Design of a Port Sampling Program for the U.S. Caribbean

(Report to the National Marine Fisheries Service Southeast Fisheries Science Center)

Todd Gedamke, John Hoenig, Lisa Ailloud and Kristen Omori

MER Consultants

December 9, 2014

Executive Summary

As the first step of this project we conducted ten days of site visits in Puerto Rico, St. Croix and St. Thomas. We met with fisheries staff, fishers, and fisheries supervisors on each island to better understand the subtleties of the fisheries in each region and begin developing a strategy for a sampling program. On all three islands we found that the list of landing locations in the NMFS database did not specifically represent landing sites (i.e. boat docks or ramps) but could apply to more general areas (e.g. a city name such as Frederiksted in St. Croix). For example, in Puerto Rico the largest reported landings come from the city of Puerto Real but this single reported landing location contains four separate places where fish could be landed. After further discussion with territorial staff we also found that sites had multiple names and others were no longer in existence.

Our first priority became revising the sampling frame (i.e. locations to be sampled). We first asked port agents and fisheries supervisors on each island to review the list of locations and provide any additional site names, clarify which names referred to single landing location and which sites could be excluded from a commercial landing survey. Each respondent then scored a spreadsheet of location names in terms of relative usage, number of boats, types of fishing, and the number of samplers necessary to see all landings. Given the responses and modifications to the sampling frame it became clear that calculating variances from the existing data may, or may not be appropriate to develop an efficient, long term, cost-effective sampling design. As a result, our primary recommendation is to conduct a pilot study in which learning about the fishery (ranking sampling sites, estimating components of variance) is an important goal. The same design can be implemented for an ongoing production survey with only minor changes in magnitude and allocation of sampling effort.

A sampling design was created for estimating the total catch weight (landings) by species for each of six regions; software was written and implemented to generate a specific sampling plan given the sampling design and specific information on number of port samplers, number of days in the survey, and definitions of sampling strata. The samplable unit is the day x sampling location combination. The sampling frame is the list of all day x location combinations. The samplable units are divided into 12 strata based on geographic location and whether the location

is a “high” or “low” usage area (i.e., 6 regions x 2 usage levels = 12 strata). Within a stratum, we considered three sampling options: 1) simple random sampling, 2) further stratification by day (thus, each region x usage level x day is a stratum, implying each day is sampled), or 3) each day is considered a cluster of locations (within a usage level x region stratum) and a random sample of clusters is selected at the first stage and one or more locations are selected within each cluster (day) at the second stage.

The final recommendation is option 3 – two stage cluster sampling within strata. The variable recorded on each samplable unit is the total weight of fish (by species) landed during the day at the site. At the same time, the number of trips in the samplable unit is recorded (to get catch per unit effort) and, if possible, observations on the number of fish landed, the type and amount of gear deployed, and the time of day when the fish were landed are also recorded in order to learn more about the fishery. The estimated total landings by species can be compared to that obtained from the commercial catch records (CCR). Additionally, the estimated total number of trips can be compared to that obtained from the CCR. It is also possible to use the CCR observations, aggregated over a suitable unit of time, as an auxiliary variable. The hope is that the CCR observations are correlated with the observations from the port sampling program. In this case, the CCR observations could be used in a ratio estimator to improve the estimates from the port sampling survey. This is true even if the CCR observations are biased. Hence, the port sampling and CCR observations can be compared in three ways (total catch, total effort, correlation).

Goals and Objectives

The goal of the project is to develop a sampling design that can be used to estimate the landings in each of four regions in Puerto Rico and in each of the two island groups (St. Croix and St. Thomas/St. John) in the US Virgin Islands with the purpose of validating or evaluating the landings obtained from commercial catch records (CCR). A secondary goal is to determine additional information that can be obtained from the same sampling program, specifically fishing effort and information that can be used to improve the sampling in the future (e.g., patterns of effort and landings over time and over space).

The objectives are to:

- Develop a sampling frame (complete list of all sampling units in the population of interest)
- Use on-site visits, interviews with knowledgeable persons, and commercial catch records to rank the sampling units in terms of landings
- Specify explicitly the sampling design
- Examine existing data from commercial catch records to obtain possible insights into sampling variances
- Write and implement a routine in R to generate a specific sampling plan given the design parameters.

Description of the Sampling Design: Structure of the Sampling

The samplable unit is the day x location combination, e.g., Fred's Dock on Wednesday, the 12th of November. It is assumed that all landings in the samplable unit can be observed by one observer. The sampling frame is the list of all known locations (landing sites) replicated for each day of the sampling program (e.g., 2 months, excluding Sundays, for an anticipated pilot project; reasons for excluding Sundays are given below under Other Considerations). For each sampled unit, the variables recorded are the landed catch in pounds of each species. (At the same time, other variables could be recorded: number of trips, amount and type of fishing gear used, amount of effort, number of fish of each species landed, and time when each trip was recorded.)

The samplable units are grouped into strata according to the desired geographical specificity of the estimates, logistics, and consideration of precision. Precision is enhanced when the strata are internally homogeneous which implies the stratum means are variable.

A constraint is that estimates are desired for 6 separate geographic regions (North, East, South and West Puerto Rico, St. Croix, and St. Thomas/St. John). The efforts to construct the sampling frame and rank the anticipated landings across sampling locations revealed a great deal of uncertainty concerning relative usage of locations. Therefore, we chose to group the locations within each region as being either "high" or "low" usage, with a few nominal sites being designated as "exclude" (don't sample) because interviewees suggested there are trivial or no landings at these sites or because some names in the NMFS database pertain to vague areas or are redundant. This led to the definition of 6 x 2 = 12 sampling strata, each stratum being defined by a set of sampling locations with the samplable unit being a day at a location. The locations defining the strata are given in Tables 1 and 2; a summary is as follows:

Stratum	Number of Locations
1) PR-North, high usage	4
2) PR-North, low usage	13
3) PR-East, high usage	8
4) PR-East, low usage	9
5) PR-South, high usage	8
6) PR-South, low usage	12
7) PR-West, high usage	11
8) PR-West, low usage	11
9) St. Croix, high usage	3
10) St. Croix, low usage	7
11) St. Thomas/St. John, high usage	7
12) St. Thomas/St. John, low usage	9

We note that the assessments of relative usage of locations by fisheries personnel conflict in some cases with the information in the CCR database. Hence, the assignments of locations to high and low use strata may not match the information in the CCR database because we chose to rely more on the fisheries professionals (c.f. Tables 1,2, and Appendix A). In addition, locations were excluded from the study due to our interviews during site visits although they may be represented within the CCR database.

Sampling within strata. Within a stratum, we considered three options: 1) simple random sampling of samplable units, 2) further stratification by day such that each region x usage level x day combination is a stratum, implying each day is sampled in each region x usage level combination, and 3) each day is considered a cluster of locations (within a usage level) and a random sample of clusters is selected at the first stage and one or more locations are selected within each cluster (day) at the second stage.

1) Simple random sampling within strata. Under this design, within each of the 12 strata a simple random sample of the samplable units (day x location combination) is selected. This design could lead to sampling plans in which some days have no locations selected and other days have multiple numbers of locations selected (Figure 1). This can cause logistical problems because some days may require more port samplers than are available. For this reason, this option was rejected.

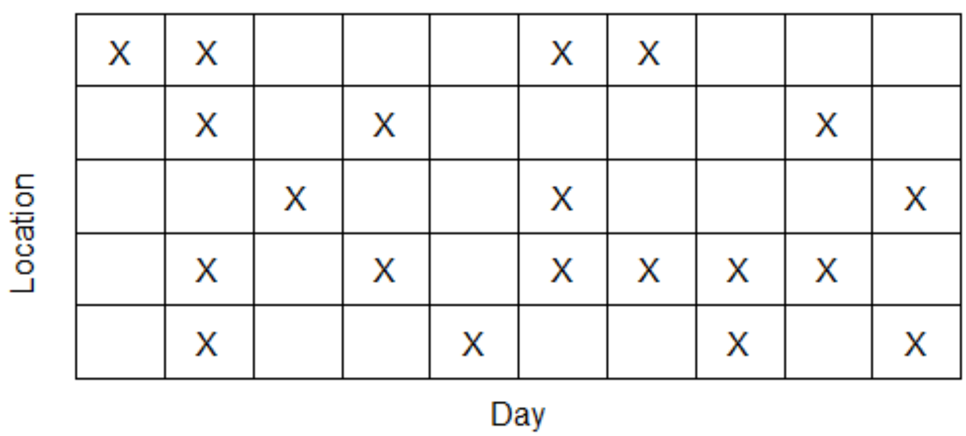


Figure 1. A simple random sample of 20 day x location combinations. Note that the number of port samplers needed varies by day.

2) Stratification by day. Under this design, a survey lasting 60 days would have 6 regions x 2 usage levels x 60 days = 720 strata. Each stratum would have a fixed number of locations, from which as many locations as there are port samplers available to the stratum would be selected for sampling (see Figure 2 for the case where there are two port samplers working each day, hence each stratum yields two observations). For example, if each port sampler worked 6 days per week, this would imply that 12 port samplers would be needed to have one person sampling each stratum, and one would obtain one observation per stratum. It is not possible to compute an unbiased estimate of the variance when only one observation is taken per stratum. However, one can use the method of collapsed strata to obtain a positively biased estimate of the variance (Sukatme et al. 1984). Under this approach, adjacent strata (days) are paired and the variance is estimated from the two observations in the pair of strata. If the stratum means in adjacent days are very close this causes little bias but, to the extent that the means differ by day, this shows up as increased estimated variance. Note that this option may cause some logistical (manpower) problems associated with requiring coverage 6 days per week.

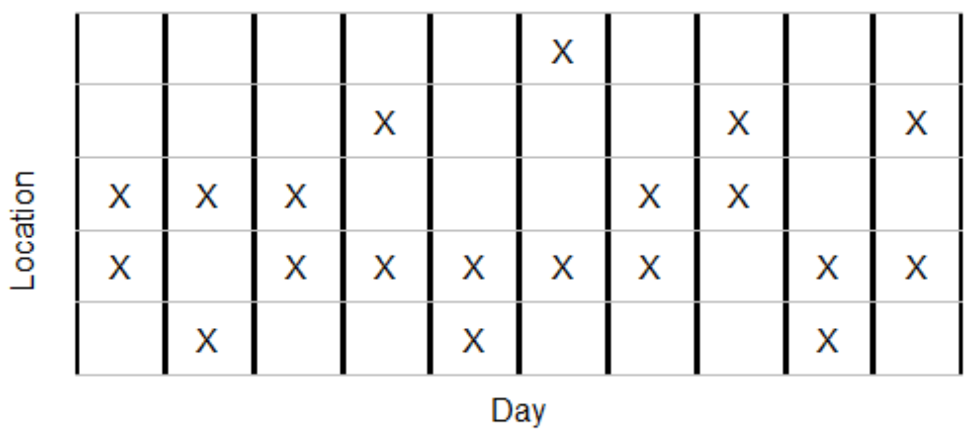


Figure 2. A stratified random sample where the days are treated as strata and 2 locations are selected for each day (stratum). The stratum boundaries are denoted by heavy lines.

3) Two-stage cluster sampling. Under this option, within each of the 12 strata (defined by region and usage level) the day is considered a cluster of sampling locations. Thus, at the first stage, a simple random sample of days is selected without replacement and, at the second stage, one or more locations are selected without replacement; the number of locations selected is fixed

(Figure 3). If there is only one port sampler available per selected cluster (i.e., there is only one port sampler available each day) the second stage variance component cannot be estimated in an unbiased way. However, it is still possible to obtain a variance estimate for the estimated total landings using the first stage variance; this will be a positively biased estimate of the variance (see Cochran 1977 p. 279).

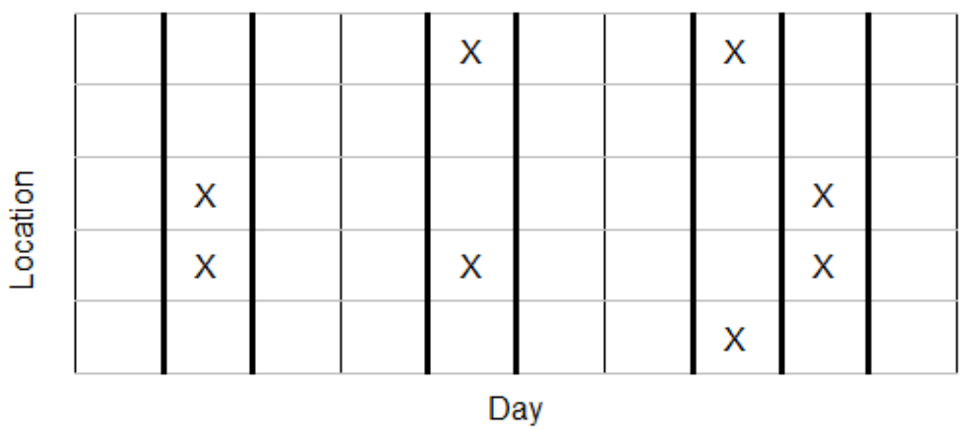


Figure 3. A two-stage cluster sample. At the first stage, 4 clusters (days) are selected. These are shown by the heavy lines. At the second stage, two locations are selected from each cluster (day).

The final recommendation is for two-stage cluster sampling because it is logistically convenient and, if two port samplers are available per stratum, it is possible to estimate the within day and between day components of the variance. We note that the sampling design may not be optimal in its allocation of two sampling agents to each stratum (it may be that some strata should have just one sampling agent and others have more than two). However, in a pilot study it is important to be able to estimate the components of variability and, with just one port sampler in a stratum, it is not possible to obtain unbiased estimates of the second stage (within-day) variance. The design advocated here can easily be modified to optimize allocation once reliable estimates of the variance components are available.

Other Considerations

Length of the fishing day and night time fishing activity. It is our understanding based on interviews with government fishery scientists and the commercial fishing industry that, apart from night landings of yellowtail and associated species, almost all landings occur within an eight hour day. Consequently, we focus on the day time landings and recognize that the design we propose will fail to quantify the landings of yellowtail. A separate program would have to be designed in order to estimate the landings from night fishing.

Excluding Sundays. We excluded Sundays from consideration based on input from government fisheries personnel. However, review of the CCR records reveals a variety of patterns of reported landings over the days of the week, with Sunday sometimes appearing as important and sometimes appearing unimportant (Figure 4). Hence, it will be necessary to verify (with an auxiliary study or spot checks during the pilot study) the assumption that Sundays are not important days for landing fishes.

Model-Based Sampling. In developing a sampling design for landings in the US Caribbean, we considered several reports prepared by Kaiser (2009a,b,c, 2010). These reports detail a model-based approach to estimating total landings by species using Bayesian statistics. The author acknowledges that the approach requires some strong assumptions and the results may not be robust to failure of assumptions. In our review, we found that the author was assuming simple random sampling of fish, i.e., that the fish sampled were representative of the population by virtue of the size of the sample, rather than by virtue of a sampling design. This approach provided no guidance on how the fish to be sampled should be obtained. Thus, the author does not distinguish between sampling all fish at one location at one time versus sampling fish at many locations or many times. We did not utilize this approach because of concerns about obtaining representative samples.

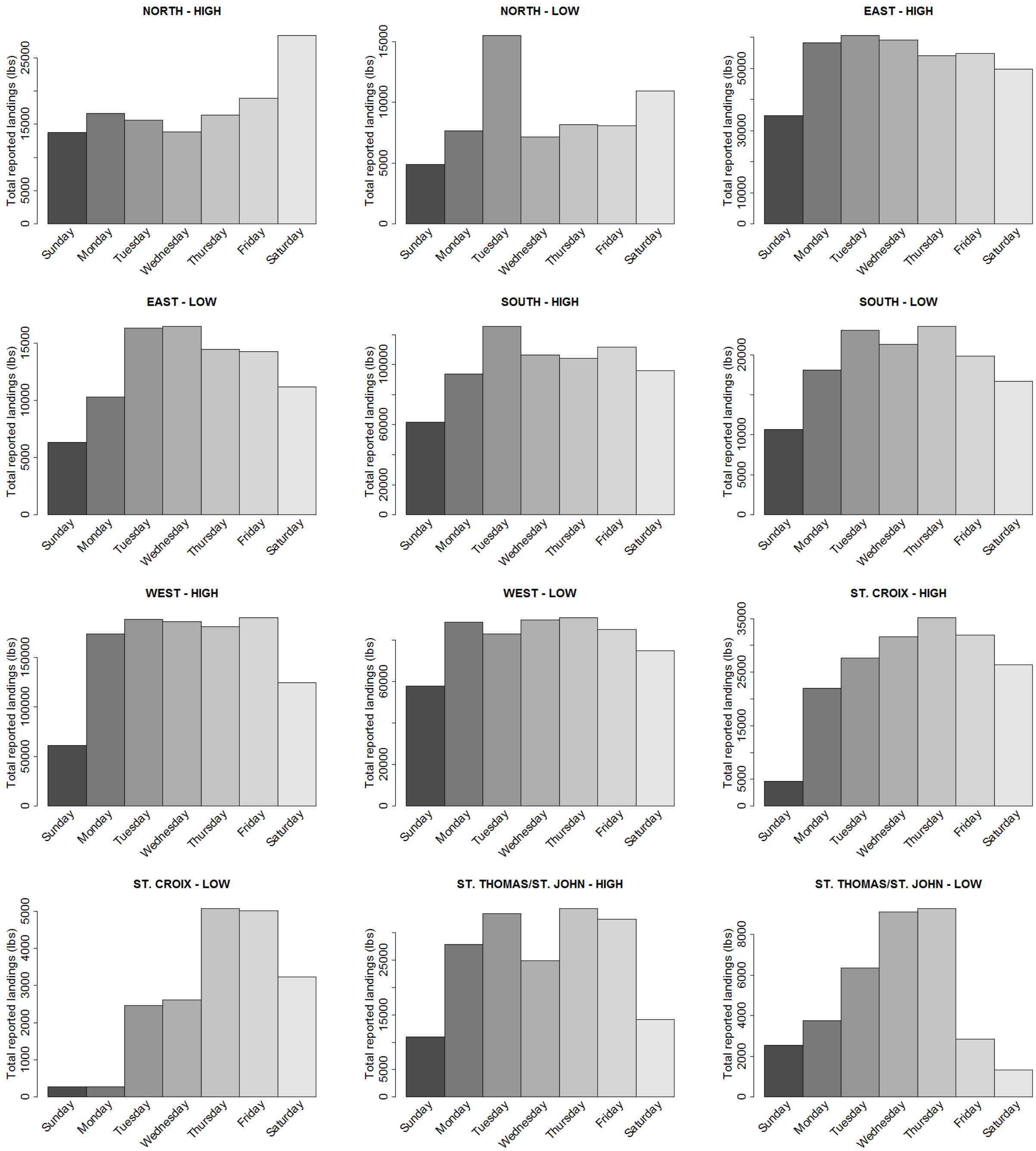


Figure 4. Pattern of daily total reported landings (lbs) over days of the week. Note the great variability in the patterns among the 12 strata. Sundays sometimes appear unimportant, especially in the Virgin Islands, and sometimes appear to have landings close to those of the other days.

Duration of the Pilot Project. We recommend a partial year pilot study with high intensity sampling rather than a full year study at low intensity, for example a two or three month long study. There are three reasons for this. First, if two port sampling agents are available per stratum, it is possible to obtain unbiased estimates of the components of variability (i.e., within cluster (locations) and among cluster (days) variability). With only one port agent per stratum, only a biased estimate of the variance of the total landings is possible for each stratum. Second, the goal is to compare CCR-based estimates with estimates from the designed port sampling program and, thus, the designed sampling program should be precise to facilitate the comparison. We believe a precise estimate over a two or three month period will allow one to determine if the CCR-based estimates differ systematically from the port sampling estimates over an appreciable period of time. In contrast, a full year survey with low sampling intensity may fail to detect systematic differences if the port sampling program produces highly variable results. Third, an important goal of the port sampling program is to provide further information that can be used to improve the sampling design in the future. We believe the design considered here, if two port samplers are available per stratum, will give much information on components of variance. Additionally, it will provide information for ranking sampling locations with time largely removed as a confounding factor.

Manpower Levels. Optimal allocation of port samplers to strata calls for the number of port samplers to be proportional to the product of the variance times the size of the stratum (number of locations within the stratum). It is possible to allocate a fractional person, e.g., 2.4 persons, to a stratum if one person works a fraction (e.g., 40%) of the time. We don't think this is advisable because part-time employees still need to be trained and they may be less likely to remain with the survey until the end if the compensation is minimal. In order to be able to compute unbiased estimates of the variance, a minimum of two port samplers is needed per stratum. However, in the event that funding is limited, it is possible to compute conservative (positively biased) estimates of the variance. We describe below how the data would be analyzed for designs employing one sampler and designs employing two samplers per stratum.

Application to Puerto Rico and the US Virgin Islands

Here, we outline in detail the design and analytical procedures if there are two and if there is just one port sampler(s) per stratum. We follow the notation of Cochran (1977, chapter 10). Consider a survey of 60 fishing days (i.e., 10 weeks of Monday through Saturday). To begin, we assume that there will be 50 of the 60 days selected at the first stage (because the port samplers will work on average 5 days a week).

Two port samplers per stratum. On each day selected, two locations will be selected for sampling. One port sampler will go to each of the two sites and remain there the whole time that fish may be landed (excluding night time landings). To illustrate the procedures, we consider the North Puerto Rico low use stratum. This gives rise to the following design parameters.

N = number of primary sampling units (PSU, days) = 60

n = number of primary sampling units sampled = 50

M = number of secondary sampling units (SSU, locations) = 15

m = number of secondary sampling units sampled from each PSU = 2.

The variable recorded, y_{ij}, is the total weight of fish (of the species of interest) landed on day i in the secondary sampling unit j (defined by location). Then, the mean of the two observations on day i is

$$\bar{y}_i = \sum_{j=1}^{m} \frac{y_{ij}}{m} = \sum_{j=1}^{2} \frac{y_{ij}}{2}, \quad (1)$$

and the mean of all the observations is

$$\bar{\bar{y}} = \sum_{i=1}^{n} \frac{\bar{y}_i}{n} = \sum_{i=1}^{50} \frac{\bar{y}_i}{50}. \quad (2)$$

We note that $\bar{\bar{y}}$ is an unbiased estimate of the population mean over all days and locations. We want an estimate of the total landings, and an unbiased estimate of this would be $NM\bar{\bar{y}} = 60(15)\bar{\bar{y}} = 900\bar{\bar{y}}$.

To estimate the variance of $\bar{\bar{y}}$, we define:

$f_1 = n/N$ = fraction of PSUs (days) sampled = 50/60 = 0.833, (3)

$f_2 = m/M$ = fraction of SSUs (locations) sampled in a day = 2/15 = 0.133, (4)

$s_1^2 = \frac{\sum_{i=1}^{n}(\bar{y}_i - \bar{\bar{y}})^2}{n-1}$ = sample variance among daily means, (5)

$s_2^2 = \frac{\sum_{i=1}^{n}\sum_{m=1}^{m}(y_{ij} - \bar{y}_i)^2}{n(m-1)}$ = sample variance among SSUs (locations) within PSUs (days). (6)

An unbiased estimate of the variance of $\bar{\bar{y}}$ can be obtained as

$$v(\bar{\bar{y}}) = \frac{1-f_1}{n} s_1^2 + \frac{f_1(1-f_2)}{mn} s_2^2 = \frac{1-.833}{50} s_1^2 + \frac{.833(1-.133)}{2*50} s_2^2 = .00334 s_1^2 + .00722 s_2^2 \ . \qquad (7)$$

Here, 1 - f_1 and 1 - f_2 are the finite population corrections (fpc) that reflect the fact that as you approach 100% sampling at a stage the variance for that stage has to approach 0. We note that if we apply this procedure to the North Puerto Rico high use stratum, the number of SSUs (locations) is only 4 so the fpc for the second stage has a much larger influence in reducing the variance, i.e., multiply by 1 – 2/4 = 0.5 instead of 1 – 2/15 = 0.867 for the low use stratum. The fpc for all other strata are between these extremes (except the St. Croix high use stratum for which the second stage fpc is 1 – 2/3 = 0.333).

We desire an estimate of the total landings over all locations and all days. As indicated above, this can be obtained as $\hat{Y}_{tot} = NM\,\bar{\bar{y}}$ = 60(15) $\bar{\bar{y}}$ = 900 $\bar{\bar{y}}$. The estimated variance of this is equal to the variance in (7) multiplied by $M^2N^2 = 15^2(60^2) = 810{,}000$. (The result may seem very large but, ultimately, it is the standard error of the estimate, not the variance, which we use to judge precision.)

The true variance of the estimated grand mean is given by (Cochran 1977 eq. 10.8):

$$V(\bar{\bar{y}}) = \left(\frac{N-n}{N}\right)\frac{S_1^2}{n} + \left(\frac{M-m}{M}\right)\frac{S_2^2}{mn} \ . \qquad (8)$$

Note that S_1^2 and S_2^2 are the true variances at the first and second stages, not estimates.

One port sampler per stratum. If funds are limited and there is only one port sampler available for each stratum, it is not possible to estimate the variance at the second stage (i.e., the variability among locations). However, Cochran (1977, pp. 278-279) shows that a positively biased estimate of the variance of $\bar{\bar{y}}$ can be obtained as (his equation 10.23)

$$v(\bar{\bar{y}}) = \frac{s_1^2}{n}. \tag{9}$$

When the number of days sampled is a small fraction of the total number of days, equation (9) is a good approximation. However, if n/N is not small, equation (9) overestimates by the amount f_1 S_1^2/n which, in the case considered here is (50/60) S_1^2 / 50 = 0.0167 S_1^2 . Here, S_1^2 is the true value of the variance at the first stage (among days), not an estimate.

Other estimators – using CCR information. At the same time that the port samplers are determining the weight of each species landed during a day at a given location, they can also get a count of the number of boats landing fish. The total number of trips in the stratum can be obtai ned as follows. Let

x_{ij} = number of trips observed on day i at location j.

Then,

$$T = \text{estimated total trips in the stratum } = NM\frac{\sum_{i=1}^{n}\sum_{j=1}^{m}x_{ij}}{nm} = NM\bar{\bar{x}}.$$

The sum of the two stratum totals in a region is an estimate of the total trips in the region. This can be compared directly to the total number of trips in the region in the CCR database. (The estimated variance of T can be obtained the same way the variance is estimated for the total landings in a stratum.) If the number of trips estimated from the port sampling is similar to the number of trips recorded in the CCR database then the total landings of a species could also be obtained by multiplying the average catch per trip estimated from the port sampling ($\bar{\bar{x}}$) times the number of trips determined by the CCR. In essence, this exercise seeks to determine if discrepancies between the port sampling and the CCR landings is due to problems with CCR catch per trip or CCR trip numbers or both.

It is also possible to use the landings data in the CCR in a ratio estimator of total landings. The trick is to aggregate the CCR landings data in a meaningful way. This is because the CCR program relies on fishers to report activities periodically and the dates may not be accurate even if the landings are accurate. Hence, by aggregating reports by a suitable time period (say, two weeks or 4 weeks) the effect of inaccurate dates can be minimized. The CCR data do not have to be unbiased; the key point is for the CCR landings to be correlated with the actual landings. Hence, CCR landings data are considered an index of actual landings. The form of the estimator, then, is

$$\text{estimated catch} = \frac{\sum_i estimated\ catch_i}{\sum_i CCR\ catch_i}\ Total\ CCR\ catch$$

where we have *n* estimates of catch from the port sampling and *n* corresponding reported catches in the CCR database. In essence, the ratio gives the calibration factor by which the total landings in the CCR database is multiplied. This could be very useful if the CCR records are found to correlate well with the survey data.

An R function called ThePlan() has been written to generate a sampling plan for a stratum. It prompts the user to enter the number of days in the survey, the number of sites in the stratum, and the number of port sampling agents available and returns the day x location combinations to be sampled. The R code and a sample output are included in the Appendix.

<u>Power analysis – effect of number of survey samplers</u>. It is possible to estimate the impact of changing the number of survey samplers if the design is held fixed and estimates are available of the components of variability. Thus, given the design proposed here, we need estimates s_1^2 and s_2^2. We attempted to obtain these estimates from the commercial catch records supplied to us by Steve Turner. There are three caveats that should be borne in mind. First, CCR records are unverified reports from commercial fishers and the reliability of the reports is open to question. Hence, it is not clear how useful they are for estimating the required variances. Second, the records do not pertain strictly to the sampling frame developed in this project. This is because some of the records are associated with non-specific locations such as “Christiansted” which is comprised of several sampling sites. Similarly, some sites, such as Puerto Real in western Puerto Rico, have been divided into two or more sampling units for the proposed design because the original site was too big to be observed in its entirety by one sampling agent. Third, some of the CCR records refer to species complexes such as “parrotfish” or “snapper” that would be identified to species in a future port sampling survey.

We obtained estimates $\bar{\bar{Y}}, s_1^2$ and s_2^2 for the 3 most commonly landed species (by weight) and for the 10th most commonly landed species for each stratum (Table 3). In ranking the species, records for large species groups (e.g., "parrotfish") were ignored (although they can comprise a large portion of the catch). Note that on some days, no landings were reported for a given location. This may mean that the landings for that day at that location were zero; it could also mean that the fishers were not careful to report reliably the days on which landings were made. In the latter case, the lack of landings for some days could be considered missing values (and the reported landings for the remaining days may (or may not) be inflated).

The percentage of day x location combinations with no landings reported for *any* species is shown in the column labelled %NA in Table 3 and ranges from 40.6% to 89.8%. This is quite troubling. Day x location combinations where there were no landings of the species of interest but there were landings of other species are considered to have zero catch of the species of interest and are treated as real zeros (as opposed to missing values). The percentage of the records with such zero catches is shown in the last column of Table 3 (labeled %ZeroCatch); values range from 4.5% to 47.3%. We do not know why the percentage of missing occurrences should be so high and find this quite troubling. Thus, we question the utility of the data for estimating the variances. Nonetheless, we have performed the power analysis.

The results for the species with the highest landed weight in each stratum are summarized in the table on the next page; results for second, third and tenth most commonly landed species are given in Table 4. The results for yellowtail snapper are not very meaningful because landings of this species cannot be surveyed with the sampling design proposed here. We present the standard errors as percentages of the landings for k = 1, 2, 3 and 4 sampling agents (samplers) per stratum. With only one sampling agent in a stratum, it is not possible to estimate the variance in an unbiased fashion from survey data; nonetheless, we can compute what the variance would be from knowledge of $S_1^2$and S_2^2.

The calculations for red hind in the north Puerto Rico high use stratum are as follows. From equation (8), with two samplers, four sampling locations, 50 out of 60 days sampled, and first and second stage variances (Table 3) of 1866.62 and 379.76, respectively, the variance of the mean over all locations and days would be

$$V(\bar{\bar{y}}) = \left(\frac{60-50}{60}\right)\frac{1866.62}{50} + \left(\frac{4-2}{4}\right)\frac{379.76}{2(50)} = 8.12.$$

The total landings is 60 x 4 x 25.1 = 6024 and the standard error of the total is 60 x 4 x the square root of 8.12, which is 683.9. Hence, the standard error is roughly 11.4% of the estimated landings and an approximate 95% confidence interval, assuming normality, would be the estimate ± 2 x standard error. Note that the North Puerto Rico high use stratum is particularly small so that the finite population correction at the second stage has a great effect in reducing the variance. In contrast, the North Puerto Rico low use stratum is particularly large and, correspondingly, the finite population correction has less impact resulting in considerably higher standard errors.

	Std. error as % of estimate with # of samplers =			
Stratum/species	1	2	3	4
PR-N, hi/red hind	14	11	10	10
PR-N, lo/lobster	53	49	47	46
PR-E, hi/mutton	14	11	9	9
PR-E, lo/mutton	10	9	8	8
PR-S, hi/yellowtail	17	12	9	8
PR-S, lo/red hind	15	13	12	12
PR-W, hi/mutton	15	11	9	8
PR-W, lo/herrings	22	17	15	14
StX, hi/yellowtail	8	5	4	3
StX, lo/yellowtail	6	6	6	5
StT/StJ, hi/yellowtail	13	11	10	9
StT/StJ lo/yellowtail	7	7	7	7

From the results above and in Table 4, it appears the standard errors can be anywhere from very small to very large. In some cases, when the second stage variance appears to be high, adding more port samplers reduces the variance considerably (e.g., herrings in the western part of Puerto Rico); in other cases, where the second stage variance is low (e.g., yellowtail in St. Thomas/St. John), adding more port samplers appears to do little to improve precision. This is based on the dubious conclusion that there is little variation among the locations in the stratum and the real source of uncertainty is the day to day variability in landings. We do not feel the CCR data give us much insight into sampling precision. It would be valuable to redo the computations once results are available from a pilot study and better estimates of the components of variability are available.

Literature Cited

Cochran, W.G. 1977. Sampling Techniques, third edition. John Wiley

Kaiser, M.S. 2009a. Sample size for estimating total trips in Puerto Rico. Dated December, 2009, with no identification of author or affiliation.

Kaiser, M.S. 2009b. Initial thoughts on sampling for USVI. Dated November, 2009, with no identification of author or affiliation.

Kaiser, M.S. 2009c. Assessing observation error in sampling for number of trips in Puerto Rico. Dated December, 2009, with no identification of author or affiliation.

Kaiser, M.S. 2010. Development of an estimator of catch in the U.S. Virgin Islands and Puerto Rico. Report prepared for Marine Resources Assessment Group Americas. March, 2010.

Sukhatme, P.C., B.V. Sukatme, S. Sukatme, and C. Asok. 1984. Sampling Theory of Surveys with Applications, third edition. Iowa State University Press.

Table 1. List of nominal sampling locations, final strata based on an evaluation of relative usage (high, low, or exclude), and number of samplers necessary to observe all landings in each location in Puerto Rico. Site or location names that are in all capital letters are identical to those in the NOAA database. Locations with two or more names will be treated as a single landing site.

NMFS Site Numbers	Location Name	NMFS "region"	Final Strata	Final Number of Samplers
230	EL FARO (MAUNABO)	EAST	HIGH	1
250	ESPERANZA	EAST	HIGH	1
201	HUCARES	EAST	HIGH	2
180	LAS CROABAS	EAST	HIGH	1
190	MACHOS BARRIO	EAST	HIGH	1
183	PLAYA PUERTO REAL	EAST	HIGH	1
182	PUERTO REAL;PUERTO DEL REY MARINA	EAST	HIGH	1
212	BUENA VISTA	EAST	LOW	1
240	CULEBRA MUNICIPIO;CULEBRA BARRIO-PUEBLO	EAST	LOW	1
200	EL CORCHO (DAGUAO)	EAST	LOW	1
251	MORROPO(ISLA DE VIEQUES)	EAST	LOW	1
220	PLAYA DE GUAYANES	EAST	LOW	1
211	PUNTA CANDELERO	EAST	LOW	1
191	PUNTA FIGUERAS	EAST	LOW	1
210	PUNTA SANTIAGO	EAST	LOW	1
181	SARDINERA BARRIO	EAST	LOW	1
221	PLAYA LUCIA(YABUCOA)	EAST	EXCLUDE	0
120	CANCHA LA PUNTILLA	NORTH	HIGH	1

50	JAREALITO	NORTH	HIGH	1
132	LA COAL(SAN JUAN)	NORTH	HIGH	1
80	PUERTO NUEVO	NORTH	HIGH	1
10	BAJURA BARRIO	NORTH	LOW	1
90	CERRO GORDO	NORTH	LOW	1
170	FORTUNA	NORTH	LOW	1
60/61/63	LA BOCA; PUNTA MANATI; PALMAS ALTAS BARRIO	NORTH	LOW	1
133	LAS MARGARITAS	NORTH	LOW	1
100	MAMEYAL	NORTH	LOW	1
160	PALMER	NORTH	LOW	1
130/131	PARADA 9 1/2(SAN JUAN); HOARE SUBBARRIO	NORTH	LOW	1
161	PUERTO MOSQUITO	NORTH	LOW	1
40	PUNTA MARACAYO	NORTH	LOW	1
30	PUNTA PENON	NORTH	LOW	1
140	TORRECILLA BAJA BARRIO	NORTH	LOW	1
150	VIEQUES BARRIO	NORTH	LOW	1
121	VIETNAM	NORTH	LOW	1
151	ANCONES (LOIZA)	NORTH	EXCLUDE	0
62	BARCELONETA BARRIO-PUEBLO	NORTH	EXCLUDE	0
41	HATILLO BARRIO-PUEBLO	NORTH	EXCLUDE	0
11	JOBOS BARRIO	NORTH	EXCLUDE	0
70	LA CALIFORNIA	NORTH	EXCLUDE	0
153	MEDIANIA BAJA BARRIO	NORTH	EXCLUDE	0
110	PALO SECO	NORTH	EXCLUDE	0

152	SUAREZ	NORTH	EXCLUDE	0
350	BAHIA	SOUTH	HIGH	1
280	JOBOS	SOUTH	HIGH	1
310	PASTILLO	SOUTH	HIGH	2
320	PLAYA DE PONCE	SOUTH	HIGH	1
290	PLAYA DE SALINAS (LAS OCHENTA)	SOUTH	HIGH	1
351/354/636	SALINAS PROVIDENCIA;PLAYA SANTA;CALETA SALINAS	SOUTH	HIGH	1
330	TALLABOA	SOUTH	HIGH	1
292	AGUIRRE BARRIO	SOUTH	LOW	1
340/646	BAHIA GUAYANILLA;PLAYA DE GUAYANILLA	SOUTH	LOW	1
260	BAJO BARRIO	SOUTH	LOW	1
281	BARRANCAS	SOUTH	LOW	1
301	PLAYA CORTADA	SOUTH	LOW	1
300	PLAYA DE SANTA ISABEL	SOUTH	LOW	1
270	PLAYA LAS PALMAS	SOUTH	LOW	1
361	PUNTA PAPAYO	SOUTH	LOW	1
360	PUNTA PARGUERA	SOUTH	LOW	2
282	PUNTA POZUELO	SOUTH	LOW	1
362	SALINAS FORTUNA	SOUTH	LOW	1
353	BALNEARIO DE CANA GORDA	SOUTH	EXCLUDE	0
352	BARRIO GUAYPAO	SOUTH	EXCLUDE	0
341	FARO (GUAYANILLA)	SOUTH	EXCLUDE	0
261	GUARDARRAYA BARRIO	SOUTH	EXCLUDE	0

322	LA GUANCHA PASEO TABLADO	SOUTH	EXCLUDE	0
291	LAS MAREAS	SOUTH	EXCLUDE	0
311	MANZANILLO	SOUTH	EXCLUDE	0
302	PLAYA DE JAUCA	SOUTH	EXCLUDE	0
321	TUQUE	SOUTH	EXCLUDE	0
371	BAHIA SUCIA	WEST	HIGH	2
401	BARRERO BARRIO	WEST	HIGH	1
372	EL COMBATE	WEST	HIGH	1
380	EL SECO(MAYAGUEZ MUNICIPIO)	WEST	HIGH	1
410	ESPINAR BARRIO	WEST	HIGH	1
381	MANI	WEST	HIGH	1
422	PLAYUELA BEACH	WEST	HIGH	1
374	PUERTO REAL	WEST	HIGH	2
403	RINCON BARRIO-PUEBLO	WEST	HIGH	1
373	BOQUERON	WEST	LOW	1
376	GUANAJIBO BARRIO	WEST	LOW	1
377	GUANIQUILLA	WEST	LOW	1
411	GUANIQUILLA BARRIO	WEST	LOW	1
420	HIGUEY	WEST	LOW	1
375	JOYUDA	WEST	LOW	1
382	MARINA MERIDIONAL SUBBARRIO	WEST	LOW	1
400	PARCELAS ESTELA(RINCON MUNICIPIO)	WEST	LOW	1
370	PUNTA PITAHAYA	WEST	LOW	1

421/423	TAMARINDO; EL GATO NEGRO RESTAURANT(CABO ROJO); EL GATO NEGRO RESTAURANT	WEST	LOW	1
390	TRES HERMANOS	WEST	LOW	1
384	CANO BOQUILLA	WEST	EXCLUDE	0
402	CORCEGA	WEST	EXCLUDE	0
383	RASQUETA(MAYAGUEZ MUNICIPIO)	WEST	EXCLUDE	0

End of Table 1.

Table 2. List of nominal sampling locations, final strata based on an evaluation of relative usage (high, low, or exclude), and number of samplers necessary to observe all landings in each location in St. Croix and St. Thomas/ St. John. Site or location names that are in all capital letters are identical to those in the NOAA database. Names containing lowercase letters were assigned to combine multiple sites or use current local names and do not correspond directly to those in the NOAA database. Locations with two or more names will be treated as a single landing site.

St. Croix

Location Name	Final Strata	Final Number of Samplers
ALTONA LAGOON, ALTONA	HIGH	1
KRAUSE LAGOON - MOLASSES DOCK	HIGH	1
MARKET SQUARE-FREDERIKSTED FISH MARKET	HIGH	1
CHRISTIANSTED HARBOR, RICHMOND, ST. CROIX SEAPLANE	LOW	2
Duggan's Reef	LOW	1
ESTATE CASTLE NUGENT	LOW	1
GREAT POND	LOW	1
Salt River Boat Ramp	LOW	1
TAGUE POINT-SKOV DOCK, TAGUE BAY	LOW	1
TURNER HOLE	LOW	1
CANE BAY	EXCLUDE	0
CHRISTIANSTED	EXCLUDE	0
Cramer's Park, Smuggler's Cove	EXCLUDE	0
EAST END CENSUS SUBDISTRICT	EXCLUDE	0
East End Marine Park	EXCLUDE	0
FREDERIKSTED	EXCLUDE	0
FREDERIKSTED PIER	EXCLUDE	0
SALT RIVER BEACH	EXCLUDE	0
SALT RIVER MARINA	EXCLUDE	0

St. Thomas/St. John

Location Name	Final Strata	Final Number of Samplers
COKI POINT	HIGH	1
FRENCHTOWN	HIGH	1
HULL BAY	HIGH	1
MANDAL	HIGH	1
MANGROVE LAGOON, Compass Point Marina, NADIR, NAZARETH, BENNER	HIGH	2
RED HOOK CENSUS DESIGNATED PLACE	HIGH	1
CHARLOTTE AMALIE, SAINT THOMAS HARBOR	LOW	1
CORAL BAY	LOW	1
CROWN BAY	LOW	1
CRUZ BAY	LOW	1
KRUM BAY	LOW	1
MAGENS BAY	LOW	1
Marine Science Center	LOW	1
Sapphire	LOW	1
SMITH BAY	LOW	1
Bolongo Bay	EXCLUDE	0
EAST END CENSUS SUBDISTRICT	EXCLUDE	0
ESTATE FRYDENHOJ	EXCLUDE	0
SAINT JOHN	EXCLUDE	0
SAINT THOMAS	EXCLUDE	0
SECRET HARBOR BEACH	EXCLUDE	0
UNKNOWN	EXCLUDE	0
WATER BAY - ST. THOMAS	EXCLUDE	0
WEST END CENSUS SUBDISTRICT	EXCLUDE	0
Yacht Haven Grand	EXCLUDE	0

End of Table 2.

Table 3. Estimates of overall mean and first and second stage sampling variances for the top 3 species and the 10th most abundant species in the landings (by weight, excluding species assemblages such as "parrotfish"). %NA refers to the percentage of day x location combinations for which there were no reported landings of any species. %Zero refers to the percentage of day x location combinations for which there was no reported catch of the species of interest but there were reported landings of other species.

Highest Ranked Species

Stratum	Species	$\bar{\bar{y}}$	s_1^2	s_2^2	%NA	%Zero
PR - NORTH - HIGH	HIND, RED	25.1	1866.62	379.76	62.15	17.49
PR - NORTH - LOW	LOBSTER, CARIB. SPINY	25	37104.31	2773.26	89.79	4.53
PR - EAST - HIGH	SNAPPER, MUTTON	31.14	1608.16	780.33	64.81	20.21
PR - EAST - LOW	SNAPPER, MUTTON	18.55	643.81	73.27	81.84	10.98
PR - SOUTH - HIGH	SNAPPER, YELLOWTAIL	45.66	1334.47	3160.92	43.05	12.49
PR - SOUTH - LOW	HIND, RED	11.19	418.34	80.29	80.33	9.51
PR - WEST - HIGH	SNAPPER, MUTTON	49.01	1551.89	2824.01	48.69	25.8
PR - WEST - LOW	HERRINGS, SARDINELLA	59.81	14922.17	7209	71.08	20.56
ST. CROIX - HIGH	SNAPPER, YELLOWTAIL	66.83	2031.51	1728.11	40.6	5.92
ST. CROIX - LOW	SNAPPER, YELLOWTAIL	25.02	548.48	23.81	76.73	7.58
ST. THOMAS/ST. JOHN - HIGH	SNAPPER, YELLOWTAIL	53.33	6309.12	1617.69	69.66	16.62
ST. THOMAS/ST. JOHN - LOW	SNAPPER, YELLOWTAIL	40.36	2402.93	23.17	81.55	4.6

Second Highest Ranked Species

Stratum	Species	$\bar{\bar{y}}$	s_1^2	s_2^2	%NA	%Zero
PR - NORTH - HIGH	CONCH, QUEEN	14.78	1597.58	282.32	62.15	30.21
PR - NORTH - LOW	CONCH, QUEEN	9.82	967.12	34.65	89.79	8.74
PR - EAST - HIGH	HIND, RED	22.68	1219.03	820.34	64.81	21.2
PR - EAST - LOW	SNAPPER, YELLOWTAIL	14.93	322.49	61.12	81.84	6.59
PR - SOUTH - HIGH	TRIGGERFISH, QUEEN	25.81	404.76	1090.81	43.05	26.32
PR - SOUTH - LOW	TRIGGERFISH, QUEEN	9.46	101.68	22.63	80.33	8.32
PR - WEST - HIGH	SNAPPER, YELLOWTAIL	37.52	2035.94	5277.55	48.69	19.29
PR - WEST - LOW	SNAPPER, MUTTON	16.32	417.23	185.22	71.08	17.53
ST. CROIX - HIGH	DOLPHIN	30.76	497.61	401.17	40.6	11.36
ST. CROIX - LOW	SNAPPER, MUTTON	7.49	336.51	11.16	76.73	19.53
ST. THOMAS/ST. JOHN - HIGH	LONGSPINE SQUIRRELFISH	38.46	6759.69	2134.66	69.66	21.84
ST. THOMAS/ST. JOHN - LOW	RUNNER, BLUE	18.22	436.63	13.17	81.55	7

Third Highest Ranked Species

Stratum	Species	$\bar{\bar{y}}$	s_1^2	s_2^2	%NA	%Zero
PR - NORTH - HIGH	LOBSTER, CARIB. SPINY	14.26	855.69	194.84	62.15	24.84
PR - NORTH - LOW	HIND, RED	8.79	325.42	16.28	89.79	7.27
PR - EAST - HIGH	SNAPPER, YELLOWTAIL	21.16	953.77	376.81	64.81	15.16
PR - EAST - LOW	HIND, RED	9.11	401.7	65.08	81.84	11.95
PR - SOUTH - HIGH	SNAPPER, MUTTON	11.07	185.09	329.84	43.05	40.04
PR - SOUTH - LOW	CONCH, QUEEN	9.11	620.53	153.25	80.33	17.66
PR - WEST - HIGH	LOBSTER, CARIB. SPINY	24.56	1111.1	2481.45	48.69	32.81
PR - WEST - LOW	SNAPPER, YELLOWTAIL	15.92	1882.6	717.71	71.08	11.65
ST. CROIX - HIGH	SNAPPER, MUTTON	19.02	813.38	320.59	40.6	38.85
ST. CROIX - LOW	RUNNER, BLUE	4.26	40.42	1.69	76.73	13.74
ST. THOMAS/ST. JOHN - HIGH	HIND, RED	37.98	3949.67	1279.42	69.66	13.32
ST. THOMAS/ST. JOHN - LOW	WENCHMAN	10.38	133.69	5.26	81.55	7.5

Tenth Highest Ranked Species

Stratum	Species	$\bar{\bar{y}}$	s_1^2	s_2^2	%NA	%Zero
PR - NORTH - HIGH	SNAPPER, LANE	2.28	45.11	17.32	62.15	31.76
PR - NORTH - LOW	OCTOPUS	3.16	156.72	7.7	89.79	8.9
PR - EAST - HIGH	SNAPPER, SILK	2.79	71.92	37.43	64.81	31.94
PR - EAST - LOW	SNAPPER, QUEEN	2.03	30.88	1.74	81.84	13.95
PR - SOUTH - HIGH	CERO	6.21	55.66	150.17	43.05	42.49
PR - SOUTH - LOW	TUNA, SKIPJACK	4.75	34.24	10.34	80.33	11.05
PR - WEST - HIGH	PORGIES	7.39	244.5	650.37	48.69	47.34
PR - WEST - LOW	JACKS	2.29	113.43	82.21	71.08	28.14
ST. CROIX - HIGH	SHARKS, REQUIEM	6.72	40.91	23.57	40.6	23.24
ST. CROIX - LOW	PARROTFISH, STOPLIGHT	1.79	21.88	2.37	76.73	18.63
ST. THOMAS/ST. JOHN - HIGH	GRUNT, BLUESTRIPED	6.58	289.68	14.68	69.66	14.14
ST. THOMAS/ST. JOHN - LOW	GRUNT, BLUESTRIPED	4.8	49.33	1.88	81.55	11.56

End of Table 3.

Table 4. Standard errors as percentages of the landings for k = 1, 2, 3 and 4 sampling agents (samplers) per stratum. Values of $\bar{\bar{y}}$, s_1^2 and s_2^2 needed for the computations are given in Table 3. The cluster sizes (number of locations in each stratum) are given in the table in the main text.

Second Highest Ranked Species

	Std. error as % of estimate with # of samplers =			
Stratum/species	1	2	3	4
PR-N, hi/conch	21	18	16	16
PR-N, lo/conch	20	19	19	19
PR-E, hi/hind, red	19	14	12	11
PR-E, lo/yellowtail	10	8	8	7
PR-S, hi/queen trigger	18	12	9	8
PR-S, lo/ queen trigger	9	8	7	7
PR-W, hi/yellowtail	27	19	15	13
PR-W, lo/mutton	13	10	9	9
StX, hi/dolphin	9	6	4	3
StX, lo/mutton	15	15	14	14
StT/StJ, hi/long. Squir.	20	16	14	14
StT/StJ lo/blue runner	7	7	7	7

Third Highest Ranked Species

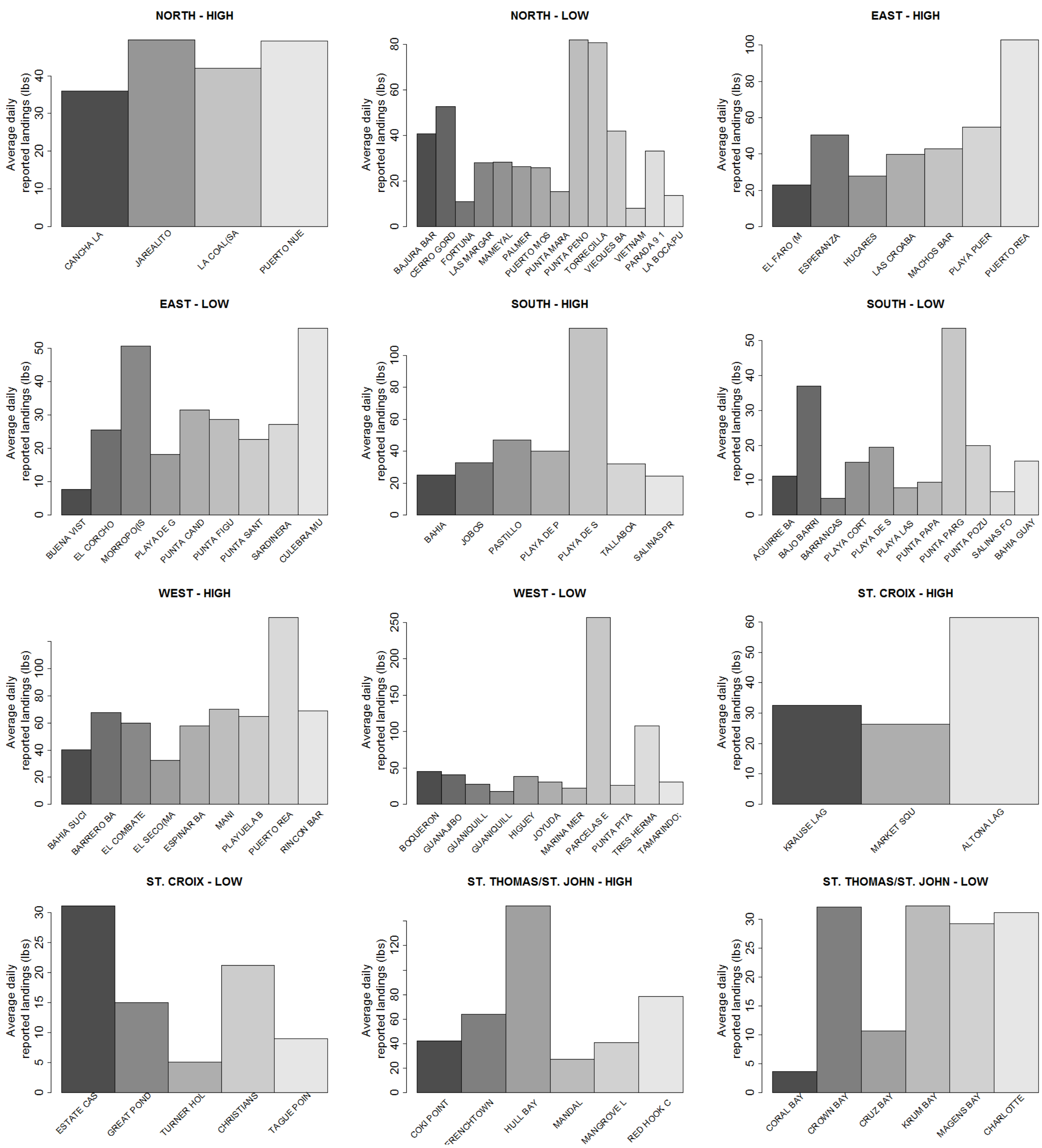
Stratum/species	Std. error as % of estimate with # of samplers = 1	2	3	4
PR-N, hi/spiny lobster	17	14	12	12
PR-N, lo/red hind	13	13	12	12
PR-E, hi/yellowtail	15	12	10	10
PR-E, lo/red hind	17	15	14	14
PR-S, hi/mutton	23	16	13	11
PR-S, lo/conch	24	20	18	18
PR-W, hi/spiny lobster	28	20	16	14
PR-W, lo/yellowtail	28	22	20	18
StX, hi/mutton	14	10	9	8
StX, lo/blue runner	9	9	9	9
StT/StJ, hi/red hind	16	12	11	11
StT/StJ lo/wenchman	7	7	7	7

Tenth Highest Ranked Species

Stratum/species	Std. error as % of estimate with # of samplers = 1	2	3	4
PR-N, hi/lane snapper	28	21	19	17
PR-N, lo/octopus	26	24	24	23
PR-E, hi/silk snapper	34	26	23	21
PR-E, lo/queen snapper	18	17	16	16
PR-S, hi/cero	27	18	15	12
PR-S, lo/skipjack	12	9	9	8
PR-W, hi/porgies	48	34	27	23
PR-W, lo/jacks	60	45	38	35
StX, hi/sharks, requiem	10	7	5	5
StX, lo/stoplight parrot.	19	17	16	16
StT/StJ, hi/bluestr. Grunt	17	16	15	15
StT/StJ lo/bluestr. grunt	9	9	9	9

End of Table 4.

Appendix A. Average daily reported landings by site for each of the 12 strata.



Appendix B. R function to generate a sampling plan. A portion of a sample output, covering the first 3 strata, is presented after the listing of the R code.

```
# Write a function to create sampling plans for a set of strata.
# User is prompted to enter:
#
# N      number of calendar weeks in the survey
# n      number of days to be sampled in the N weeks (e.g., in an 8 week survey we may sample
40 d)
# K      number of strata
# Size   vector of length K with the number of locations in each of the K strata
# P      number of port samplers assigned to each stratum

# Output is a list with K components; each component is a matrix of 0's and 1's designating
# which day x location combinations are to be sampled (1 = sample)

# It is assumed no sampling will take place on Sundays.

# J. Hoenig    11/10/14 for MER Consultants (todd@merconsultants.org)


ThePlan <- function() {
  ### Data input
  cat("enter number of calendar weeks in survey, then hit return")
  N = eval(parse())
  cat("enter number of days to be sampled (e.g., in an 8 week survey we may sample 40 d), then
hit return")
  n = eval(parse())
  cat("enter number of strata, then hit return")
  K = eval(parse())
```

```
Size <- rep(NA,K)
cat("enter vector with number of locations in each stratum, one stratum per line")
cat("\n", K, " lines of input needed")
for (i in 1:K)  Size[i] <- eval(parse())
cat("enter number of port samplers per stratum, then hit return")
P = eval(parse())
### Print the inputs to screen
cat("\nSummary of Inputs:\n")
cat("number of weeks in survey = ", N, " (", 6*N, "days, excluding Sundays)\n")
cat("number of days to be sampled = ", n, "\n")
cat("number of strata = ", K, "\n")
cat("number of locations in each stratum = ", Size, "\n")
cat("number of port samplers per stratum", P, "\n\n")

# generate samples
plan <- list()
days <- N*6
daynames <- rep(c("Mon", "Tues", "Wed", "Thurs", "Fri", "Sat"), N)
weeks <- rep(1:N, each=6)
for (i in 1:K) {  # for each stratum
  x <- matrix(0, days, Size[i])
  rnames <- NULL
  for(j in 1:days) rnames[j] <- paste("week",weeks[j],",",daynames[j])
  rownames(x) <- rnames
  plan[[i]] <- x
  names(plan)[[i]] <- paste("stratum",i)
  # now pick the n days out of days to be sampled
  pickeddays <- sample(1:days, n)
  # now pick P locations to sample for each day (row) sampled
  for (j in pickeddays) {
    pick <- sample(1:Size[i], P)
```

```
    plan[[i]][j,pick] <- 1
   }
  }
 return(plan)
}

########################################################################

ThePlan()

########################################################################
```

```
Summary of Inputs:
number of weeks in survey =  8  ( 48 days, excluding Sundays)
number of days to be sampled =  40
number of strata =  12
number of locations in each stratum =  4 13 8 9 8 12 11 11 3 7 7 9
number of port samplers per stratum 2

$`stratum 1`
               [,1] [,2] [,3] [,4]
week 1 , Mon      0    0    1    1
week 1 , Tues     0    1    0    1
week 1 , Wed      0    0    0    0
week 1 , Thurs    1    0    0    1
week 1 , Fri      1    1    0    0
week 1 , Sat      1    1    0    0
week 2 , Mon      1    0    0    1
week 2 , Tues     1    1    0    0
week 2 , Wed      0    1    1    0
week 2 , Thurs    1    0    0    1
week 2 , Fri      0    1    0    1
week 2 , Sat      1    0    0    1
week 3 , Mon      0    0    1    1
week 3 , Tues     1    0    0    1
week 3 , Wed      0    1    1    0
week 3 , Thurs    0    0    0    0
week 3 , Fri      0    0    0    0
week 3 , Sat      0    0    0    0
week 4 , Mon      0    1    0    1
week 4 , Tues     1    1    0    0
week 4 , Wed      1    1    0    0
week 4 , Thurs    1    0    1    0
week 4 , Fri      0    0    0    0
week 4 , Sat      1    1    0    0
week 5 , Mon      1    0    0    1
week 5 , Tues     1    0    0    1
week 5 , Wed      1    1    0    0
week 5 , Thurs    1    0    1    0
week 5 , Fri      0    0    0    0
week 5 , Sat      1    1    0    0
week 6 , Mon      0    0    0    0
week 6 , Tues     0    0    1    1
week 6 , Wed      0    1    0    1
week 6 , Thurs    1    0    1    0
```

```
week 6 , Fri        1    1    0    0
week 6 , Sat        0    0    0    0
week 7 , Mon        0    0    1    1
week 7 , Tues       0    0    1    1
week 7 , Wed        1    0    1    0
week 7 , Thurs      1    1    0    0
week 7 , Fri        0    0    1    1
week 7 , Sat        1    1    0    0
week 8 , Mon        0    1    1    0
week 8 , Tues       0    1    1    0
week 8 , Wed        0    0    1    1
week 8 , Thurs      1    0    1    0
week 8 , Fri        1    0    0    1
week 8 , Sat        1    0    0    1
```

```
$`stratum 2`
               [,1] [,2] [,3] [,4] [,5] [,6] [,7] [,8] [,9] [,10] [,11] [,12] [,13]
week 1 , Mon      0    0    0    0    0    0    0    0    0     0     0     0     0
week 1 , Tues     0    0    0    0    0    0    0    0    0     0     0     0     0
week 1 , Wed      0    0    0    0    1    1    0    0    0     0     0     0     0
week 1 , Thurs    0    0    1    0    1    0    0    0    0     0     0     0     0
week 1 , Fri      0    0    1    0    1    0    0    0    0     0     0     0     0
week 1 , Sat      0    0    0    0    0    0    0    1    0     0     0     1     0
week 2 , Mon      0    1    1    0    0    0    0    0    0     0     0     0     0
week 2 , Tues     0    0    0    0    0    1    0    0    1     0     0     0     0
week 2 , Wed      0    0    0    0    0    0    0    0    0     0     0     0     0
week 2 , Thurs    0    0    0    0    0    0    0    0    0     0     0     0     0
week 2 , Fri      0    0    0    0    0    1    1    0    0     0     0     0     0
week 2 , Sat      0    1    0    0    0    0    0    0    1     0     0     0     0
week 3 , Mon      1    0    0    1    0    0    0    0    0     0     0     0     0
week 3 , Tues     0    0    0    0    0    0    0    0    0     0     0     0     0
week 3 , Wed      1    0    0    0    0    0    0    0    1     0     0     0     0
week 3 , Thurs    0    0    0    0    1    0    0    1    0     0     0     0     0
week 3 , Fri      0    0    1    0    0    0    1    0    0     0     0     0     0
week 3 , Sat      0    0    0    0    0    1    0    0    0     1     0     0     0
week 4 , Mon      0    0    0    0    0    0    1    0    0     0     0     1     0
week 4 , Tues     0    0    0    0    0    0    0    0    0     1     0     0     1
week 4 , Wed      0    0    0    0    0    0    1    0    0     1     0     0     0
week 4 , Thurs    0    0    1    0    0    0    0    1    0     0     0     0     0
week 4 , Fri      0    0    0    0    0    0    1    0    1     0     0     0     0
week 4 , Sat      0    0    0    1    1    0    0    0    0     0     0     0     0
week 5 , Mon      0    0    0    0    0    0    0    0    0     1     0     1     0
week 5 , Tues     1    0    0    0    0    0    1    0    0     0     0     0     0
week 5 , Wed      0    0    0    0    0    0    0    0    0     0     0     0     0
week 5 , Thurs    0    0    1    0    0    0    0    0    1     0     0     0     0
week 5 , Fri      1    0    0    0    0    0    0    0    0     0     0     1     0
```

week 5 , Sat	0	0	0	1	0	0	0	0	0	0	0	0	1
week 6 , Mon	1	0	0	0	1	0	0	0	0	0	0	0	0
week 6 , Tues	0	0	0	0	0	1	0	1	0	0	0	0	0
week 6 , Wed	0	1	0	0	0	0	0	0	0	0	0	1	0
week 6 , Thurs	1	0	0	1	0	0	0	0	0	0	0	0	0
week 6 , Fri	0	0	0	0	0	1	0	0	0	0	0	1	0
week 6 , Sat	0	0	0	0	0	1	0	0	1	0	0	0	0
week 7 , Mon	0	0	0	0	0	0	0	1	1	0	0	0	0
week 7 , Tues	0	0	0	0	0	0	0	0	0	0	0	0	0
week 7 , Wed	0	0	0	0	0	0	1	0	0	0	0	0	1
week 7 , Thurs	1	0	0	0	0	0	0	1	0	0	0	0	0
week 7 , Fri	1	1	0	0	0	0	0	0	0	0	0	0	0
week 7 , Sat	0	0	0	0	0	0	1	0	0	0	1	0	0
week 8 , Mon	1	0	0	0	0	1	0	0	0	0	0	0	0
week 8 , Tues	0	0	1	0	1	0	0	0	0	0	0	0	0
week 8 , Wed	1	1	0	0	0	0	0	0	0	0	0	0	0
week 8 , Thurs	0	0	0	0	0	0	0	0	0	0	0	0	0
week 8 , Fri	0	0	0	0	1	0	0	1	0	0	0	0	0
week 8 , Sat	0	0	0	0	0	0	0	1	0	0	1	0	0

$`stratum 3`

	[,1]	[,2]	[,3]	[,4]	[,5]	[,6]	[,7]	[,8]
week 1 , Mon	0	0	0	0	0	0	1	1
week 1 , Tues	1	0	0	0	0	1	0	0
week 1 , Wed	0	0	0	0	1	0	0	1
week 1 , Thurs	0	1	0	1	0	0	0	0
week 1 , Fri	1	0	0	0	0	1	0	0
week 1 , Sat	1	0	1	0	0	0	0	0
week 2 , Mon	1	0	0	0	0	0	0	1
week 2 , Tues	0	0	1	0	0	0	0	1
week 2 , Wed	1	1	0	0	0	0	0	0
week 2 , Thurs	1	0	0	0	0	1	0	0
week 2 , Fri	0	0	0	0	1	0	1	0
week 2 , Sat	0	1	0	0	1	0	0	0
week 3 , Mon	0	0	0	1	1	0	0	0
week 3 , Tues	0	0	0	0	0	1	0	1
week 3 , Wed	1	0	0	0	0	0	1	0
week 3 , Thurs	0	0	0	0	0	0	0	0
week 3 , Fri	0	0	0	0	0	0	0	0
week 3 , Sat	0	0	0	0	0	1	1	0
week 4 , Mon	0	0	1	1	0	0	0	0
week 4 , Tues	0	0	0	0	0	0	0	0
week 4 , Wed	1	0	0	0	0	1	0	0
week 4 , Thurs	0	1	0	0	0	0	0	1
week 4 , Fri	0	0	0	0	0	1	1	0

week 4 , Sat	1	0	0	1	0	0	0	0
week 5 , Mon	0	0	0	0	0	0	0	0
week 5 , Tues	0	0	0	0	0	0	0	0
week 5 , Wed	0	1	0	0	0	0	1	0
week 5 , Thurs	0	0	0	0	0	0	0	0
week 5 , Fri	0	0	0	0	0	0	0	0
week 5 , Sat	0	1	0	0	0	0	1	0
week 6 , Mon	0	0	1	0	0	0	0	1
week 6 , Tues	0	1	0	0	0	1	0	0
week 6 , Wed	0	1	0	0	0	0	1	0
week 6 , Thurs	0	1	0	1	0	0	0	0
week 6 , Fri	0	1	0	0	1	0	0	0
week 6 , Sat	0	0	0	0	0	0	1	1
week 7 , Mon	0	0	0	0	0	1	0	1
week 7 , Tues	1	0	1	0	0	0	0	0
week 7 , Wed	0	1	0	0	0	1	0	0
week 7 , Thurs	0	0	0	1	1	0	0	0
week 7 , Fri	0	0	0	0	1	0	0	1
week 7 , Sat	0	0	0	0	0	0	1	1
week 8 , Mon	0	1	0	1	0	0	0	0
week 8 , Tues	0	1	0	0	1	0	0	0
week 8 , Wed	0	0	1	0	0	0	0	1
week 8 , Thurs	0	1	0	1	0	0	0	0
week 8 , Fri	0	0	0	1	0	0	0	1
week 8 , Sat	0	0	0	0	0	0	0	0