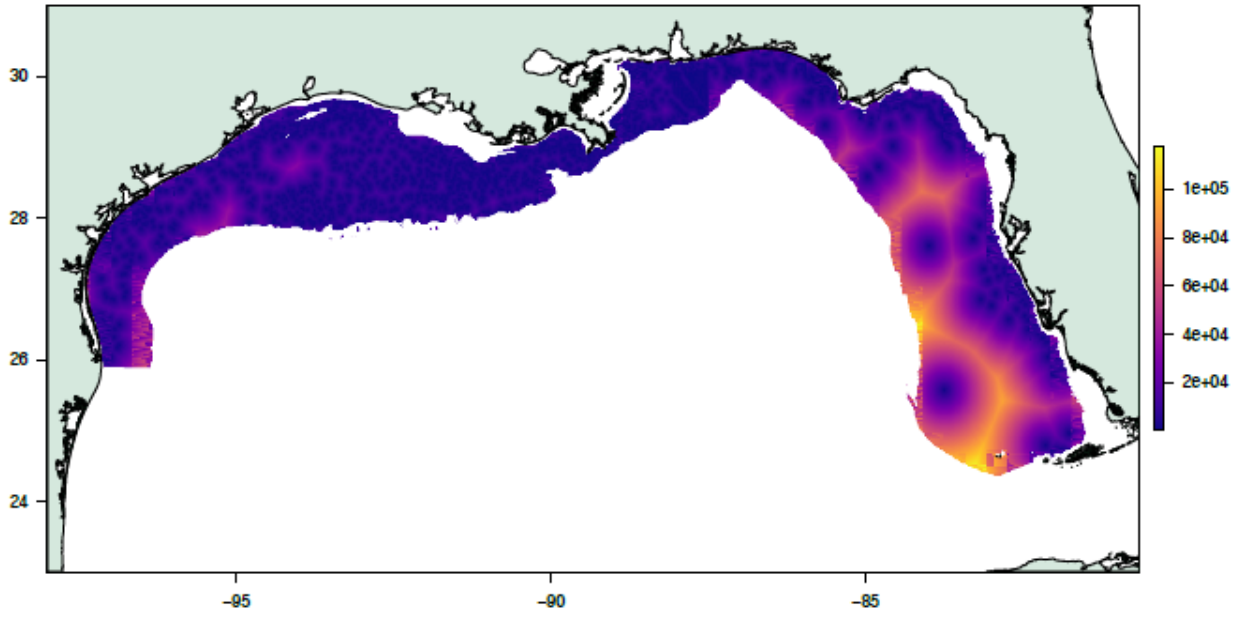


FIGURE 8

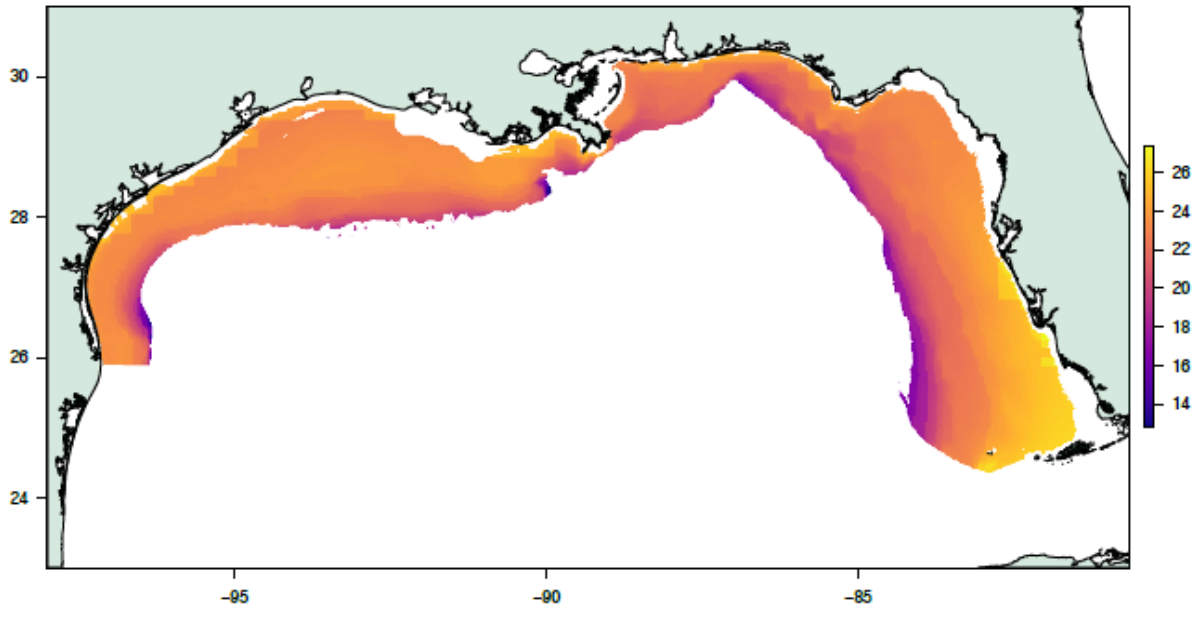


FIGURE 9

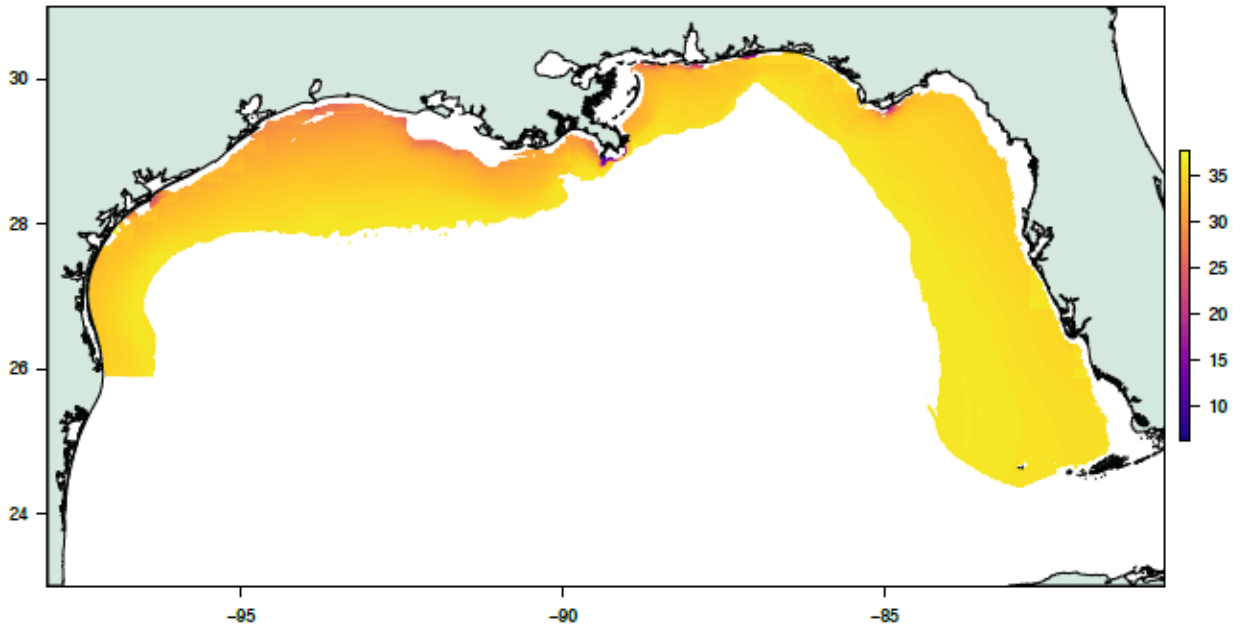


FIGURE 10

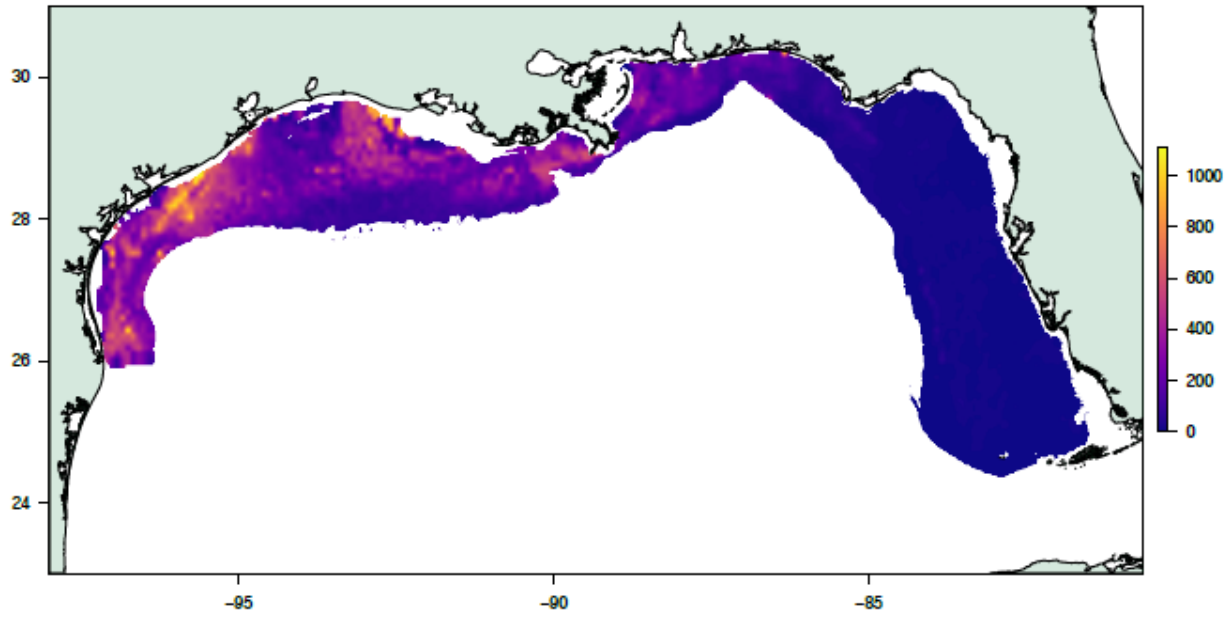


FIGURE 11

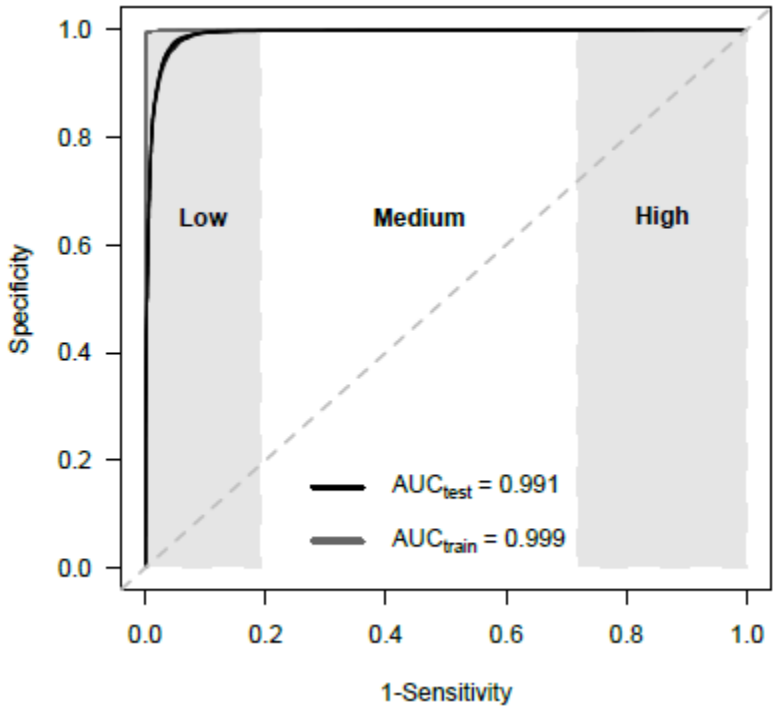


FIGURE 12

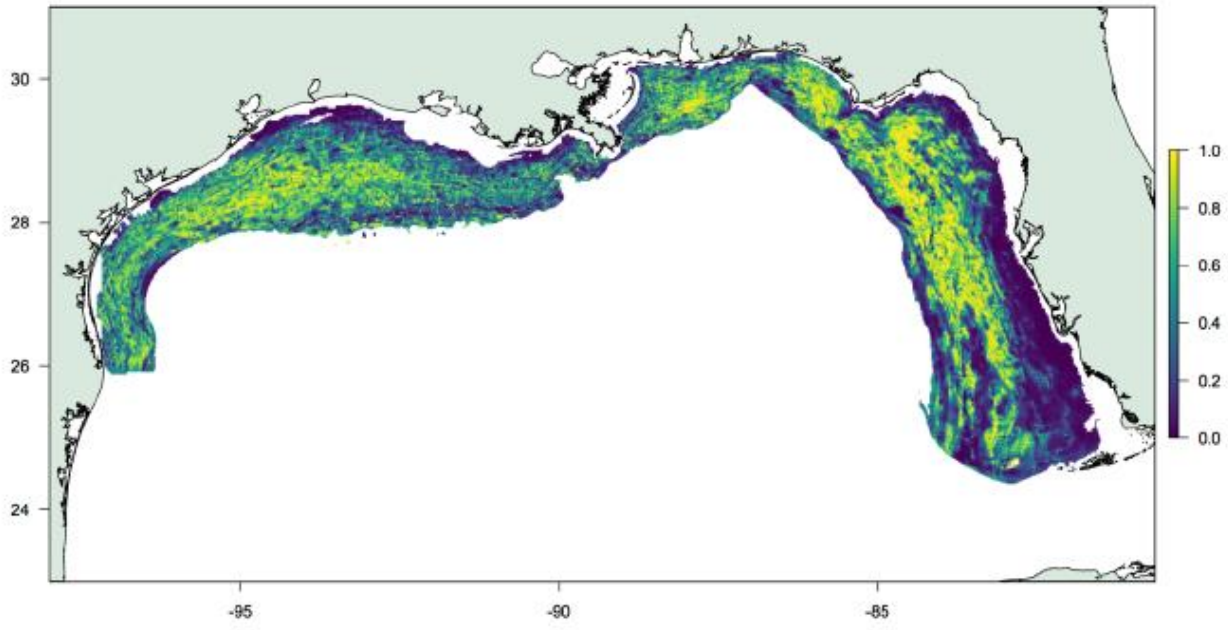
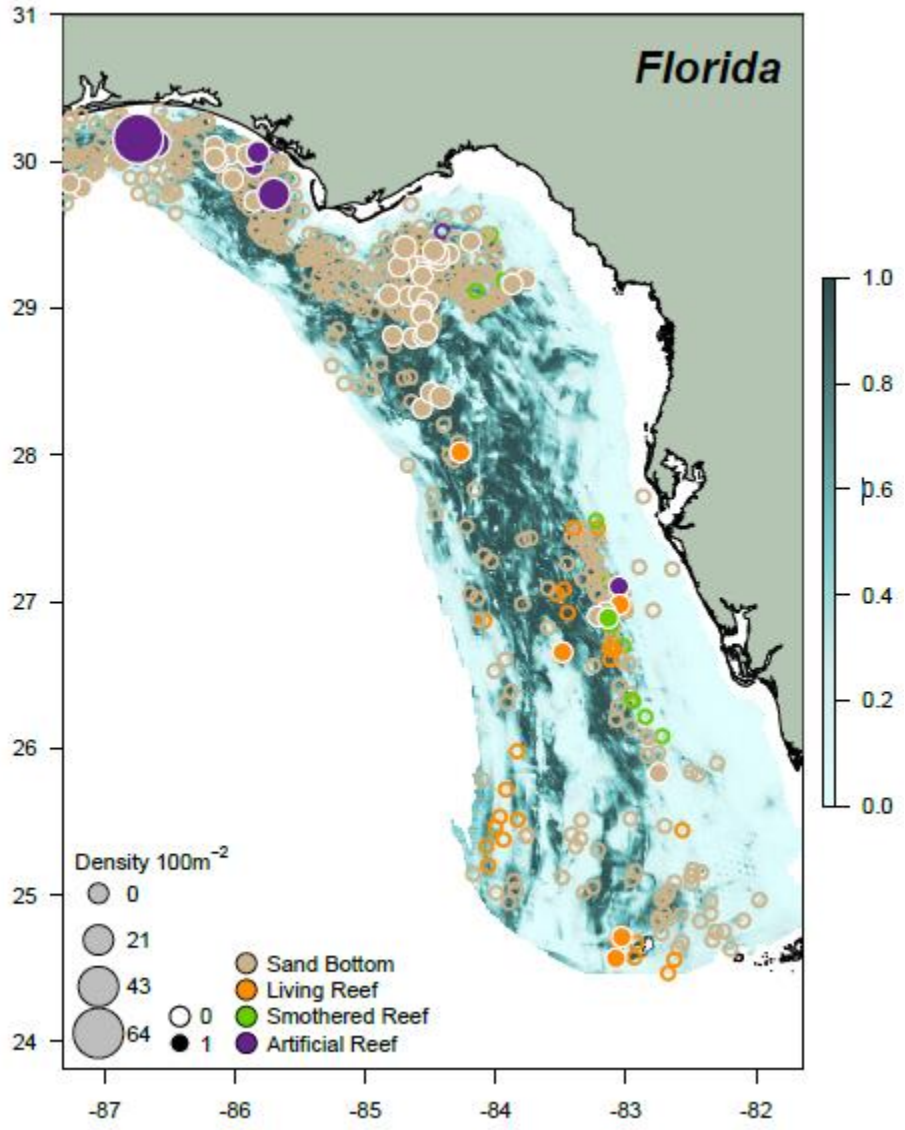


FIGURE 13


Florida
31
30
29
28
27
26
25
24
-87
-86
-85
-84
-83
-82
1.0
0.8
0.6
0.4
0.2
0.0
Density 100m^{-2}
0
21
43
64
0
1
Sand Bottom
Living Reef
Smothered Reef
Artificial Reef