DEP AGREEMENT NO. CZ529

Evaluation of sample size to assess managed-area-level trends in oyster habitats
John Handley, PhD

Task 3: Apply updated model and workflow to managed areas with sufficient monitoring data and completed SEACAR oyster analyses;

Deliverable 3b: A report summarizing the results for each managed area that includes reason(s) for any parameter data that were deemed insufficient and summary points written for relevant staff to easily absorb the key takeaways that are most relevant to their work.

This report extends the work from Task 3, CZ325 from 2Q2023 to other managed areas and parameters. In the original work, the goal was to develop a methodology to determine how many spatial samples are needed for a certain precision when estimating a parameter, and from where to collect them. In that original work, we focused on shell height within Estero Bay Aquatic Preserve, while in this work we also included percent live and density. The basic approach was to use oyster shell height measurements from 36 reefs to build a spatial model for mean shell height and then use it to predict values for the remaining 549-36 = 513 reefs.

We then used a standard spatial sampling method called Generalized Random Tessellation Stratified (GRTS) to select a spatially representative subset of sites (reefs) (Stevens and Olsen, 2004; https://archive.epa.gov/nheerl/arm/web/pdf/grts_ss.pdf) and parameterized it to prioritize reefs with the greatest uncertainty in their mean shell height predictions for selection because less is known about them. This method is consistent with well-established efficient sampling survey methods, which prioritize collecting samples where the variance or uncertainty in the target parameter is deemed to be the greatest. This probability model-based procedure helps to prevent selection bias (i.e., selecting sites that might be skewed one direction or another) while planning for oyster monitoring, resulting in a more spatially representative and statistically informative sampling plan capable of producing better estimates of the parameter of interest.

In the work reported here, we extended that approach to a total of three parameters (shell height, percent live and density) and nine managed areas (with GTM NERR separated into seven sub-areas, as requested by GTM NERR staff; Table 1). In addition, the modeling results and GRTS workflow were integrated into interactive Shiny dashboard applications to facilitate interpretation of the results as well as their use in planning for future oyster monitoring work. The following sections describe the interactive dashboard's components and intended use, how the parameter uncertainties were combined to allow for selection of a single set of potential sample locations, and the sample size calculation procedures foundations and assumptions. Appendix A includes tables for each parameter, shell height (Table A1), percent live (Table A2) and density (Table A3), that summarize the analyses results for the 9 managed areas (with GTM NERR separated into seven sub-areas, as requested by GTM NERR staff). Each Appendix A results table illustrates sample selection results based on the produced modeling workflow and dashboard interface for given standard deviations for the relevant parameter to facilitate comparisons across managed areas. Finally, Appendix B summarizes the foundational process of computing appropriate sample sizes for simple random samples.

Table 1. Managed Areas in CZ529 study.

Office	Managed Area
Apalachicola National Estuarine Research Reserve	Apalachicola Bay Aquatic Preserve
Apalachicola National Estuarine Research Reserve	Apalachicola National Estuarine Research Reserve
Estero Bay Aquatic Preserve	Estero Bay Aquatic Preserve
GTMNERR	Guana River
GTMNERR	Salt Run
GTMNERR	Tolomato River
GTMNERR	St Augustine
GTMNERR	Pellicer Flats
GTMNERR	Butler Beach
GTMNERR	Fort Matanzas
GTMNERR	Guana River Marsh Aquatic Preserve
Indian River Lagoon Aquatic Preserves	Indian River-Vero Beach to Ft. Pierce Aquatic Preserve
Indian River Lagoon Aquatic Preserves	Jensen Beach to Jupiter Inlet Aquatic Preserve*
Charlotte Harbor Aquatic Preserves	Lemon Bay Aquatic Preserve
Charlotte Harbor Aquatic Preserves	Pine Island Sound Aquatic Preserve

*Jensen Beach had only one sampled site and was not included in this work.

Interactive dashboard

We developed a software suite to facilitate use of the site selection procedure. First, there is an R script, GeneralWorkflow.R, to fit a model and predict on unsampled sites. The script also uses simulations to show how standard deviation decreases as more GRTS-selected sites are included. Second, we provide a geographic information system to view the managed area along with the simulation results to guide the number of sites to be selected and view them on a map along with metadata. This application is based on a Shiny interface to a Leaflet mapping application (Figure 1).

Figure 1. Screen shot of GIS application for site selection.

At the top is a scrollable "readme" file that explains the basic functionality of the system and how to use it. On the upper left is a number of layer selections. These include prior sampling locations (red), bathymetry and mapped oyster reefs compiled by the Florida Fish and Wildlife Conservation Commission (shown in Figure 1).

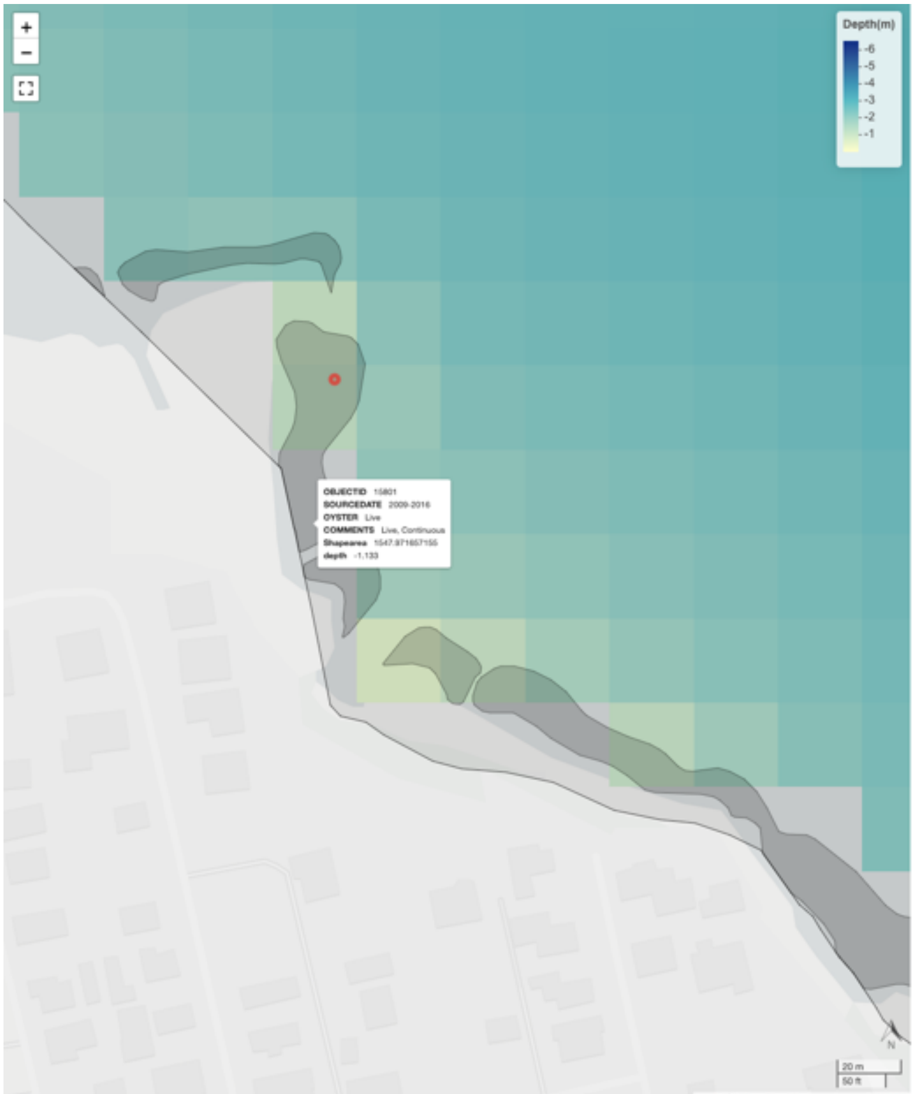

Figure 2. Blowup showing FWC oyster reef polygons in gray.

Recall that site selection is based on the GRTS algorithm but informed by estimation uncertainty from the spatial models for the three parameters. We can also select layers showing estimation uncertainty for shell height, percent live and density. The size of each point represents the relative uncertainty as seen in Figure 3 (larger circles represent greater uncertainty, and thus higher priority for GRTS selection).

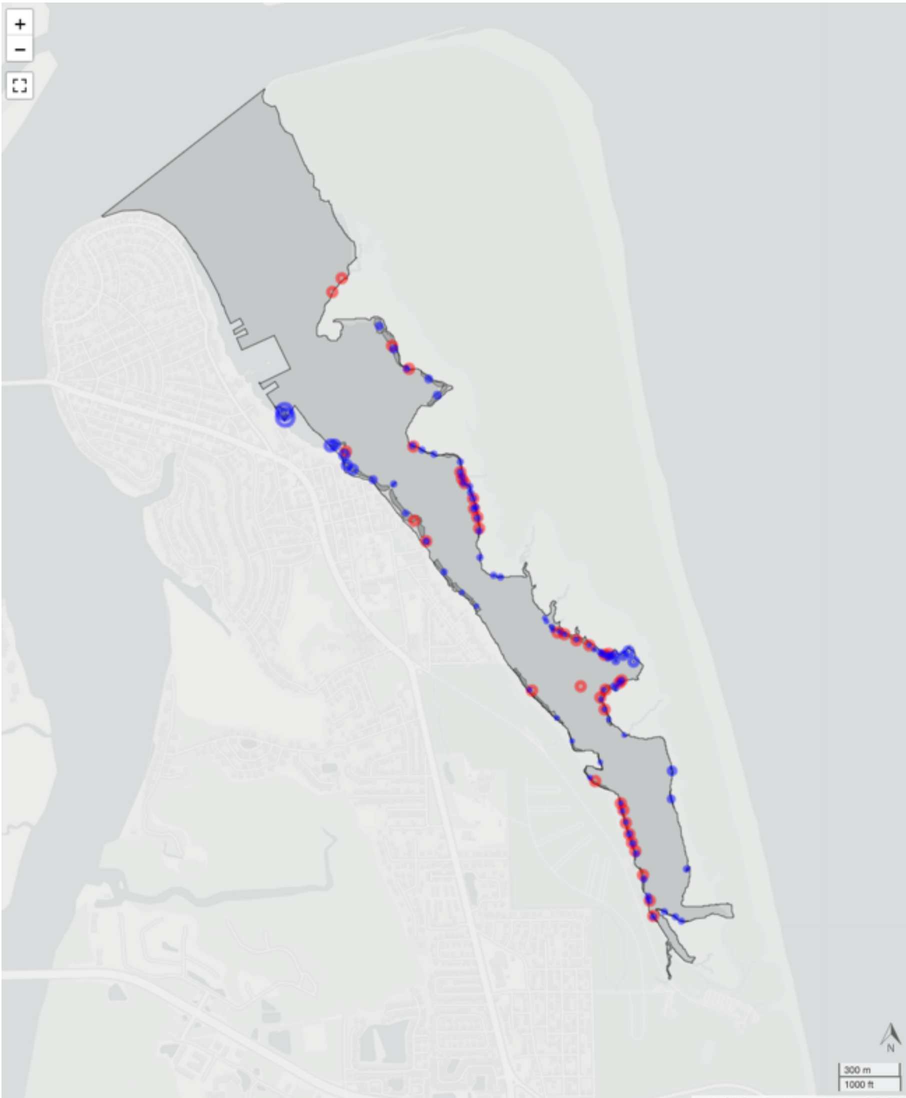

Figure 3. Previously sampled sites in red and estimation uncertainty for shell height in unsampled sites in blue. Larger circles indicate greater uncertainty.

Though sites with greater prediction uncertainty will be prioritized in GRTS site selection, intuitively, sites further away from existing sites naturally have greater uncertainty, but this also depends on the underlying data (number of samples at each site and the variability of the measurements).

One can also choose to include sites based on a range of water depths. The interface calculates the average depth for each FWC oyster reef polygon and presents the range for selection by the user. Similarly, the FWC oyster layer contains the areas for the reef polygons and the user can also select a range of areas to include in the candidate set of sites. One must be careful to decide a priori on the depth and area criteria rather than exploring the data to decide. Using data to decide selection criteria will violate statistical assumptions of randomness for subsequent analyses.

Finally, the user interface also allows one to deselect sites *a priori* by uploading a csv file —it is understood that some sites are not candidates because of ecological sensitivity or logistical considerations (e.g., accessibility). One can then select the number of sites of desired (guided by the plots) and then generate those sites in real time using the GRTS algorithm that is informed by modeling uncertainty.

We recommend that users choose more sites than the minimum they require because some sites after the fact may not be accessible. The important aspects are that the sites are chosen randomly and representative of the spatial distribution of candidate sites; if a site must be rejected the user should move to the next site on the list, in order, to preserve the randomness and spatial representativeness of the final sample collection.

The center part of the interface shows two columns of graphs for selecting the number of samples per site and the number of sites to sample (Figure 4). These issues are related but statistically complicated to deal with simultaneously.

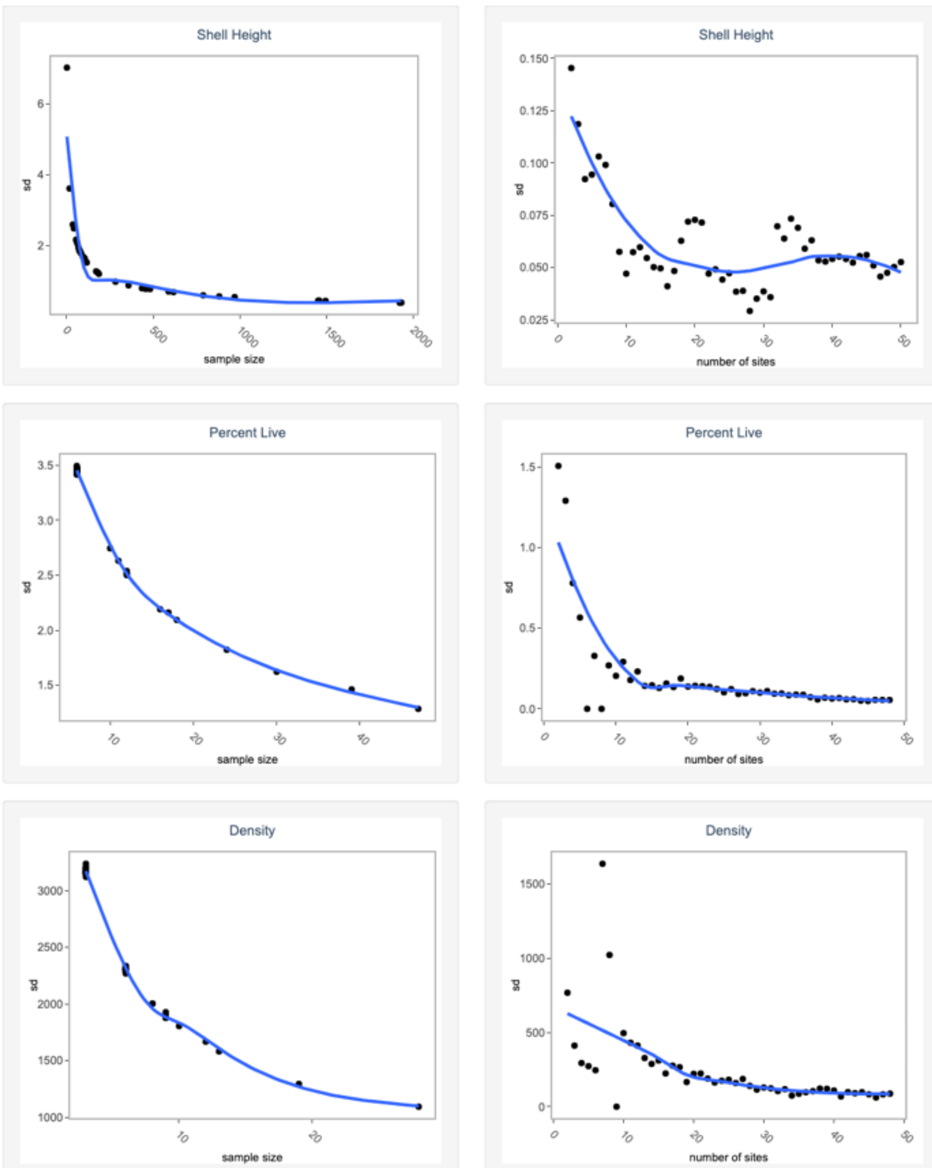

Figure 4. Plots of standard deviation by sample size and number of sites.

The number of samples per site is fairly straightforward. Within a reef, as the number of samples collected increases, the precision of the target parameter estimate, as measured by standard deviation, also increases. These graphs vary by parameter (shell height, percent live and density). Generally, there are far more shells collected to measure shell height than quadrats measured for percent live or density. The number of samples overall is determined by the desired precision.

The second column represents a simulation-based approach to estimate the relationship between area-wide uncertainty for the target parameter and the number of spatial samples collected. There is scant guidance in the literature on determining the number of spatial samples needed to attain an area-wide estimate for a target parameter with a given level of precision with confidence. The estimation is complicated by the fact that most of the possible spatial samples are unsampled and they vary across space, meaning the precision in the target parameter estimate one obtains from a given collection of samples depends on which set of locations are included in addition to the number of locations. The spatial variation, which is unknown, presents the greater challenge. Thus, the general recommendation is to attempt estimation by simulation. The graphs presented in the UI are based on fitting a model on the available existing data, and then repeatedly (50 times) invoking the enhanced GRTS algorithm for each of a sequence of numbers of sites, each time predicting the mean target parameter values for each site, and then averaging them. The standard deviations of those mean values are then proxies for the precision. The assumption of this approach is that the general shape of the resulting curve represents the diminishing standard deviation as a function of increasing the number of sample sites. We go into more detail at the end of this report.

After deciding on a number of sites to select, the user can enter it as the number of requested GRTS samples and the requested set of sites will be generated and displayed on the map (Figure 5) and can also be downloaded in a format of the user's choice.

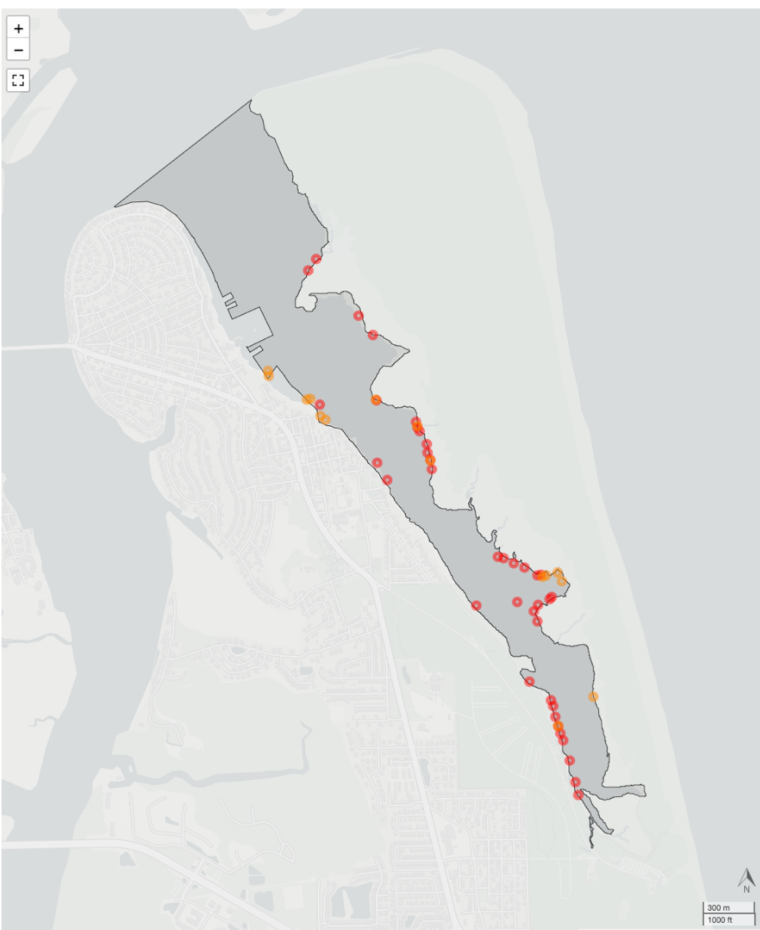

Figure 5. 15 GRTS generated sites for future sampling (orange points), overlayed on the set of sites with existing data (red points).

As a greater spatial extent is sampled, the advantages of including estimation uncertainty diminish. Likewise, if only a limited part of a managed area is sampled, the extrapolation to other parts of the area is unreliable, making the uncertainty across other possible sample sites uniformly high. For example, Figure 6 shows Pine Island Sound Aquatic Preserve, where all three sites with existing data are in the southern portion of the managed area while a majority of the unsampled locations are to the north. In this case, the uncertainty estimates for these unsampled locations are uniformly high, meaning any new samples should come from the north.

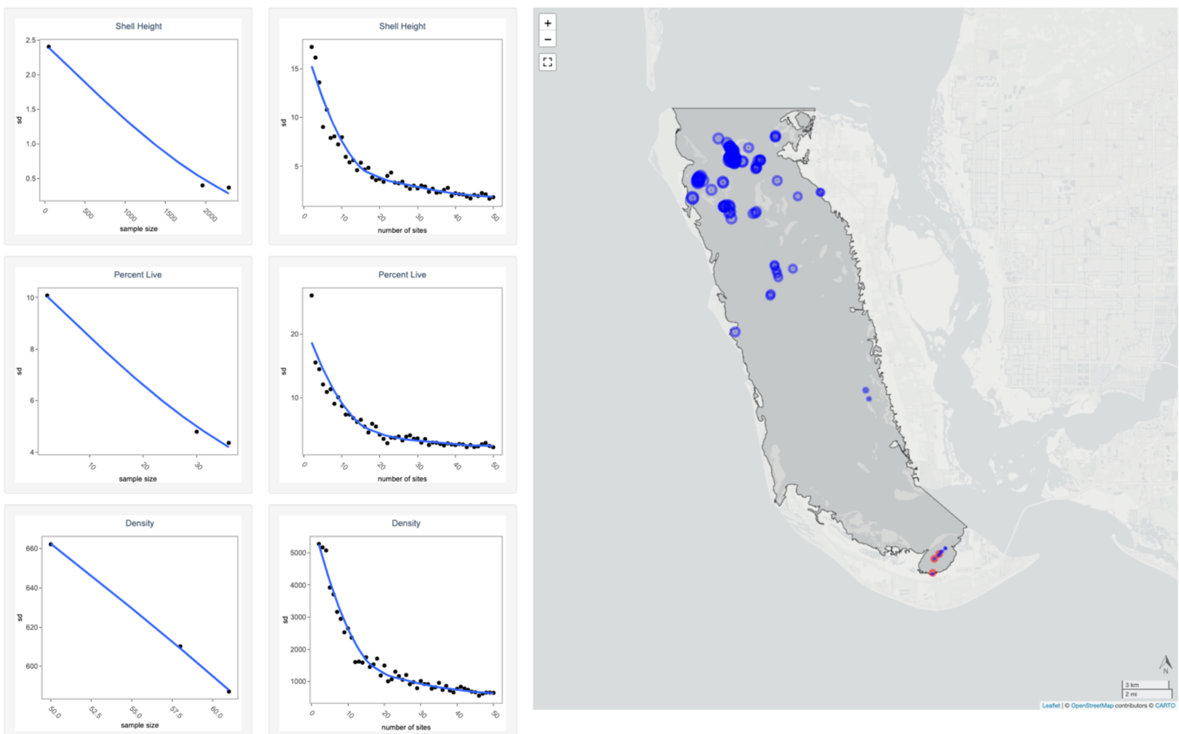

Figure 6. Pine Island Sound Aquatic Preserve with sample locations with existing data (red points) to the south and a majority of unsampled locations (blue points) to the north.

Nevertheless, the resulting samples are generated by GRTS and the uncertainty, being high across all the unsampled sites, will have little practical effect. One would simply get a GRTS sample.

Combining uncertainties for GRTS

We have up to three parameters to drive site selection (i.e., shell height, density, and percent live) in each managed area but we really only want one set of sites based on all three parameters. Because GRTS is influenced by modeling uncertainty, we wanted to test if the uncertainties for all three parameters would be approximately the same. It turned out that uncertainty estimates for the three parameters do tend to be positively correlated with each other, but there were a few managed areas where they were not well correlated (Table 2).

Table 2. Correlation of prediction uncertainties for all three parameters.

Managed Area	**Shell Height, Percent Live**	**Shell Height, Density**	**Percent Live, Density**
Estero Bay Aquatic Preserve	0.95	0.36	0.40
Guana River Marsh Aquatic Preserve	0.61	0.50	0.81
Apalachicola Bay Aquatic Preserve	0.11	0.97	0.04
Apalachicola NERR	0.27	0.45	0.17
Indian River-Vero Beach to Ft. Pierce Aquatic Preserve	0.98	0.98	0.98
Lemon Bay Aquatic Preserve	0.12	0.13	0.99
Pine Island Sound Aquatic Preserve	0.98	0.99	1.00
GTMNERR Guana River	0.50	0.37	0.81
GTMNERR Salt Run	0.62	0.82	0.83
GTMNERR Tolomato River	0.63	0.80	0.71
GTMNERR Pellicer Flats	0.96	0.76	0.73
GTMNERR St Augustine	1.00	0.97	0.97
GTMNERR Butler Beach	0.77	0.94	0.71
GTMNERR Fort Matanzas	0.93	0.93	0.98

For example, in Apalachicola Bay Aquatic Preserve, uncertainties for shell height were uncorrelated with those for percent live, and the uncertainty estimates for percent live were also uncorrelated with those for density (Table 2). Maps of these parameter uncertainties show that percent live tends to be most uncertain in the southern portion of the managed area, while both density and shell height were most uncertain in the western part of the bay (Figure 7). By using the average of the uncertainties when generating 20 GRTS samples, both areas were covered (Figure 7).

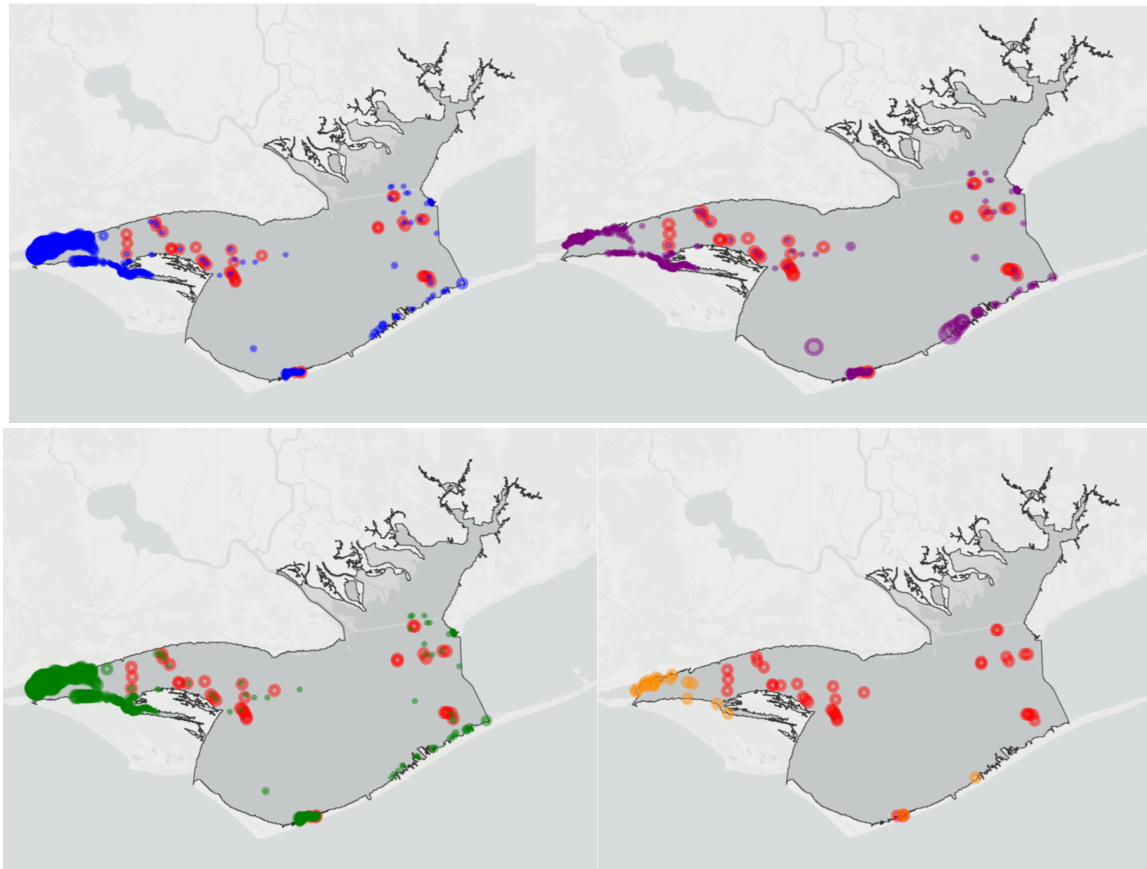

Figure 7. Apalachicola Bay Aquatic Preserve mean parameter uncertainties and the corresponding GRTS sample location selections based on the averaged uncertainties of the individual parameters. Upper left: Shell Height (blue points); Upper Right: Percent Live (purple points); Lower Left: Density (green points); Lower Right: GRTS samples (orange points). Locations with existing parameter data are shown as red points in each panel.

Another example is Lemon Bay Aquatic Preserve, where shell height uncertainty was poorly correlated with the estimated uncertainties for both percent live and density. Similar to the Apalachicola Bay Aquatic Preserve example, there was one region (in the northwest) where shell height uncertainties differed from those of percent live and density (Figure 8). Again, 20 GRTS generated samples based on the average uncertainties across the parameters captured the areas of highest uncertainty for all parameters (Figure 8).

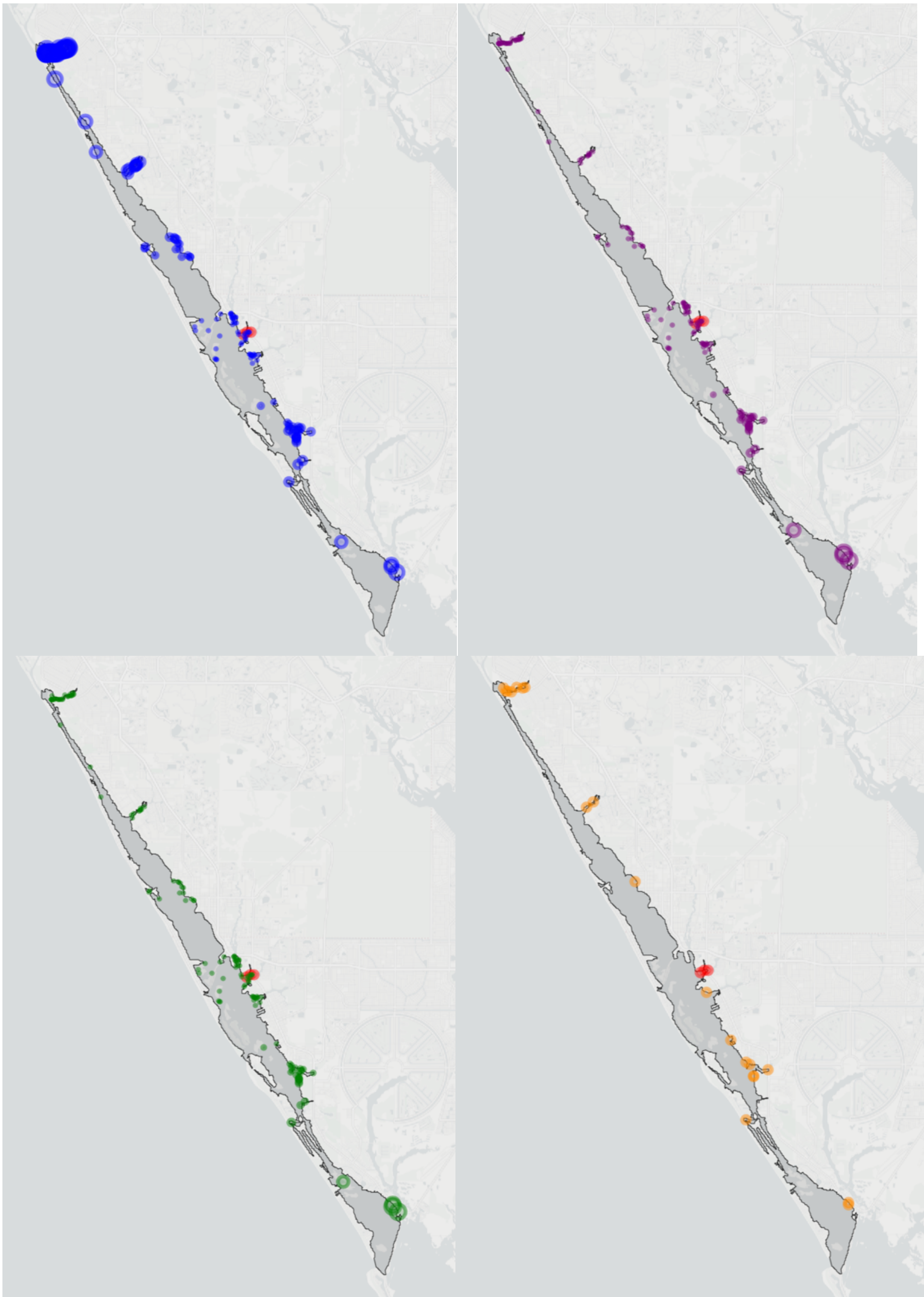

Figure 8. Lemon Bay Aquatic Preserve estimated uncertainties and the corresponding GRTS sample locations selected using the averaged uncertainties of the individual parameters. Upper left: Shell Height (blue points); Upper Right: Percent Live (purple points); Lower Left: Density (green points); Lower Right: GRTS samples (orange points). Locations with existing parameter data are shown as red points in each panel.

In summary, we used the average of the individual parameter uncertainties to drive the selection process in order to capture areas where the parameter uncertainties may differ in a single set of target sample site selections. In cases where the uncertainties differed, the spatial coverage aspect of the GRTS algorithm had the desired effect, as illustrated by the examples just discussed (Figures 7, 8). This approach is an "engineering" compromise between the added complexity of reconciling three different sets of "bespoke" spatial samples and the potentially reduced representativeness of a single set of spatial samples for any individual parameter by combining the three uncertainties into one. We believe the practical effect is minimal, given the uncertainties are often positively correlated (Table 2), and that the fundamental property of the GRTS algorithm to generate spatially representative samples provides the desired result.

Sample size calculation

For completeness, Appendix B contains a review of sample size calculation in the case of simple random sampling.

To review, sample size calculation depends on knowing (or estimating) properties of an underlying population (which are unknown because a sample has not yet been obtained). Further, there are assumptions made on the sampling process itself, like simple random sampling, where each item has the same probability of appearing in a sample. The upshot is that sample size calculations are approximate in practice. The statistical properties of the sampling process are rarely known in the sense that samples may not be truly random or biased in some unknown way.

Given a set of prior samples, as above, one can empirically approximate sampling effort by plotting the standard deviation of a value (say, shell height) against sample size. We show such a graph Figure 4.

In the context of spatial sampling, the situation is even more complicated because of the many additional unknown statistical properties. Often the task is to produce sample sites that are representative of the spatial arrangement of all the possible sites. It is much harder to estimate the number of additional sites owing to the spatial variability. We provide some guidance, however using the simulation procedure previously described.

Given a set of “legacy sites” we estimate a model to predict a response values. The Bayesian approach we use provides posterior distributions for model parameters. For unseen sites, we can estimate their mean values using a predictive algorithm using estimated parameter values and the coordinates of the new sites.

We use the Generalized Random Tessellation Stratified (GRTS) (Stevens and Olsen 2004) algorithm to select these sites. We also inform the algorithm to “prefer” sites with greater estimation uncertainty. The stratification heuristic is that one should sample more where the uncertainty is greater. The algorithm requires the number of sites m to be specified. For a given m, the algorithm selects m sites and we can estimate uncertainties of the estimated global mean from those m sites. We can take repeated samples of m from the GRTS algorithm and do the same thing. In our simulation, we do 50 for each value of m. For m from 2 to 50, we collect the standard deviations. We get an array of sample sizes from 2 to 50, each with 50 simulated standard deviations. If increased sample sizes were to increase precision, one would expect the means of the standard deviation distributions to decrease with sample size. This is a noisy process because repeated GRTS samples are all different and capture spatial variation, some of which was not incorporated in the original “legacy” data and model fit (i.e., for locations that have never been sampled, but do have predicted uncertainty values derived from the model fit).

As in the simple random sample case (see Appendix B), sample size calculation is subject to unknown uncertainties. This is more problematic in the case of spatial sampling because the spatial sampling itself is highly variable. That is, samples depend on *where* in addition to *how many*. But one could draw some basic conclusions about diminishing marginal returns – after some point, even in the presence of all that uncertainty, taking more samples simply won’t provide measurable improvements in the parameter estimates.

In practice, spatial sampling involves additional considerations beyond the simple statistical argument as well. Some sites may not be accessible to sample or there may be other known aspects about a site that make it an unsuitable candidate for a given monitoring program, which is why inclusion and exclusion facilities were included in the GIS application.

In summary, we aimed to provide a user-friendly application to facilitate a statistically principled approach to representative spatial sample site selection (i.e., determining how many sites to sample and an appropriate spatial arrangement). At the core, the application is a geographical information system (GIS) with basic mapping of oyster reefs, previously sampled sites and water depth, with site selection driven by the standard GRTS algorithm, enhanced by spatial uncertainty predictions derived from a kriging model fit to existing parameter data.

Reference

Stevens Jr, D. L. and A. R. Olsen (2004). Spatially balanced sampling of natural resources. Journal of the American Statistical Association, 99(465):262-278.

Appendix A. Summaries of sample size calculations

To provide some overall assessment of sample size calculation, we provide a table for each parameter. Using the interface for each of 14 managed areas and sub areas, we used the graphs to identify sample sizes needed to achieve a certain standard deviation. For some managed areas the standard deviation was already achieved according to the model while others would need samples beyond what was in the data and requires extrapolation outside the data range. The same is true for the number of reefs. For illustration, we chose standard deviations for which most of the curves provided values in the range of our simulations, 2 to 50. Readers are referred to the individual interactive dashboards for visualizations of the analyses results.

As a reminder, the standard deviation of a parameter as a function of the number of reefs is based on a simulation by repeatedly generating candidate sites using GRTS, extending the model to them and computing the standard deviation of the average of those estimated means. The goal is to show how precision increases with more GRTS samples. It cannot be directly used to assess what a managed area - level standard deviation would be, which would include sample sizes at each reef.

Some general conclusions are evident from the various managed area results. For example, when we extend far beyond the samples as in Pine Island Sound, the standard deviation estimates are not good, and in the inverse case, when there are few candidate samples left for the GRTS algorithm to choose, as in St. Augustine, there is not enough data to simulate a full sample size by standard deviation a curve. Further, in cases when a managed area is sparsely sampled to begin with, the modeling approach does not contribute much additional information, so in such cases the GRTS algorithm we provide in the interface essentially (and appropriately) defaults to the basic algorithm. Lastly, we note that bimodality in the raw data (e.g., as we found for in the percent live data for Apalachicola Bay, which tended to be either close to zero or very large and close to 100) poses problems for our model, which is based on normal-like distributions. A bimodal distribution is difficult to interpret; although it can be calculated, a simple summary statistic like a mean is basically meaningless for such distributions because one really has a mixture of two distributions that would be more appropriately summarized by two means. We do not recommend this procedure for these kinds of data, but if such cases arise, sample site selection can likely still be confidently guided by the other parameters, assuming their raw data are unimodally distributed. Fortunately, we encountered this issue for only one parameter in one managed area in our analyses for this project.

Overall, the analyses results highlight the significant spatial heterogeneity between oyster populations among different managed areas (and within managed areas, as illustrated by the GTM NERR sub-area results). Appropriate sampling plans to achieve a given level of precision depend on the variation in the existing data, the number of reefs in the managed area, and the size of the managed area, which can contribute to the spatial distance among sampled and unsampled reefs. Given this heterogeneity, we offer some managed-area-specific summary points here.

Apalachicola Bay AP and Apalachicola NERR

The highest parameter uncertainties in Apalachicola Bay AP and Apalachicola NERR were modeled in the far western part of the Bay, in the vicinity of Little St. George Island/Indian Pass and western St. Vincent Sound, depending on the parameter. In light of this result and due to the relatively high concentrations of mapped reefs, many of the GRTS-selected reefs tend to be in St. Vincent Sound. As mentioned previously in this report, the bimodality of the raw percent live data for these managed areas violated some assumptions of the spatial model and led to poor sample size simulation results for that parameter, with unstable standard deviation estimates, especially at lower simulated sample sizes. The shell height and density sample size simulation results appeared reasonable, however, and suggested that standard deviations of <3mm are possible for even 10 or fewer monitoring locations for shell height, and that standard deviations below 500 m^{-2} are possible with a minimum of about 10 monitoring locations, respectively.

Estero Bay AP

The highest parameter uncertainties in Estero Bay AP were modeled in the northwestern part of the Bay because most oyster monitoring data have been collected from the northeastern, central, and southern portions. Modeled uncertainties also tended to be locally higher near some freshwater inputs along the east/northeastern coast of the Bay. These results, combined with the numerous unsampled reefs in the northwestern portion of Estero Bay, led the GRTS procedure to select many of the target reefs there. Overall, the parameter sample size simulations for Estero Bay suggested that relatively low standard deviations may be possible even when sampling only a modest number of reefs, especially for shell height and percent live parameters, which each showed very low uncertainties even at low sample sizes, e.g., <1 mm and <2% standard deviations possible, respectively, with fewer than 10 reefs sampled. Density, unsurprisingly, was more variable but the simulated standard deviations still dropped below 300 m^{-2} after sample sizes reached about 15 reefs.

Guana Tolomato Matanzas NERR and Guana River Marsh AP

The GTM NERR analyses were split into seven sub-areas at the request of the managed area staff representatives (see Task 2 report; Table 1). The two northern-most GTM NERR sub-areas, Guana River and Tolomato River, have substantial overlap with the Guana River Marsh AP (GRMAP) boundary, meaning the models for these two GTM NERR sub-areas were essentially using subsets of the same dataset that was used for the GRMAP analyses. The spatial parameter uncertainties were correlated for all three areas, but the uncertainties in the Tolomato River sub-area tended to be greatest in its northern and southern ends, while the estimated uncertainties for the Guana River sub-area were highest in the northwest corner and were also elevated along central and southern stretches of the river's eastern bank. In the case of the GRMAP spatial model, similar patterns in estimated uncertainties are evident, but the largest ones again fall in the northern and southern ends of the Tolomato River portion of the managed area. The corresponding GRTS reef selections followed the expected pattern based on the spatial uncertainties for the two sub-areas individually. Interestingly, however, the majority of the GRTS selections for the GRMAP area fell within its Guana River portion despite the higher modeled uncertainties in parts of the Tolomato River area, likely due to the very large number of mapped, but unsampled, reefs in the Guana River. Unsurprisingly, the sample size simulation results for the three areas were fairly similar, especially for percent live, for which all three simulations estimated that standard deviations of <1.5% are possible with sample sizes of fewer than 10 reefs, suggesting there is low spatial variability in percent live across these areas. In contrast, the results for shell height and density both suggested that the Guana River sub-area has greater variability than the Tolomato River sub-area, with the Guana River simulations suggesting that 10 or more reefs should be sampled to achieve standard deviations of <2mm and <500m^{-2} for shell height and density, respectively, while comparable standard deviations from the Tolomato River simulations were <0.5mm and <200m^{-2}.

Continuing south, the next GTM NERR sub-area is St. Augustine. This sub-area had relatively few mapped reefs, so a significant proportion of them have existing data. Most of the previously monitored reefs are in the northern half of the sub-area though, so the model estimated uncertainties for all parameters were highest in the southern end of the sub-area. As was the case in other managed areas, this spatial pattern of reefs with existing data and the resulting spatial model uncertainties led the GRTS procedure to select many reefs from the southern end of the sub-area. Unfortunately, due to the small number of mapped reefs, the simulation approach to estimating the relationship between number of reefs sampled and the precision in parameter estimation performed poorly for all three parameters for this sub-area, limiting its utility for estimating the number of reefs to sample. However, the maximum simulated standard deviations for shell height, percent live, and density were approximately 0.45mm, 2%, and 1700 m^{-2}, indicating that variability in at least shell height and percent live appears to be low and possibly justifying a target sample size of fewer than 10 reefs across the area.

The Salt Run sub-area encompasses a tributary that stretches south/southeast from the northern portion of the St. Augustine sub-area. The spatial uncertainty estimates for Salt Run tended to be largest along the central portion of its western bank and at its southern end, the former of which is also an area where few of the mapped reefs had existing data. Although the GRTS procedure selected a number of reefs in these areas with elevated uncertainties, the full set represented reefs from both banks of Salt Run and along its entire length. The sample size simulations for Salt Run again suggested that monitoring data from a relatively modest sampling effort of about 10 reefs may be capable of estimating parameter means with standard deviations of only 0.5mm, 0.4%, and 500 m^{-2} for shell height, percent live, and density, respectively.

The next sub-area south from St. Augustine and Salt Run is Butler Beach. The pattern of modeled uncertainties for Butler Beach differed slightly between parameters, despite high overall correlations (Table 2). The largest uncertainty estimates for shell height occurred along the eastern bank of the river in the northeastern portion of the sub-area, while those for percent live were located along the western bank of the river in the south-central portion of the sub-area, and the greatest uncertainties for density were estimated for reefs along the eastern bank of the river in the northwestern portion of the sub-area. These disparate areas of greatest uncertainty did not unduly affect the GRTS procedure, however, likely because it utilized average uncertainties across the parameters and due to the parameter uncertainties' otherwise high correlations. The GRTS procedure's reef selections were spread across the entire sub-area, with only the area of highest uncertainty for percent live not represented. Similar to the results for a number of other sub-areas, the sample size simulations for Butler Beach suggested that monitoring data from approximately 10 reefs may be sufficient to estimate managed-area-level means for shell height and percent live with standard deviations of only about 1mm and 0.2%, respectively. The Butler Beach density simulations yielded a sample size of 13 to achieve standard deviations below 300 m^{-2}.

The next GTM NERR sub-area to the south of Bulter Beach is Fort Matanzas. Previously sampled reefs in this sub-area are split between the northern portion and the southern-most portion, with a substantial number of unsampled reefs in between. Owing to this pattern, the spatial model is essentially interpolating over the unsampled reefs in the center of the sub-area, leading to the greatest uncertainties in those reefs. As expected, the GRTS procedure selected many reefs in the center of the area, which corresponded to the areas of higher uncertainty and was spatially representative of the reefs. The sample size simulations for Fort Matanzas suggested a shell height standard deviation of 0.2 mm can be obtained with six reefs, a standard deviation of 0.5% for percent live can be obtained with 20 reefs, and a density standard deviation of 300 m^{-2} can be reached with six reefs.

The southernmost sub-area within GTM NERR is Pellicer Flats. The reefs in Pellicer Flats that have previously been monitored are distributed mainly throughout the middle of the area and provided good initial spatial coverage except for the northern portion and along the south/southwestern border of the area, where the estimated uncertainties for all three parameters were greatest. Owing to the spatial coverage at the center of the area, models for each parameter predict well in the center of the sub-area, but have greater uncertainty toward the north and south -- the GRTS procedure selects those sites. Simulations for shell height means show a standard deviation of 2 mm is achieved at 28 reefs. Thirteen reefs are needed to obtain a percent live standard deviation of 0.5%. A density standard deviation of 300 m^{-2} would require more than the maximum 50 reefs used in the simulation. All three curves from the simulations show the expected smooth decreases in standard deviation.

Indian River-Vero Beach to Ft. Pierce AP and Jensen Beach to Jupiter Inlet AP

This area has only six reefs that have previously been sampled, which limited the estimation of the spatial model. Four of the reefs with existing data are clustered in the southeast where there are many other unsampled reefs, leading to the greatest prediction uncertainty being estimated for reefs to the north. The GRTS procedure selected those reefs as well as more in the southeast, where the bulk of the unsampled reefs are. Sample size simulations for shell height suggested it would take close to 24 more reefs to achieve a managed-area-level estimate of the mean with a standard deviation of 0.2 mm. For percent live, the simulations indicated that a sample size of 19 reefs would be needed to achieve a mean estimate with a standard deviation of 0.5%. At a sample size of 24 reefs, the density simulations estimated that the managed-area-level mean would have a standard deviation of 300 m^{-2}.

Analyses could not be conducted for Jensen Beach to Jupiter Inlet AP because existing monitoring data are only available for one oyster reef from that managed area (Table 1).

Lemon Bay AP

This area has three previously sampled reefs all clustered together in the middle of the area on the northeast side. This spatial arrangement makes it challenging to estimate a model that can predict across the large, unsampled region. Thus, it is not surprising that model uncertainties for all parameters were greatest to the far northwest and southeast portions of the managed area. The GRTS procedure still selected a set of reefs that was well distributed throughout the area, however. Sample size simulations suggested that sampling about 12 reefs would be required to achieve a standard deviation of 0.2 mm for managed-area-level estimates of mean shell height, a sample size of 20 reefs could yield a standard deviation of 0.5% for percent live, and sampling only five reefs could achieve a density standard deviation of 300 m^{-2}.

Pine Island Sound AP

The three sampled reefs in this area are confined to a small region in the south, while the bulk of the reefs in the managed area are distributed to the north. This small number of previously sampled reefs and their unrepresentative spatial distribution made the spatial model unreliable, leading to high model uncertainty across the northern portion of the managed area where the bulk of the unsampled reefs are. The GRTS samples selected were as expected, including many reefs across the northern portion of the managed area, where the estimated uncertainties across all parameters were greatest and that were farthest away from sampled reefs. Because there are ample unsampled reefs, the sample size simulations of parameter standard deviations showed an expected decrease as a function of the number of reefs sampled. However, the simulation estimated that it would require 46 samples to achieve a standard deviation of 2 mm for shell height. For percent live, the simulations estimated that a standard deviation of about 10% would be achieved by sampling approximately 10 reefs, and that a comparable sample size (10 reefs) would lead to standard deviations of about 3000 m^{-2} for density. These sample size estimates should be interpreted with caution though, because the data are sparse and confined to one region, meaning the results are likely to change as data from additional reefs are added.

Table A1. Shell Height

Managed Area	**# Prior samples**	**Total Oyster Reefs**	**Sample size to achieve sd of 2**	**Number of GRTS samples to achieve sd of 0.2**
Estero Bay Aquatic Preserve	36	459	64	*2
Guana River Aquatic Preserve	93	1692	119	38
Apalachicola Bay Aquatic Preserve	38	433	46	>50
Apalachicola National Estuarine Research Reserve	50	724	45	50
Indian River-Vero Beach to Ft. Pierce Aquatic Preserve	6	38	*190	24
Lemon Bay Aquatic Preserve	3	130	*658	12
Pine Island Sound Aquatic Preserve	3	104	390	>50
GTMNERR Guana River	50	1273	137	47
GTMNERR Salt Run	38	86	146	20
GTMNERR Tolomato River	46	437	91	23
GTMNERR St Augustine	6	15	136	**
GTMNERR Pellicer Flats	21	1026	*190	28
GTMNERR Butler Beach	35	1121	92	18
GTMNERR Fort Matanzas	41	725	104	6

* Standard deviation already achieved, showing minimum sample size from data
** Curve poorly estimated – does not decrease

Table A2. Percent Live

Managed Area	# Prior samples	Total Oyster Reefs	Sample size to achieve sd of 3	Number of GRTS sample to achieve sd of 0.5
Estero Bay Aquatic Preserve	23	459	33	11
Guana River Aquatic Preserve	93	1692	8	13
Apalachicola Bay Aquatic Preserve	49	433	95	*
Apalachicola National Estuarine Research Reserve	65	724	95	*
Indian River-Vero Beach to Ft. Pierce Aquatic Preserve	6	38	**15	19
Lemon Bay Aquatic Preserve	10	130	46	20
Pine Island Sound Aquatic Preserve	3	104	**35	>50
GTMNERR Guana River	50	1273	8	10
GTMNERR Salt Run	38	86	9	8
GTMNERR Tolomato River	46	437	8	12
GTMNERR St Augustine	6	15	**5	***
GTMNERR Pellicer Flats	22	1026	10	13
GTMNERR Butler Beach	35	1121	7	2
GTMNERR Fort Matanzas	41	725	8	20

* Percent live distribution is bimodal; spatial model fit is poor; distribution mean is meaningless for a bimodal distribution
** Standard deviation already achieved, showing minimum sample size from data
*** Curve poorly estimated – does not decrease

Table A3. Density

Managed Area	# Prior samples	Total Oyster Reefs	Samples to achieve sd of 3000	Number of GRTS sample to achieve sd of 300
Estero Bay Aquatic Preserve	34	459	10	15
Guana River Aquatic Preserve	93	1692	6	11
Apalachicola Bay Aquatic Preserve	35	433	*43	*2
Apalachicola National Estuarine Research Reserve	44	724	*4	11
Indian River-Vero Beach to Ft. Pierce Aquatic Preserve	5	38	*8	5
Lemon Bay Aquatic Preserve	7	130	6	24
Pine Island Sound Aquatic Preserve	3	104	*50	>50
GTMNERR Guana River	50	1273	7	11
GTMNERR Salt Run	38	86	4	16
GTMNERR Tolomato River	46	437	4	*2
GTMNERR St Augustine	6	15	7	**
GTMNERR Pellicer Flats	20	1026	>12	>50
GTMNERR Butler Beach	35	1121	3	13
GTMNERR Fort Matanzas	41	725	5	6

* Standard deviation already achieved, showing minimum sample size from data
** Curve poorly estimated – does not decrease

Appendix B. Sample size calculation for simple random samples
Simple random sampling, where each item has the same probability of appearing in a sample, is challenging in practice, but it is a useful assumption for illustrating the basic principles of sample size calculations. These calculations often begin by specifying a desired precision between the average population value $\bar{Y}$ and an average estimate $\bar{y}$,

$$P(|\bar{y} - \bar{Y}| \leq e) = 1 - \alpha$$

e is called the margin of error, and a typical target value is 0.03. A typical target value of α, the probability of a Type 1 error (or "significance" level), is 0.05. One would express this in words as a 95% confidence that the margin of error is no greater than 3%. How big of a sample does one need to achieve this approximate precision?

Using a normal approximation, a little algebra shows,

$$e = z_{\alpha/2} \sqrt[2]{\left(1 - \frac{n}{N}\right)} \frac{S}{\sqrt[2]{n}}$$

where z is the critical value corresponding with the target value of a, n is the sample size, N is the size of the population being sampled, and S is the standard deviation of the value of interest. This equation shows that precision is affected by the sampling proportion n/N but also by the population standard deviation. To get a sample size for given e and α (and assuming S), solve for n,

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$

where

$$n_0 = z_{\alpha/2}^2 S^2/e^2$$

To achieve the margin of error e, assuming N is very large, the required sample size is then

$$n_0 = z_{\alpha/2}^2 S^2/e^2$$

In general, S is unknown, but some assumptions can be made. For a political simple survey to determine the proportion p of the population preferring one candidate over another, one can use the maximum variance value of a Bernoulli distribution when $p=1/2$: $S^2 = 0.25 = p(1-p)$ and $z_{0.025} = 1.96$; an upper bound on the sample size is

$$n_0 = 1.96 * 1.96 * 0.25/(0.03 * 0.03)$$

or approximately 1,067 for e = 0.03.