N A T U R A L R E S O U R C E S

Variance Estimation for STR Transect Sampling

April 22, 2021

PUGET SOUND ECOSYSTEM MONITORING PROGRAM

WASHINGTON STATE DEPT OF NATURAL RESOURCES

Variance Estimation for STR Transect Sampling

April 22, 2021

Pete Dowty
Bart Christiaen
Helen Berry

Nearshore Habitat Program
Aquatic Resources Division

Puget Sound Ecosystem Monitoring Program

Washington State Dept of Natural Resources

Acknowledgements

The Nearshore Habitat Program is part of the Washington Department of Natural Resources' (DNR) Aquatic Resources Division, the steward for state-owned aquatic lands. The Nearshore Habitat Program monitors and evaluates the status and trends of marine vegetation for DNR and the Puget Sound Partnership. Program funding is provided primarily through DNR, in support of its mandate to sustainably manage aquatic resources.

Washington State Department of Natural Resources
Aquatic Resources Division
1111 Washington St. SE
P.O. Box 47027
Olympia, WA 98504-7027

www.dnr.wa.gov

Contents

Executive Summary

The Washington Department of Natural Resources (DNR) has conducted annual monitoring of the seagrasses in greater Puget Sound since 2000 through the Submerged Vegetation Monitoring Program (SVMP) (Christiaen et al. 2019). For selected sites, the SVMP estimates the abundance of native seagrasses and the change in abundance over time. These estimates are based on randomly selected transects that are surveyed with underwater video and classified for the presence of native seagrass. These site estimates, in turn, are the basis for regional estimates and the eelgrass vital sign indicator reported by the Puget Sound Partnership.

Initially transects were selected at a site by simple random selection (SRS). Unbiased estimators were available for seagrass area and variance we well as seagrass change and its variance (Skalski 2003). While the set of all possible SRS samples has the desirable property of being unbiased, when the population sampled is heterogeneous and the SRS sample by chance exhibits clumping of transects, the sample can be a poor representation of the population. In such a case, the result obtained may depart strongly from the true value to the point where this can be discerned just upon inspection.

In response to these concerns, starting in 2013 transects were also selected by stratified random selection with one transect per stratum (STR). The STR transects were first used on an exploratory basis but have since been made the primary method for SVMP transect selection.

The benefit of STR over SRS is better spatial distribution over the sampling area and greater precision in the estimate of site seagrass area, especially at sites with spatial heterogeneity in seagrass abundance or in change in abundance. The challenge with STR sampling is that there is no generally unbiased variance estimator for these samples. This challenge also applies to systematic sample selection (SYS), which has been the subject of work to assemble a set of alternative variance estimators (Wolter 1984) and apply them to vegetation sampling with transects (McGarvey et al. 2016). The performance of these variance estimators was found to be mixed in the case of SYS sampling (McGarvey et al. 2016).

In this study, we evaluate these SYS variance estimators modified for analysis with the SVMP STR samples. Our approach was to construct spatial models of native seagrass for six contrasting sites based on available SVMP data. We then drew and analyzed a large number of Monte Carlo transect samples from these site models. We analyzed these results in aggregate to assess the performance of the variance estimators in terms of precision and bias. We also evaluated performance in terms of confidence interval coverage and bias in interval width.

We compared the performance of the standard SRS variance estimator (denoted v1) with six alternative estimators developed for analysis of SYS samples. None of these estimators performed well consistently across the six site models. The v1 estimator was the only one to not exhibit negative bias for any of the six models. This is important because negative bias leads to greater risk of spurious results (e.g., a conclusion that seagrass area is significantly different than some reference value, or that a change estimate is significant).

Normally the concern about a positively biased variance estimator is that precision is diminished and the power to detect change is reduced. However, if the change from SRS to STR transects is considered with other methodological changes made by the SVMP (repeat transect sampling and full site sample polygons), power to detect change is increased in most cases and at least not reduced in the most challenging cases.

We recommend the use of the v1 estimator for the analysis of SVMP STR samples based on the low risk of negative bias as and the expectation of greater power when the suite of methodological changes is considered together. However, due to the potentially high positive bias in variance estimates, results must be treated with greater caution than in the SRS case. Consideration of the distribution of values across the transects in a sample can aid in interpreting results.

In the future as the STR data record is extended, the SVMP should place greater weight on trend analyses that do not require variance estimates as input. This approach will gain the benefit of greater precision with STR while avoiding the difficulty in variance estimation. In cases where STR variance estimates are used and there is concern about bias, development of additional site models may be helpful in quantifying the level of bias.

1 Introduction

1.1 The Submerged Vegetation Monitoring Program

The Submerged Vegetation Monitoring Program (SVMP) has conducted annual monitoring of the seagrasses in greater Puget Sound since 2000 (Christiaen et al. 2019). For native seagrasses, estimates are made of their abundance and the change in their abundance at a sample of sites across the study area. These sites are selected under a regional monitoring design that also provides for estimation of regional seagrass abundance and change based on the site-level estimates.

The Washington State Department of Natural Resources (DNR) implements the SVMP. DNR initiated seagrass monitoring in its role as steward of state-owned aquatic lands and the attached or embedded resources such as seagrass. State-owned aquatic lands in Washington total 2.6 million acres (1.1 million hectares) and include all subtidal areas and a substantial amount of the state's intertidal lands.

The monitoring (or sampling) design specifies how data is collected and analyzed and is critical to ensure analyses and associated estimates are reliable. As the monitoring record has grown this has allowed for assessment of the sampling design performance. Also, the priorities of the monitoring program have shifted since the onset of the program. For example, there has been an increasing focus on change and trend detection as opposed to seagrass area estimation. Also, there has been an increasing focus on the quality of site level results independent of the quality of the soundwide results. Both the work on sampling design performance and the shifting priorities have prompted adjustments to the design. This report assesses one of these adjustments and makes a recommendation for associated analysis.

1.2 Stratified Random Sampling (STR) and Variance

This report focuses on the use of stratified random sampling with one unit per stratum (STR) that was introduced into SVMP site sampling in 2013. It was initially introduced on an exploratory basis but the program has now made this the primary method for transect selection in sampling of most sites. The move to STR transect selection was part of a suite of methods changes that included transect sample replacement policy and sampling area delineation.

STR sampling presents a unique challenge because there is no generally unbiased variance estimator for STR samples. The standard variance estimator is unbiased only for simple random sampling (SRS). The same issue applies to systematic sampling (SYS) and the SVMP has relied on previous work on SYS variance estimators to inform evaluation of potential STR variance estimators (Wolter 1984, McGarvey et al. 2016).

Wolter (1984) and McGarvey et al. (2016) evaluated a collection of variance estimators denoted v1 through v8 (McGarvey et al. also evaluated several additional estimators that are not considered here). The estimator denoted as v1 is just the standard SRS variance estimator. McGarvey et al. found the v8 estimator to be the most promising for the SYS transect sampling of different simulated vegetation patterns, but they also noted that its performance was inconsistent and could be very poor in some scenarios.

A previous SVMP study showed that the bias incurred by using the standard variance estimator on STR samples in SVMP sampling can be large and highly variable between sites with different spatial patterns of seagrass and change (Dowty et al. 2017).

Based on the results of McGarvey et al. (2016), the previous SVMP work (Dowty et al. 2017) included evaluation of the v8 estimator for contrasting site seagrass models based on data from core001 and flats26. That work showed that the v8 estimator outperformed the v1 estimator for some STR site sampling scenarios, but that it was inconsistent and, most importantly, could have negative bias, thereby increasing the likelihood of elevated Type I error (false positives, or spurious results). Dowty et al. recommended further work that evaluated the other Wolter estimators (v2-v7) on SVMP site models. That recommendation led to the work presented in this report.

Additional SVMP work developed a broader suite of site models to support expanded modelling work. While the focus of that work was to evaluate sensitivity of change detection to sample size, the additional site models are more generally useful and were included in the work reported here.

1.3 Objectives and Approach

The overall goal of this work was to contribute to the development of SVMP operational procedures for analysis of STR data. Five alternatives were identified:

1. Use the standard SRS variance estimator (denoted here as v1) for STR samples.
2. Select one of the alternative Wolter variance estimators (v2-v8) for STR samples
3. Select a variance estimator on a site-by-site basis based on an assessment of site characteristics.
4. Select a variance estimator on a site-by-site basis based on modelling studies that use site models developed for each site.
5. Deprecate the presentation and use of results that rely on STR variance estimation, regardless of which estimator is used based on the premise that no consistently highly performing variance estimator is available. This would preclude change analysis between two sampling occasions (variance estimates are required for this) but would not preclude trend analysis that estimates variance from residuals.

The goal, then, was to narrow down this list of alternatives or highlight top alternatives for further consideration.

More specifically, the objectives of this work were to:

- Evaluate the entire suite of Wolter (1984) variance estimators (v1-v8) for SVMP site sampling with STR.
- Synthesize this work with previous work to present current understanding of issues surrounding the use of STR transect selection.
- Recommend an STR variance estimator for immediate operational use, and possibly a path forward to reach a longer-term recommendation based on further analysis.

2 Methods

This study uses a modelling approach to assess the performance of variance estimators. First, spatial models of seagrass distribution at six contrasting sites were selected from previous SVMP work (Dowty et al. 2017 and unpublished analysis of sample size). Then these site models were subjected to simulated transect sampling in a Monte Carlo framework using 10,000 iterations for each variance estimator and site. Each Monte Carlo iteration involves selection of an STR transect sample of size n=10 transects and the calculation of sample estimates. The resulting population of sample estimates comprises the sampling distribution and allows for calculation of population parameters and the performance of sample estimates in terms of proximity to the population parameters.

The Monte Carlo sampling, the associated calculations and the visualization of results was conducted in R.

The SVMP is concerned with three categories of estimates at the site level:

(1) seagrass area estimates at a given sample occasion
(2) change in seagrass area between two sampling occasions
(3) trend in seagrass area over a series of sampling occasions.

Only the estimates of site seagrass area and change in seagrass area will typically require accompanying estimates of sample variance. Trend estimation, e.g., by linear regression, will typically assume variance in seagrass area estimates is the same across all estimates and estimate that variance as part of the regression.

Variance estimators may perform quite differently when estimating variance of site seagrass area and when estimating variance of change in site seagrass area. This is because the distributions of transect fraction and change in transect fraction may be quite different.

The scope of the modelling work conducted for this study was restricted to site seagrass area estimates.

2.1 Site Models

The site models used are simplified spatial representations of the seagrass distribution at selected sites surveyed by SVMP. The models represent sites as a set of 1000 transects of 1 meter width that comprehensively cover the site in the longshore dimension and span the sampling area in the cross-shore dimension. They include detail down to the length of each transect and the vegetation length of each transect based on existing SVMP transect data and some form of interpolation to fill in between surveyed transects. The model does not represent the spatial arrangement of the vegetation along the transects which is not relevant to the estimators used.

Two of the site models used were generated in the work reported in Dowty et al. 2017. These sites are core001 and flats26. Both sites have seagrass distributed through the entire longshore dimension and a simple linear interpolation was used to construct the site models.

Four additional site models were generated in the work conducted by Bart Christiaen in 2015 that includes sites with discontinuous seagrass in the longshore dimension and used an interpolation technique that relied on a spline. This work is documented in a summary presentation, R code and the resulting site models. The site models used from this work include core004, cps1035, cps1054, and cps2565. The cps1035 site model seemed to depart from the available survey data in the seagrass distribution and the transect lengths but it was used in this study as is. Given that each site is modelled as having a 1000 m longshore dimension, the core001 model departs from the actual larger site in terms of spatial scale while retaining the characteristic spatial distribution of seagrass at the site.

The site models are presented in Figure 2-1 and the SVMP survey data used to generate the site models is presented in Figure 2-2.

2.2 Wolter Estimators Adapted for SVMP

Wolter (1984) first compiled the eight variance estimators (v1-v8) for SYS samples. These were later assessed for sampling of vegetation spatial patterns with transects selected by SYS (McGarvey et al. 2016). Here, we utilize these same estimators but apply them to STR samples and make modifications to conform to SVMP requirements.

The v7 estimator was excluded from this study because it was not completely clear how to implement the estimator as presented by Wolter (1984). Given the poor performance of v7 reported by McGarvey et al. (2016), this estimator was dropped from this study rather than dedicating effort to resolve this ambiguity. The R code used by McGarvey et al. (2016) was obtained from the lead author so their specific implementation of v7 is available if assessing this estimator becomes a priority in the future.

The v1-v8 estimators of Wolter (1984) and McGarvey et al. (2016) give an unweighted sample estimate. These estimators had to be adapted for use within the SVMP framework because site sample estimates are weighted by transect length. When the standard variance estimator (v1) is weighted (e.g., to estimate the variance associated with the weighted mean) each term of the sum of squares represents one sample unit and it is multiplied by an attribute of that sample unit to serve as a weighting. The v2-v8 estimators have an added complication that each term of the sum of squares involves multiple sample units. The approach taken here was that each weight was a sum of terms involving transect lengths that followed the form of the term being squared. This is shown in the adapted estimators used in this study (Table 2-1).

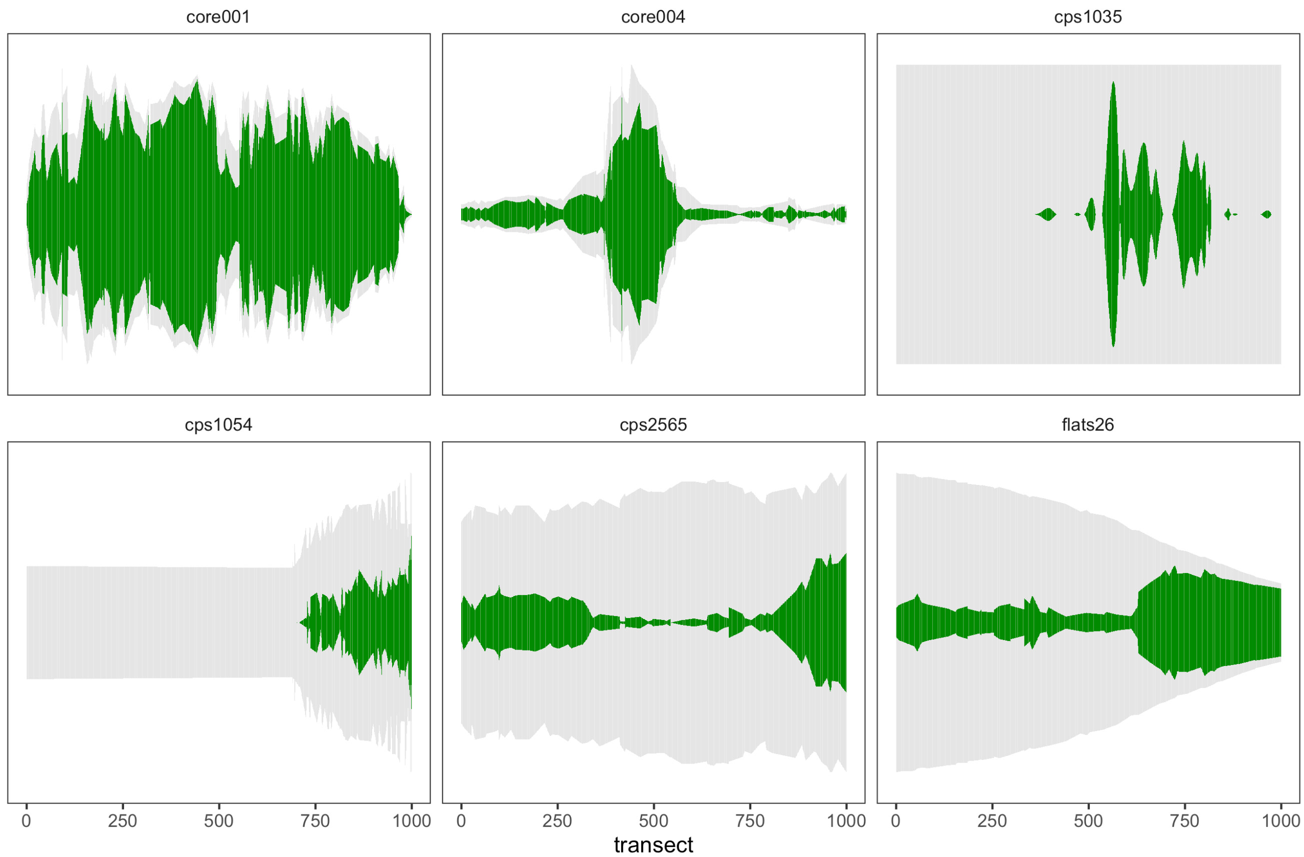

Figure 2-1. The six site models used in this study. The longshore dimension of each site is represented as 1000 1m wide transects arranged along the x-axes in these figures. The y-axes of the figures represent the cross-shore dimension. Green areas represent the vegetated length along each transect. Gray areas represent segments of the transects where seagrass is absent. The site survey data that were the basis for these site models are presented in Figure 2-2.

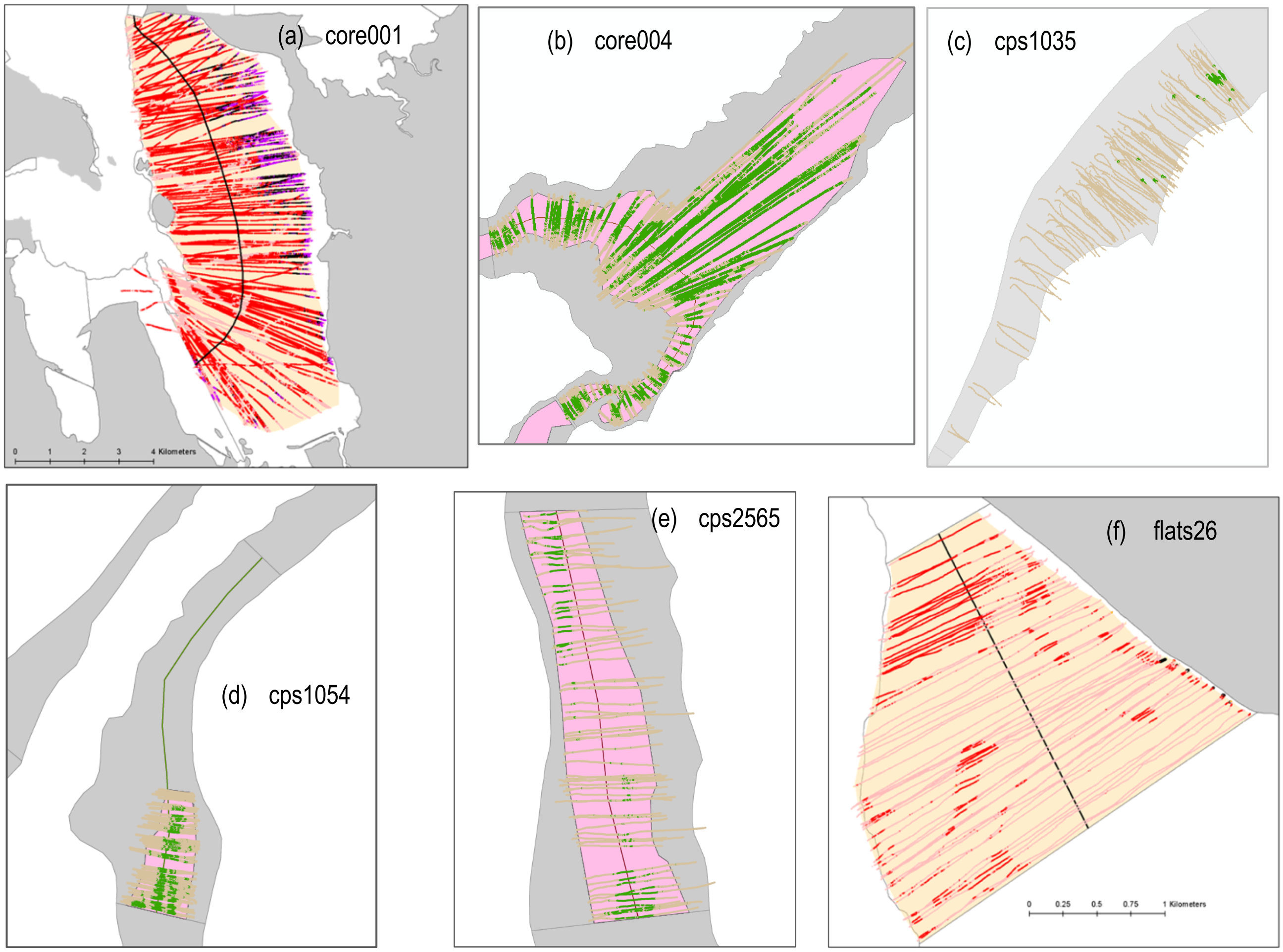

Figure 2-2. SVMP site survey data that was used to develop the site models in Figure 2-1. In (a) and (f), red indicates eelgrass, black represents *Z. japonica* and purple is mixed. For the remaining panels, green represents native seagrass.

Table 2-1. The variance estimators of Wolter (1984) adapted with weighting for use as SVMP site estimators. The v1 estimator (denoted here as $\widehat{Var_1}$) is the standard SRS site variance estimator from Skalski (2003).

$\widehat{Var_1}\,(\hat{\bar{p}}) = \dfrac{\sum_{i=1}^{n}(l_i - \hat{\bar{p}}L_i)^2}{n(n-1)\bar{L}^2}$	$\widehat{Var_5}(\hat{\bar{p}}) = \dfrac{\sum_{i=5}^{n} w_i\left({p_i}/{2} - p_{i-1} + p_{i-2} - p_{i-3} + {p_{i-4}}/{2}\right)^2}{(\sum_{i=2}^{n} w_i)3.5n(n-4)}$ $w_i = {L_i}/{2} + L_{i-1} + L_{i-2} + L_{i-3} + {L_{i-4}}/{2}$
$\widehat{Var_2}(\hat{\bar{p}}) = \dfrac{\sum_{i=2}^{n} w_i(p_i - p_{i-1})^2}{(\sum_{i=2}^{n} w_i)2n(n-1)}$ $w_i = L_i + L_{i-1}$	$\widehat{Var_6}(\hat{\bar{p}}) = \dfrac{\sum_{i=9}^{n} w_i\left({p_i}/{2} - p_{i-1} + p_{i-2} - p_{i-3} + \dots {p_{i-8}}/{2}\right)^2}{(\sum_{i=9}^{n} w_i)7.5n(n-8)}$ $w_i = {L_i}/{2} + L_{i-1} + L_{i-2} + L_{i-3} + \dots {L_{i-8}}/{2}$
$\widehat{Var_3}(\hat{\bar{p}}) = \dfrac{\sum_{i=1}^{n/2} w_i(p_{2i} - p_{2i-1})^2}{\left(\sum_{i=1}^{n/2} w_i\right)n^2}$ $w_i = L_{2i} + L_{2i-1}$	$\widehat{Var_8}(\hat{\bar{p}}) = \begin{bmatrix} \widehat{Var_1}(\hat{\bar{p}})\left[1 + {2}/{ln(\widehat{\rho_k})} + {2}/{\left(\widehat{\rho_k}^{-1} - 1\right)}\right] & \widehat{\rho_k} > 0 \\ \widehat{Var_1}(\hat{\bar{p}}) & \widehat{\rho_k} \le 0 \end{bmatrix}$
$\widehat{Var_4}(\hat{\bar{p}}) = \dfrac{\sum_{i=3}^{n} w_i(p_i - 2p_{i-1} + p_{i-2})^2}{(\sum_{i=2}^{n} w_i)6n(n-2)}$ $w_i = L_i + 2L_{i-1} + L_{i-2}$	$\widehat{\rho_k} = \dfrac{\sum_{i=2}^{n}(L_i + L_{i-1})(p_i - \hat{\bar{p}})(p_{i-1} - \hat{\bar{p}})}{\left(\sum_{j=2}^{n}(L_j + L_{j-1})\right)n(n-1)\widehat{Var_1}(\hat{\bar{p}})}$

Symbols:
l_i = the vegetated length of transect i
L_i = the total length of transect i
$\hat{\bar{p}}$ = the sample estimate of mean transect vegetated fraction at a site
$\bar{L}$ = mean transect length at a site
w_i = weighting term

2.3 Calculation of Estimator Performance Metrics

For each Monte Carlo iteration, a sample estimate of site seagrass area (A) is calculated using the SVMP estimator from Skalski (2003), i.e., the product of the weighted estimate of mean transect fraction ($\hat{\bar{p}}$) and the sample polygon area (E), or

$$\hat{A} = \hat{\bar{p}} \cdot E$$

The sample polygon area in m^2 is simply the sum of all transect lengths (in meters) in the site model (n=1000) since the transects are 1 m wide. This expressed as

$$E = \sum_{i=1}^{1000} L_i$$

Similarly, the true seagrass area of each site model is

$$A = \sum_{i=1}^{1000} l_i$$

The variance estimators (Table 2-1) are first evaluated on their precision and accuracy. These performance metrics are easily compared across variance estimators by comparing the associated box plots that show dispersion in the variance estimates and position of the estimates relative to the true variance. Each box summarizes the variance estimates from the 10,000 Monte Carlo iterations for each site and variance estimator.

The true variance is calculated as the variance of the set of all site seagrass area estimates for a site (7 sets x 10,000 iterations = 70,000 estimates).

The sample confidence interval widths were calculated with *z* rather than *t* statistics. While the *t*-based interval is more appropriate with n=10 transects, it was not clear how the degrees of freedom vary across the variance estimators and it was assumed that the differences in results across estimators would be much larger than the difference between *z*- and *t*-based confidence intervals. The sample estimate of the 95% confidence interval half-width based on the j^{th} variance estimator was calculated as

$$\widehat{CI_{95}} = z_{(1-\alpha/2)} \cdot \sqrt{\widehat{Var_j}(\hat{A})}$$

with $\alpha = 0.05$ and the sample confidence interval was determined as

$$\hat{A} \pm \widehat{CI_{95}}$$

The confidence interval coverage for a given variance estimator gives the proportion of the sampling distribution encompassed within the confidence interval. In principle, a 95% confidence interval should encompass 95% of the sampling distribution. In practice, any given confidence interval estimate can be expected to vary from the nominal value. We calculate the confidence interval coverage here simply as the proportion of 10,000 sample estimates that fall within the estimated confidence interval. This is given by

$$CI_{cover} = \frac{\sum_{i=1}^{10000} m_i}{N_{tot}}$$

where $N_{tot} = 10{,}000$ and

$$m_i = \begin{cases} 0 & A < \left(\hat{A} - \widehat{CI_{95}}\right) \; OR \; A > \left(\hat{A} + \widehat{CI_{95}}\right) \\ 1 & A \geq \left(\hat{A} - \widehat{CI_{95}}\right) \; AND \; A \leq \left(\hat{A} + \widehat{CI_{95}}\right) \end{cases}$$

The true 95% confidence interval width for each site, CI_{95}, is found iteratively starting with an initial guess. The bisection root finding method is used to evaluate the coverage of the initial guess, to improve the guess and then evaluate again. The true 95% confidence interval was found when subsequent guesses did not vary by more than 0.0001 ha.

Confidence interval bias was the departure of the mean of the confidence interval estimates from the true confidence interval. This is expressed as

$$Bias = \overline{\widehat{CI_{95}}} - CI_{95}$$

3 Results

3.1 Variance Results

The performance of the variance estimators is shown in Figure 3-1 when estimating site seagrass area on a given sampling occasion with STR transects. In general, the variance estimators are highly biased. Of the 42 scenarios assessed (6 sites x 7 estimators), the boxes in Figure 3-1 only span the true variance in 8 scenarios. In some cases, the entire variance sampling distribution is disjunct from the true variance.

More problematic for the operational utility of these variance estimators is the fact that the bias for four of the seven estimators (v1, v3, v6, v8) clearly varies between sites. The bias is not only large but also variable. Estimators v2-v4 appear consistently negatively biased across all sites but some lesser variation would be apparent with an expanded scale. The v1 estimator is the only one whose bias is positive across all sites although the magnitude of the bias shows the greatest variation among the estimators.

3.2 Confidence Interval Results

The confidence interval coverage results are presented in Figure 3-2. The ramifications of negative bias in variance estimation are clearly seen in the low coverage of the confidence interval that reach as low as 40%, rather than the nominal 95%. The ramifications of positive bias in variance estimation are not as apparent because the coverage scale only extends to 100% regardless of how oversized the confidence intervals are.

One simple way to compare variance estimators is the proximity of coverage to the nominal 95%. If we define coverage values in the range of 90-99% as in close proximity to the nominal 95%, then three of the seven variance estimators lead to at least one instance with coverage that is proximal to the nominal value:

- v1: coverage proximal to 95% for three sites (core001, core004, cps1035)
- v3: coverage proximal to 95% for one site (cps2565)
- v8: coverage proximal to 95% for four sites (core001, core004, cps1035, cps1054)

Bias in confidence interval width is presented in Figure 3-3. If we define the range of -20% to +20% as having modest levels of bias, then there are three variance estimators leading to at least one instance of modest bias:

- v3: modest bias for two sites (cps2565, flats26)
- v6: modest bias for three sites (core001, cps1035, flats26)
- v8: modest bias for one site (cps1035).

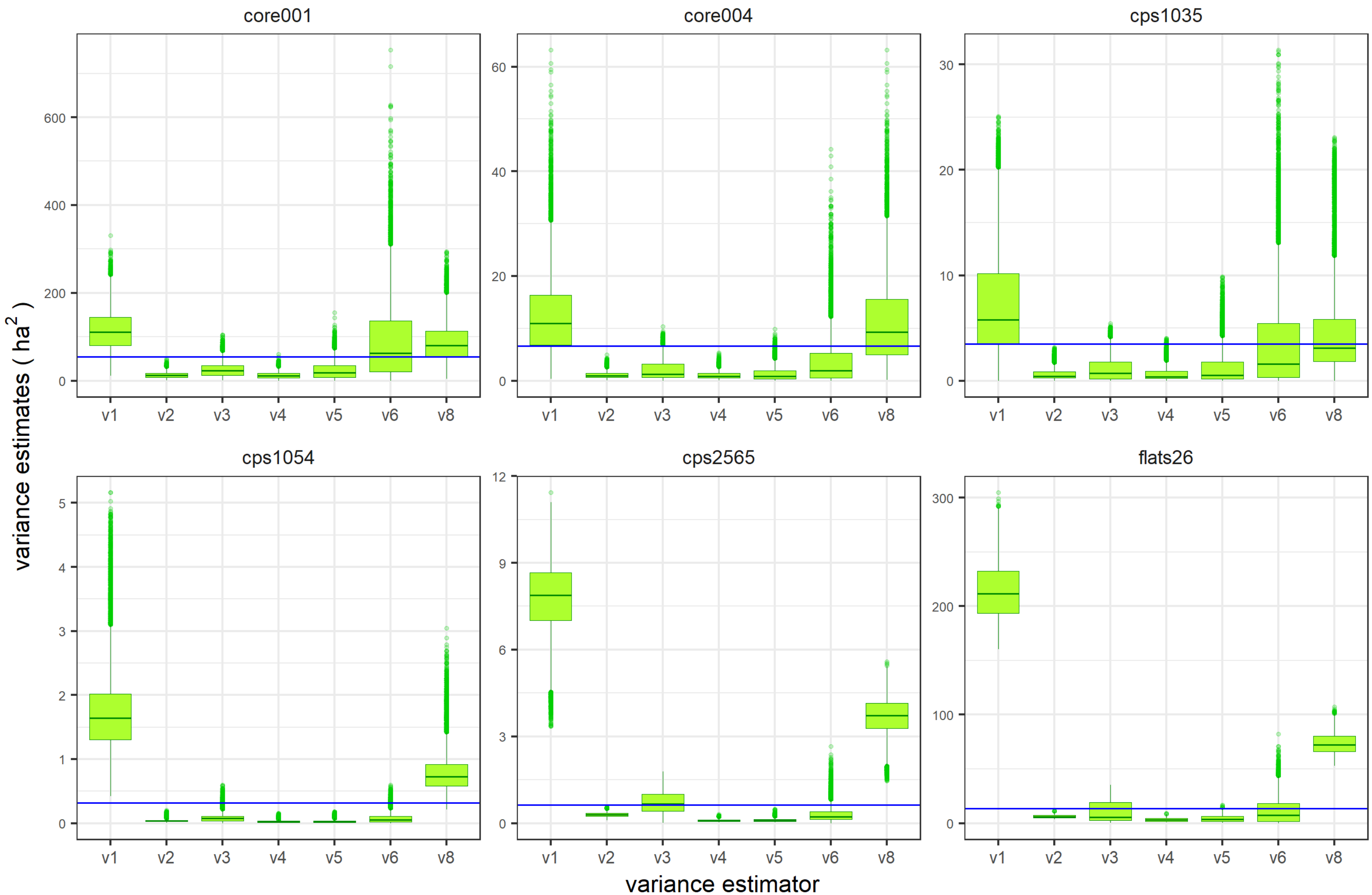

Figure 3-1. Box plots that depict performance of the v1-v8 variance estimators when estimating site seagrass area with STR transect samples. The six panels show results for the six different site models that were sampled. The horizontal blue lines represent the true variance that is being estimated. The box size represent precision (dispersion in estimates) and the position of the box relative to the blue line represents accuracy (distance from blue line to box mean [not shown] is bias).

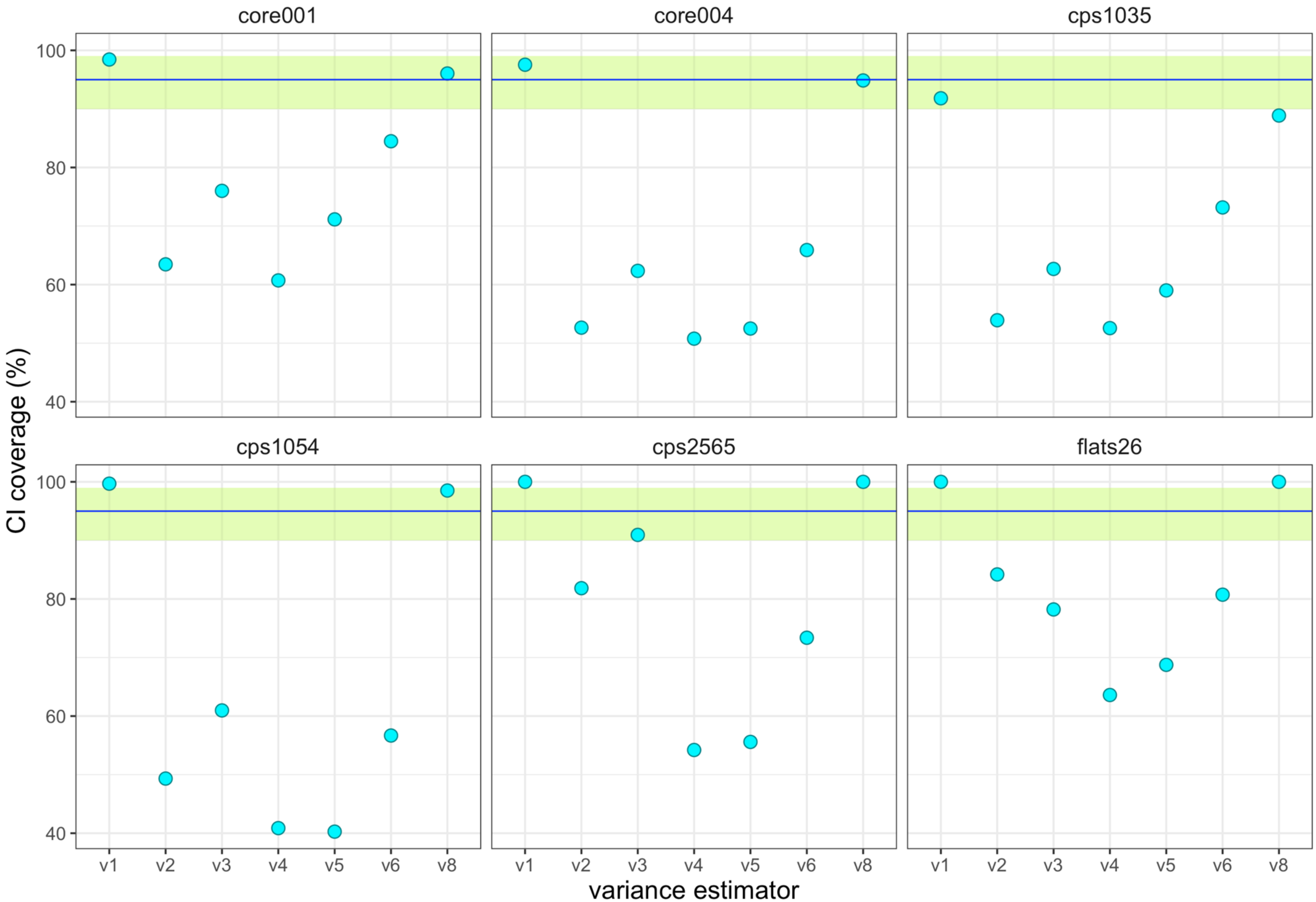

Figure 3-2. Coverage of the 95% confidence intervals based on the v1-v8 variance estimators. The six panels show results for the six different site models that were sampled. The horizontal blue line indicates 95%. The green region highlights a region in the vicinity of the target 95% (90-99%).

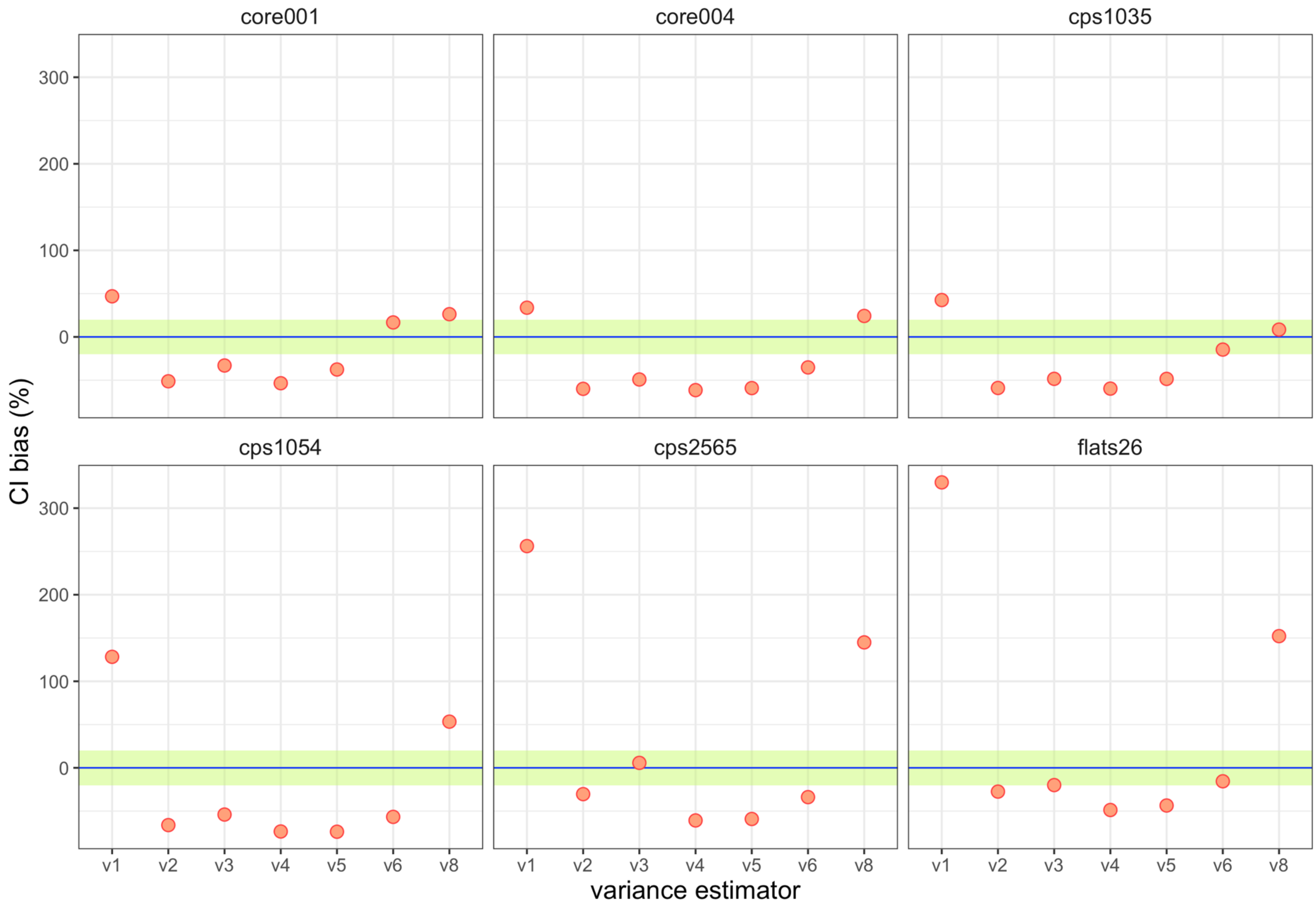

Figure 3-3. Bias in the 95% confidence interval estimators based on the v1-v8 variance estimators. The green region highlights a region in the vicinity of zero bias (-20 to 20%).

4 Discussion

4.1 Current Understanding of STR Analysis Options

Earlier we demonstrated that in SVMP sampling scenarios the standard variance estimator (v1) tends to be biased with STR samples (Dowty et al. 2017). In that work, we also demonstrated that the v8 estimator, given a mixed recommendation by McGarvey et al. (2016), tended to have a lower magnitude bias but it tended to be negative while the v1 bias tended to be positive. The v8 estimator had potential but the negatively biased variance seemed problematic.

The main purpose of this study was to extend that investigation to include the other estimators in the v1-v8 set presented by Wolter (1984). Now, all the Wolter estimators (except v7) have been evaluated in simulated SVMP scenarios.

We have now seen that none of the v1-v8 estimators is consistently unbiased with STR sampling and they are typically biased. In addition, no estimator was clearly better performing than the v1 estimator.

While the earlier work showed that the v1 estimator tended to be positively biased with STR samples, there was also an instance of negative bias. The fact that the bias could possibly be of either sign makes inference more challenging than if bias was known to be of one sign.

But the shortcomings of the v1 estimator in the STR context must be considered in the context of the suite of site sampling method changes that have been made by the SVMP in recent years. In particular, the shift from sample replacement (new sample draw) each occasion to sample retention (repeat transects), in conjunction with the shift to full site polygons, can be expected to result in a large net gain in power to detect change. This is shown clearly in the simulation results in Figure 4-1 which were drawn from Dowty et al. (2017).

In Figure 4-1, there is a strong gain in power to detect change when new draw SRS transects are replaced with repeat STR transects. This is true even with the shortcomings of the v1 estimator when used with STR samples. This gain in power is seen in almost all of the different spatial change scenarios investigated.

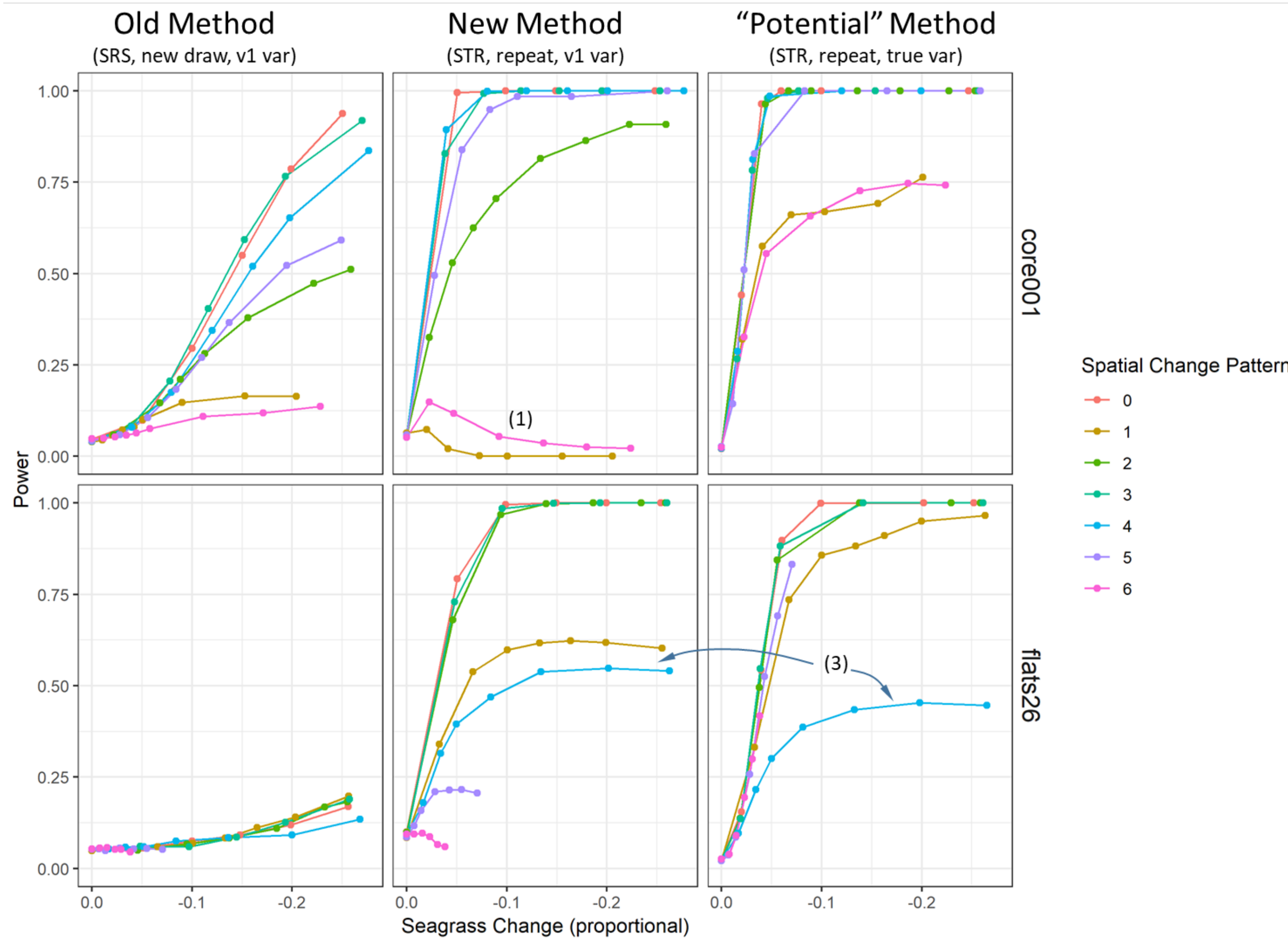

Figure 4-1. Power to detect change (y-axis) of increasing intensity (x-axis is proportional change) under the "old" site sampling design (left), the "new" sampling design (middle) and a hypothetical potential design (right) for the core001 site model (top row) and flats26 model (bottom row). The old method is SRS transect selection with newly drawn transects each occasion and use of the v1 variance estimator. The new method utilizes STR, repeat transects and the v1 estimator. The potential method also utilizes STR and repeat transects but instead of estimating variance it uses the true variance available from the Monte Carlo results. Each curve represents a specific spatial pattern of change with spatial pattern 0 being homogeneous change throughout the site while change in the other scenarios is restricted to different subsets of the site. These results were extracted from Dowty et al. (2017).

Figure 4-1 also compares the power to detect change using the “new” method (STR, repeat, v1 estimator) with the “potential” power that would hypothetically be achieved with a perfect variance estimator. This comparison does not show strong loss of power associated with the use of the v1 estimator with STR. The “potential” power was assessed by conducting statistical testing with the true value of variance obtained from the Monte Carlo results. This result suggests that even if a more reliable STR variance estimator could be identified, the gain in power would be modest relative to usage of the v1 estimator. Given the attention this problem has received in the literature (Cochran 1977, McGarvey et al. 2016, Wolter 1984) as well as the work conducted here and in Dowty et al. 2017, without successfully identifying a reliable alternative variance estimator, it is not clear that continued effort along these lines would be successful. Certainly, such an effort would be a major time-intensive undertaking to move beyond the work already reported.

An alternative approach given at the beginning of this report (section 1.3, p.4) was that the optimal STR variance estimator among v1-v8 could be selected on a site-by-site basis. This approach acknowledges that there is no single optimal STR estimator, but that one of the v1-v8 estimators may be optimal for a particular situation based on a site’s seagrass spatial pattern and the spatial pattern of change. Our assessment from the results presented in this study is that this approach is not promising. For example, the site models for core004, cps2565 and flats26 (Figure 2-1, p.9) have basic similarities in being essentially two-phased beds that were continuous in the longshore dimension but estimator performance varied strongly among these cases (Figure 3-1, p.14). While the reason for the variable performance among these could presumably be identified through focused modelling work, it seems ambitious to think a general set of rules could be developed to reliably assign all sites to an optimal estimator based on inspection.

Given the challenges of variance estimation with STR samples, it is relevant to ask if there is a benefit to STR over SRS. This question was not addressed in the work conducted for this study, but Dowty et al. (2017) showed how STR achieves a major gain in precision, particularly in spatially heterogeneous cases where SRS performs poorly. This is the motivation for integrating STR selection into the SVMP site sampling.

4.2 Recommendation for STR Analysis

For estimating variance of statistics derived from STR samples, we recommend using the v1 estimator. This includes usage for both confidence interval estimation and statistical testing.

Given the expected bias associated with using the v1 estimator with STR samples, the results must be used and interpreted with caution. This applies to the variance estimates themselves, derived confidence interval estimates and testing results. This is especially true since the bias is of uncertain magnitude or sign. In the expert review of site results, this essentially lessens the weight placed on the statistical results and puts more weight on other available results and the expert assessment.

Deprecating the influence of the statistical result seems to parallel the drive to change the scientific community’s reliance on p-values (Smith 2020, Wasserstein and Lazer 2016).

They both emphasize inference based on weighing of multiple pieces of information rather than deference to the clarity of the threshold-based statistical test.

We also recommend that as soon as a data record of STR repeat transects begins to accumulate, we should look for opportunities to apply trend analysis. As long as the assumption of uniform variance (homoscedasticity) is reasonable, a trend analysis such as linear regression does not require variance estimates as input. This effectively bypasses the challenges of variance estimation with STR samples. This doesn't hold where we cannot assume uniform variance in which case the variance estimates are needed for a weighted regression.

4.3 Future Work

With a recommendation in place, there is no further work identified that would be needed for implementation of STR sample analysis. This position should be re-evaluated after STR analyses have been implemented for some time. If it becomes clear that the lack of reliable variance estimates is problematic in site analysis, then further work may be warranted.

One path toward more reliable variance estimates would be to develop spatially-explicit site models for every site sampled. These models would be subjected to Monte Carlo STR sampling to determine exact variance for sampling at a given modelled site. The reliability of these results relies mainly on the integrity of the spatial representation in the model. In the context of change analysis, the reliability depends on the integrity of the spatial representation of the pattern of change (rather than abundance). Change models will be more challenging to construct than abundance models.

Even in cases where integrity of the model representation is in question for a particular site, it may be possible to run multiple sets of Monte Carlo iterations with contrasting model characteristics that are thought to bracket the true value. Comparison of the variance from the bracket points and the v1 sample variance should indicate whether the v1 estimate needs adjustment. Such adjustment could be informed by the model results.

There are certainly other approaches that could be pursued to improve the STR variance estimation. But this future work should only be considered if the initial implementation indicates it is necessary.

5 References

Berry, H.D., A.T. Sewell, S. Wyllie-Echeverria, B.R. Reeves, T.F. Mumford, Jr., J.R., Skalski, R.C. Zimmerman and J. Archer. 2003. *Puget Sound Submerged Vegetation Monitoring Project: 2000-2002 Monitoring Report*. Nearshore Habitat Program, Washington Department of Natural Resources, Olympia, Washington. 60pp. plus appendices. Retrieved from http://www.dnr.wa.gov/publications/aqr_nrsh_00_02svmp_rpt.pdf

Christiaen, B., L. Ferrier, P. Dowty, J. Gaeckle and H. Berry. 2019. Puget Sound Seagrass Monitoring Report: Monitoring Year 2016-2017. Nearshore Habitat Program. Washington Department of Natural Resources. Olympia, WA. Retrieved from https://www.dnr.wa.gov/publications/aqr_nrsh_svmp_report_2016_2017_data.pdf?e7vfl5

Cochran, W.G. 1977. *Sampling Techniques*. Wiley.

Dowty, P., B. Christiaen, H. Berry. 2017. *Assessment of Sample Selection and Replacement for Eelgrass Monitoring*. Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, WA. Retrieved from: https://www.dnr.wa.gov/publications/aqr_nrsh_site_samp_assess.pdf?xa6yhn

McGarvey, R., R. Burch and J.M. Matthews. 2016. Precision of systematic and random sampling in clustered populations: habitat patches and aggregating organisms. *Ecological Applications*. 26(1):233-248.

Skalski, J.R. 2003. Statistical Framework for Monitoring *Zostera marina* (Eelgrass) Area in Puget Sound. *In*: Berry et al. 2003. *Puget Sound Submerged Vegetation Monitoring Project: 2000-2002 Monitoring Report*. Appendix L. Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, Washington. Available online: http://www.dnr.wa.gov/publications/aqr_nrsh_00_02svmp_rpt.pdf.

Smith, E.P. 2020. Ending Reliance on Statistical Significance Will Improve Environmental Inference and Communication. *Estuaries and Coasts*. 43:1-6.

Wasserstein, R.L. and N.A. Lazer. 2016. The ASA Statement on p-Values: Context, Process and Purpose. *The American Statistician*. 70(2):129-133.

Wolter, K.M. 1984. An Investigation of Some Estimators of Variance for Systematic Sampling. *Journal of the American Statistical Assocation*. **79(388)**: 781-790.