A Study of Sampling and Analysis Methods: Submerged Vegetation Monitoring Project At Year 4

January 2005

Pete Dowty
Nearshore Habitat Program

Aquatic Resources Division
Washington Department of Natural Resources

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Doug Sutherland - Commissioner of Public Lands

PSAMP
Puget Sound Ambient
Monitoring Program

The Submerged Vegetation Monitoring Project is implemented by the
Nearshore Habitat Program. It is a component of the
Puget Sound Ambient Monitoring Program.

WASHINGTON STATE DEPARTMENT OF
Natural Resources
Doug Sutherland - Commissioner of Public Lands

Nearshore Habitat Program
Aquatic Resources Division
Washington State Department of Natural Resources
1111 Washington St SE, 1st floor
P.O. Box 47027
Olympia, WA 98504-7027
U.S.A.

(tel) 360.902.1100
http://www2.wadnr.gov/nearshore/

January 2005

Sampling and Analysis Issues:
Submerged Vegetation Monitoring Project
At Year 4

January 2005

Pete Dowty
Nearshore Habitat Program

Aquatic Resources Division
Washington Department of Natural Resources
Olympia, Washington

Acknowledgements

John Skalski provided valuable comments and guidance during a series of meetings after completion of the first draft of this report. He also provided useful comments on later drafts. His contributions are gratefully acknowledged.

Helen Berry introduced me to the statistical aspects of the project and was very helpful in formulating the questions that motivated this study. She shared her critical thinking in many helpful discussions as this work developed.

Amy Sewell and Betty Bookheim provided the assessment of *Z. marina* area ratio for all the major flat sites that helped develop the model flats dataset.

Blain Reeves was very helpful in sharing his insight and questions particularly with respect to site level analysis issues. Discussions with Jim Norris and Ian Fraser were also very helpful in understanding many of the statistical issues addressed here.

This report also benefited from review and comments from Sandy Wyllie-Echeverria and Tom Mumford.

I also want to thank RoseLynn Wright for her work on preparations for the 2004 sampling season, which allowed me to focus on the work in this report.

Contents

APPENDICES

EXECUTIVE SUMMARY

The Submerged Vegetation Monitoring Project (SVMP) began tracking status and trends in *Zostera marina* (eelgrass) in Puget Sound in 2000 (Berry et al. 2003). This project was initiated by the Washington State Department of Natural Resources and is part of the Puget Sound Ambient Monitoring Program (PSAMP), a multi-agency effort to monitor key components of the Puget Sound environment in support of resource management.

At the core of the SVMP is a sophisticated statistical framework that was designed without the benefit of pre-existing data to characterize the population to be monitored (Skalski 2003). Now that the SVMP has four years of results available (2000-2003), it is an opportune time to assess whether the statistical framework is performing as intended and to formulate specific refinements to the sampling and analysis procedures where appropriate. This is the overall purpose of the work described in this report.

Key outcomes fall into four general categories.

1. Confirmation that the project will meet our key monitoring objective to reliably detect a 20% loss of *Z. marin*a in Puget Sound over 10 years;
2. Recommendations for improvements to the study design that increase our ability to meet monitoring and management needs within current resource constraints;
3. Validation that assumptions made in the initial sampling design before sample data were available were appropriate.

The work reported here was guided by six research questions that emerged during analysis of 2000-2003 data. Each of the main chapters of this report focuses on one of these questions.

A significant portion of this report focuses on sampling and analysis issues specific to the flats sampling stratum, although this was just one of four strata sampled over the 2000-2003 period. The flats stratum was problematic because it had the highest variability and decreased the overall precision of the monitoring results. However, it also provided the greatest opportunities for improvements and for both of these reasons became an important focus of the work reported here.

A series of specific recommendations emerged from this work, some of which were implemented in the 2004 sampling season. In addition, a number of important issues were identified that fell outside the scope of this project. These issues were prioritized for future work to further improve the sampling and analysis components of this monitoring project.

The main results described in the report are highlighted below in the categories of key findings, recommendations and future priorities.

KEY FINDINGS

1. The SVMP is well-positioned to meet its primary monitoring goal of detecting a trend of 20% loss in *Z. marina* over 10 years with an acceptable level of statistical power. We achieve sufficient precision to meet our goal using different methods than originally envisioned. We increased precision by defining an additional flats sampling stratum (finding #3 below). This is a more conservative approach than relying on retrospective adjustment (finding #2 below).
2. The unexpected retrospective adjustment results previously reported (Berry et al. 2003) deviate from theoretical expectations. This is likely explained by a combination of (a) low sample size in the flats stratum coupled with a low rate of site rotation and (b) a highly skewed distribution in the flats stratum and (c) the role of measurement error at the site-level which is not considered in the calculation of theoretical expectations.
3. A systematic procedure identified the three flat sites (flats11, flats12 and flats20) that bear most responsibility for the high variance (low precision) in the flats stratum. When these sites are placed in a separate stratum, precision improves dramatically. This greatly improves the power to detect change at no additional cost to the project. It also minimizes potential problems associated with outlier sites rotating in and out of the sample pool.
4. The ability to detect change in eelgrass abundance depends in part on the pattern of change. Our power to detect change decreases as the pattern deviates from a linear trend. Therefore, interpretation of results should include consideration of the pattern of change.
5. The 20% rate of site rotation used for 2001-2003 sampling is consistent with statistical design considerations and is appropriate for the SVMP. The report discusses the key issues in balancing estimates of status against estimates of trend.
6. The general target of sampling 11 transects per site was supported by two findings:
    a. This target produces adequate estimates at the site level as demonstrated from modeling six sites that were intensively sampled in 2003.
    b. This target appropriately balances competing management goals for precise estimates at both the site scale and the sound-wide scale. The goal to produce precise results at the site level alone calls for sampling fewer sites more intensely. In this case more than 11 transects would be sampled at a site, but with fewer sites the sound-wide results would be unreliable. In contrast, the goal to produce precise results at the sound-wide scale alone calls for sampling more sites by sampling fewer transects at each site, thereby producing very crude estimates at the site scale. There was demonstrated by an analysis of uncertainty in the sound-wide results which reinforces the use of 11 transects as an appropriate balance of SVMP goals.

RECOMMENDATIONS

1. The three flat sites that are most detrimental to precision of stratum estimates should be placed in a new stratum (flats11, flats12 and flats20). All three sites should be measured every year. This recommendation was implemented in the 2004 sampling season.

2. Four recommendations were made for changes to the division of large embayments into discrete flats sampling sites. These changes improve the consistency of the flats sampling frame. These recommendations were implemented in the 2004 sampling season.
3. When the 5- and 10-year tests for decline are implemented, power analysis results should be consulted to ensure the appropriate significance level is used to achieve acceptable power. New power analyses should be performed if the sound-wide coefficient of variation (CV) diverges from current estimate.
4. Retrospective adjustment should be used with caution until the effects of changes to the flats strata are shown to eliminate concerns.
5. The current rate of site rotation (20%) should be maintained.
6. The current target of sampling a minimum of 11 transects at each site should be maintained.
7. The procedure developed to identify outlier flat sites should be applied in the future if new data suggest other sites are overly detrimental to precision. Any future changes to flats strata, however, should be avoided and flat sites should be moved only after they have completed their role as randomly selected sites in the rotational sampling.

FUTURE PRIORITIES

Priority issues for future work are:

1. A power analysis should be performed for the paired site change detection tests, both at the sound-wide and regional scales. This work is critical since paired site analysis is our principle technique for trend detection over short time periods. This power analysis will be used to further interpret results such as the significant decline in eelgrass abundance between 2002 and 2003.
2. A simulation study should be conducted to assess paired site change detection tests for bias at the small sample sizes available for regional estimates. When recommendations for minimal sample sizes are generated, the optimization of site rotation should be revisited.
3. A Monte Carlo study based on an expanded model flats dataset should be conducted to definitively answer whether there is significant bias in the retrospective adjustment results associated with low sample size, skewed distributions and the presence of measurement error.
4. Multi-parameter tests for detecting long-term trend should be developed as alternatives to the 5- and 10-year tests as currently envisioned. The purpose would be to devise tests that have greater power to detect change and are not based on a linear change model and therefore are less sensitive to the pattern of change.
5. Multi-parameter tests should be developed for identifying individual sites undergoing change in *Z. marina* area. The purpose would be to produce a list of candidate sites that might benefit from increased scrutiny and might be of particular interest to local community and governmental entities.

6. A study should be conducted to determine the optimal allocation of sampling effort among all strata. Results in this report suggest that the wide fringe strata would benefit the most from additional sites and it may be appropriate to shift resources from the narrow fringe stratum to the wide fringe stratum.

Introduction

The Submerged Vegetation Monitoring Project (SVMP) in the Washington State Department of Natural Resources began collecting data on the status of *Zostera marina* (eelgrass) in Puget Sound in 2000 (Berry et al. 2003). This project is part of the Puget Sound Ambient Monitoring Program (PSAMP), a multi-agency effort to monitor key aspects of the Puget Sound environment to support resource management.

At the core of the SVMP is a sophisticated statistical framework that was designed without the benefit of pre-existing data to characterize the population to be monitored (Skalski 2003). Now that the SVMP has four years of results available (2000-2003), it is an opportune time to assess whether the statistical framework is performing as intended and if refinements to the sampling and analysis procedures are warranted. This is the overall purpose of the work described in this report.

The objective of the study was to prioritize and address in depth a number of specific sampling and analysis issues that had emerged during analysis of 2002 data and in various discussions within the SVMP team. In December 2003, the initial list of questions was prepared to serve as the basis for this effort. These questions eventually ended up in the form presented in Appendix A. The first six have been addressed in this report and are listed below. In addition to producing analysis pertaining to these specific questions, an important outcome of this report was the identification of priority issues for future work.

This effort paralleled a much broader review of the SVMP, involving both internal and external reviewers. The results of this broader review will be reported separately (Sewell et al., in prep.).

Each question below is addressed in a separate chapter. In addition, other work that has been produced on an ad-hoc basis over 2003-04 to aid SVMP analysis has been included in appendices. The fifth year of SVMP sampling was completed in 2004. Some recommendations from the work reported here were implemented in the 2004 sampling.

The six questions addressed in this report were formulated and prioritized by the SVMP staff. They are questions that pertain to performance of statistical estimators and reliability of the monitoring results.

The six questions are given in the following table with their prioritized ranking and the relevant spatial scale of the question.

Rank	Basin[1] or Site-Level	Question and Approach
1	Basin	What is the relationship between number of sites sampled per season and our change detection limit at the sound-wide and regional scales? Would an increase in sites within budget constraints lead to a substantial improvement in results?
2	Basin	Is the large improvement in the coefficient of variation (CV) that resulted from adjusting the 2002 results using 2003 data statistically defensible? If this can be considered in instability in the analysis, how can the methodology be improved?
3	Basin	What are the key considerations in dividing the current flats stratum into two strata? Would the rotation effects be reduced given the loss in sample size?
4	Basin	Under our current study design, what criteria are suitable for optimizing the level of site rotation? Given these criteria, what is the optimal level of rotation?
5	Basin	How do the effects of site rotation compare to our stated detection limit? Given these effects, can we meet our detection limit objective?
6	Site-Level	Using data from the 2003 intensive sites, what relationships emerge between number of random transects and variance?

---

[1] 'Basin' denotes questions relevant to analysis at the scale of either the SVMP regions or the entire sound-wide study area.

1 Flats Stratification Issues

Results from the first three years of SVMP sampling indicate that a change to the stratification of the flat sites may be appropriate. This conclusion is based on two observations:

1. Between 2000 and 2001, the estimate of *Z. marina* area in the flats stratum increased by more than 150%, roughly an order of magnitude more than the next largest increase (+16%, wide fringe stratum). This was the primary explanation for the 43% increase overall during this same time period. An analysis by Norris et al. (2001) attributes this change to the random addition of a single anomalous site to the sample, flats11 (Samish Bay North), as part of normal sample rotation.
2. In the 2001-2003 results, all of which include flats11, variance estimates for *Z. marina* area in the flats stratum are more than an order of magnitude greater than in the other strata.

These observations suggest that flats11 is an outlier in the sample pool and that placing this site in a separate stratum would lead to more homogenous strata and lower overall variance.

Also, the SVMP team has reconsidered the current policy of dividing large flats into multiple sampling sites. This issue is separate from that of stratification but clearly must be resolved before any plan to revise the flats stratification can be developed and implemented.

This chapter has four central goals:

- Document the effects of flats11 on estimates for the flats stratum as defined for sampling in 2000-2003.
- Identify key considerations related to changing the policy of dividing large flats and make a recommendation regarding changes to this policy.
- Devise a systematic procedure for ranking flat sites according to their influence on stratum estimates and identify sites that have an inordinate influence on current estimates.
- Develop a recommendation for changes to the stratification of flat sites.

The recommendations for changes to the flats sampling frame and stratification were implemented in the 2004 sampling season based on an early draft of this report.

1.1 Flats11 Effects

Flats11 (Samish Bay North) is an outlier with respect to other sites in the 2000-03 flats sampling stratum in terms of site *Z. marina* area. This is clearly seen in Figure 1-1(a) which shows that flats11 is more than two standard deviations above the mean of *Z. marina* area estimates in the stratum. It is important to note that flats11 is not the largest site yet sampled, in terms of *Z. marina* area. As seen in Figure 1-1(b), core001 (Padilla Bay) has more than twice the *Z. marina* area of flats11, but it is not in the flats sampling stratum and is not a factor in the assessment of changes to this stratum.

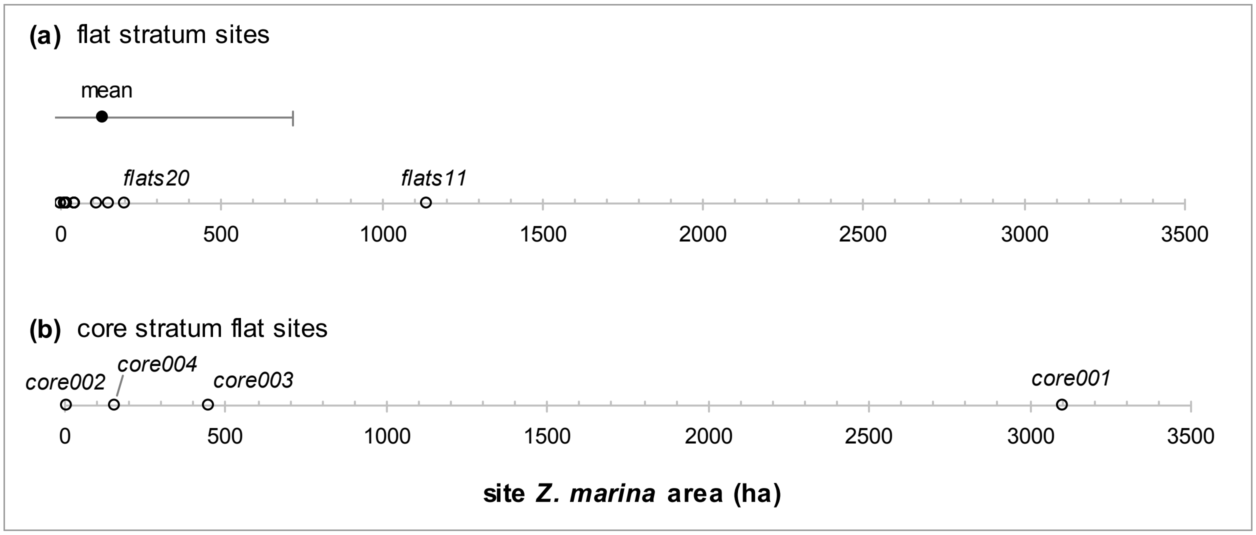


Figure 1-1. Distribution of estimates of flat site *Z. marina* area for (a) sites sampled in the 2000-03 flat sampling stratum and for (b) flat sites in the core sampling stratum. In (a), the mean of the fourteen points is also indicated with a 2× standard deviation error bar. Values shown are averages of all available 2000-2003 estimates for each site.

To assess the effect of flats11 on flats stratum results, the estimates of flats *Z. marina* area and variance were recalculated with flats11 dropped from the dataset. However, it is necessary to drop each site in turn from the calculations so that the influence of flats11 can be assessed relative to the influence of the other individual sites. This was done for each sample year, 2000-2003, resulting in 29 sets of calculations. These calculations were performed with an analysis program developed in C for this purpose. The results are summarized in Figure 1-2, Figure 1-3, Figure 1-4 and Figure 1-5 and discussed in the remainder of this section.

Flats11 was not sampled in 2000. The 2000 results in Figure 1-2 show that while some sites had more influence than others, no single site had a large impact on *Z. marina* area and all variance estimates were very low relative to the calculations for 2001-2003 shown in subsequent figures.

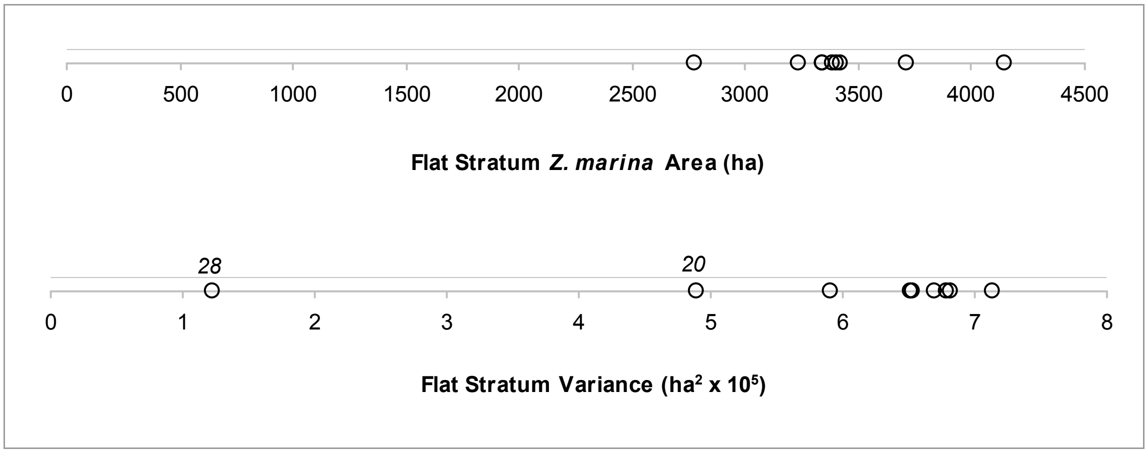


Figure 1-2. Recalculated 2000 flat stratum estimates. Each of the nine flat sites sampled in 2000, was in turn dropped from the calculation, resulting in nine calculations. Numbers above data points indicate which site was dropped for that calculation.

A consistent pattern is apparent in the 2001-2003 estimates when individual sites are dropped from the calculations (Figure 1-3, Figure 1-4, Figure 1-5). In each year, removal of flats11 results in a much lower estimate of flats stratum *Z. marina* area. However, flats11 is not the only site with an overwhelming effect on the results. Removal of flats20 also has a strong effect on the results but of an opposite sign – its removal produces a much higher estimate.

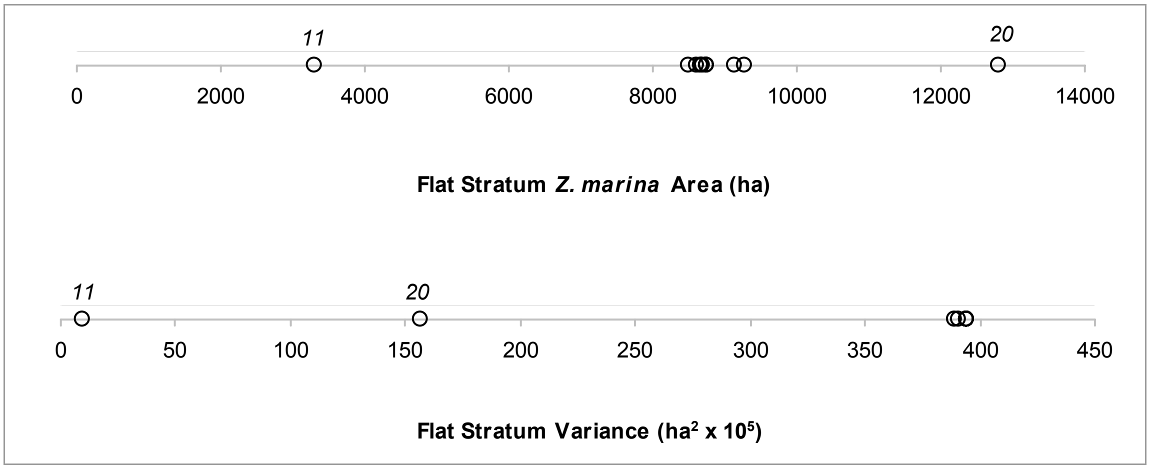


Figure 1-3. Recalculated 2001 flat stratum estimates. Each of the ten flat sites sampled in 2001, was in turn dropped from the calculation, resulting in ten calculations. Numbers above data points indicate which site was dropped for that calculation.

The influence of a particular site can be quantified as the difference between the result when that site is withheld from the calculation and the average of *Z. marina* area estimates as other sites are withheld. By this measure, flats11 has a stronger influence on the results than flats20, but only marginally so in 2001.

The variance estimates calculated with sites withheld support this assessment. In each year of 2001-2003 (Figure 1-3, Figure 1-4, Figure 1-5), removal of either flats11 or flats20 from the calculation dramatically reduces the estimate of flats stratum variance. It is clear in each year, however, that removal of flats11 has a much stronger effect on the variance.

These results support the conclusion of Norris et al. (2001) that flats11 is unique within the flats stratum. Removal of this site from the stratum would dramatically reduce the variance estimates. However, these results also show that flats20 also has an inordinately strong but opposite influence on stratum estimates. This is an important consideration for the development of objective criteria for implementing changes to the stratification of the flats, which is discussed in section 1.3.5 (p.34).

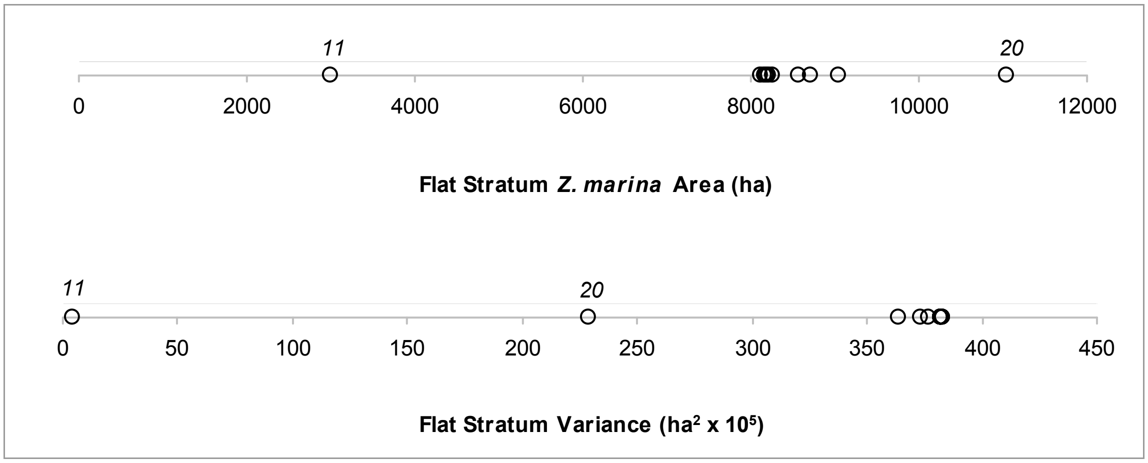


Figure 1-4. Recalculated 2002 flat stratum estimates. Each of the ten flat sites sampled in 2002, was in turn dropped from the calculation, resulting in ten calculations. Numbers above data points indicate which site was dropped for that calculation.

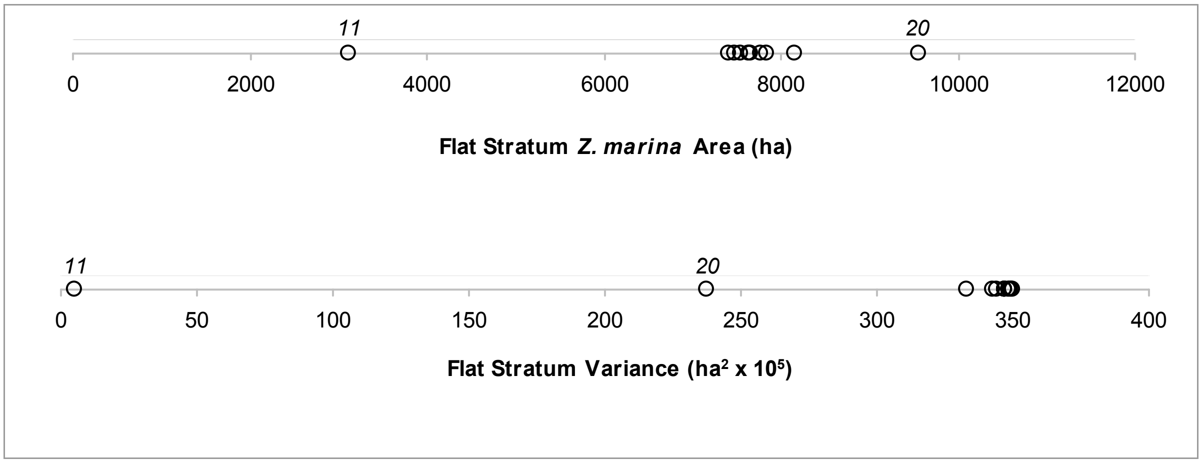


Figure 1-5. Recalculated 2003 flat stratum estimates. Each of the ten flat sites sampled in 2003, was in turn dropped from the calculation, resulting in ten calculations. Numbers above data points indicate which site was dropped for that calculation.

1.2 Revisiting the Division of Flats

This section first reviews the issues surrounding potential changes to the existing SVMP policy of dividing large geomorphic flats into multiple sampling sites. Then the distribution of flat areas is presented for both intact geomorphic flats and the 2000-03 flat sampling frame that includes subdivision. Finally, a recommendation is made regarding changes to the subdivision policy and the underlying sampling frame.

1.2.1 Issues

There are two conflicting principles that are relevant to this discussion:

- Geomorphic flats *may* function as ecological units that should be studied and sampled as a whole.
- Large flat sites should be divided to make sampling more manageable and to eliminate extreme discrepancies in sampling rates across sites (i.e. ratio of total transect length to sample polygon area).

Large sites were initially divided in the sampling design to ensure that each flat site could be sampled in a single day. Following 2000, an exception was made for Padilla Bay (core001) so that the site boundary coincided with the National Estuary Research Reserve (NERR) boundary. Padilla Bay currently takes approximately 1.5 days to sample.

Dividing large flat sites is problematic for the following reasons.

1. Whole geomorphic flat sites (as delineated by 0 and –20ft [-6.1m] bathymetric contours) may be more appropriate sampling units because of characteristic patterns of within-site heterogeneity and interdependent ecosystem functions.

   With respect to heterogeneity, due to bathymetry and locations of freshwater inputs, it may be impossible to divide some flat sites into comparable subunits without resorting to tortuous boundaries that are not feasible from a sampling standpoint. Hence, some subunits may be potential *Z. marina* habitat while others may not. Figure 1-6 shows an example of a site that was flagged for division into three subunits in the 2000-03 sampling frame but does not easily divide into subunits with comparable *Z. marina* habitat characteristics.

   Division of large flats into disparate subunits does not violate any theoretical sampling requirements but it would tend to increase variance estimates for the stratum.

2. There are working SVMP guidelines for identifying which flats should be divided and how the subdivisions are to be delineated, but these are not precisely defined and do not seem to have been applied consistently.

   The division of large flats directly affects the distribution of sampling site areas. The general guideline has been to divide large flats so that each subunit could be sampled in a single day. There is no definitive size threshold associated with this guideline. Similarly the guideline for delineating subdivisions was that boundaries should generally follow the depth gradient, but in some instances the bathymetry is such that this guideline does not provide clear boundaries.

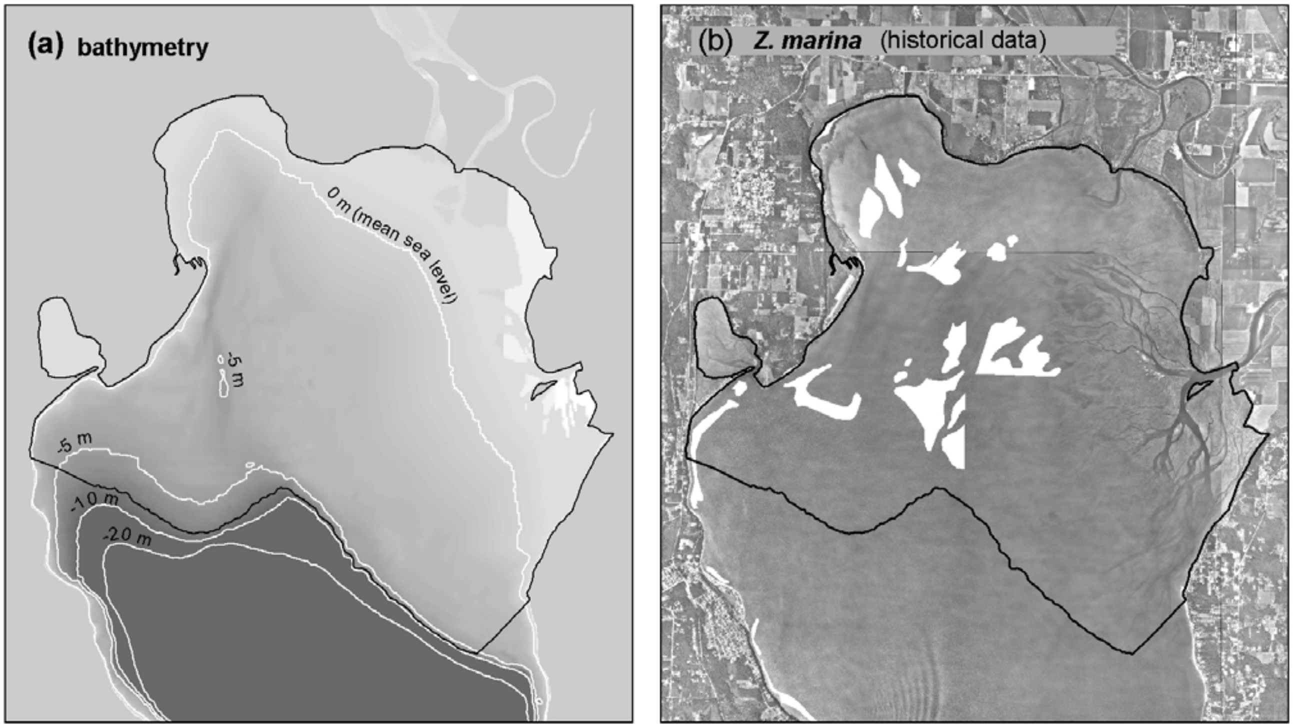


Figure 1-6. Port Susan embayment as delineated by the SVMP with (a) bathymetry (Finlayson et al., 2000) and (b) orthophotos and historical (1970s) *Z. marina* distribution (Uhrich and McGrath 1997). This is an example of a site that is not easily divided into comparable subunits. This site was originally designated for division into three subunits (flats22, flats23 and flats24) but had not yet been divided as preparations began for the 2004 field season. The sharp, linear boundaries in the *Z. marina* beds seen at the lower right of (b) appear to be an artifact of the generation of GIS data (Uhrich and McGrath 1997; Needham and Lanzer 1993) from the original *Zostera* spp. data layer in the Puget Sound Environmental Atlas (Evans-Hamilton Inc. and D.R. Systems, Inc. 1987), in turn based on the Coastal Zone Atlas of Washington (Washington State Department of Ecology 1979).

On the other hand, transitioning to the sampling of whole geomorphic flats is problematic for the following reasons.

1. The largest sites may be so dominant that they need to be shifted to a second flats stratum that is sampled every year (effectively increasing the number of core sites).

2. Results would not be directly comparable to existing 2000-2003 results thereby diminishing the value of three years of effort and prolonging the period until SVMP produces meaningful multiyear trend analyses.

3. There could be dramatically different levels of sampling density among sites. Here sampling density is understood to be total sampling transect length per sample polygon area.

1.2.2 Flat Area Distributions

The flats stratum as sampled in 2000-03 contains 67 flats (not including flats in the core stratum). Over 2000-2003, a total of 14 of these were sampled and seven of those were subunits of large geomorphic flats.

The distribution of flats area is shown below for both intact geomorphic (undivided) flat sites in the flats stratum (Figure 1-7 and Table 1-1) and flat sampling sites as delineated in the 2000-03 sampling frame (Figure 1-8). The distribution of geomorphic flat areas is very skewed with a few very large sites, especially Skagit Bay (Figure 1-7). The 2000-03 sampling frame, in which large geomorphic flats have been subdivided, has virtually the same level of skewness (Figure 1-8) although the range in areas has been compressed. The two largest flat sites (flats21 and flats20) are two of the subdivisions of Skagit Bay.

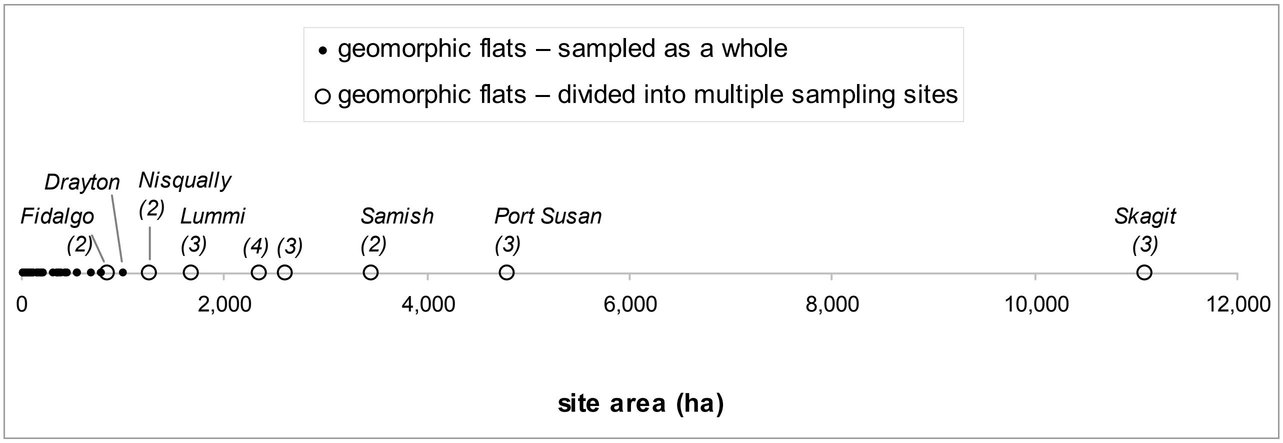


Figure 1-7. Distribution of area in geomorphic (undivided) flats within the flats stratum. Geomorphic flats flagged for subdivision into smaller sampling units (sites) in the 2000-03 sampling frame are shown as open circles. The number of subunits is shown in parentheses. Values are shown in Table 1-1. Names refer to embayments (e.g. Samish Bay) or prominent local features (e.g. Nisqually river delta).

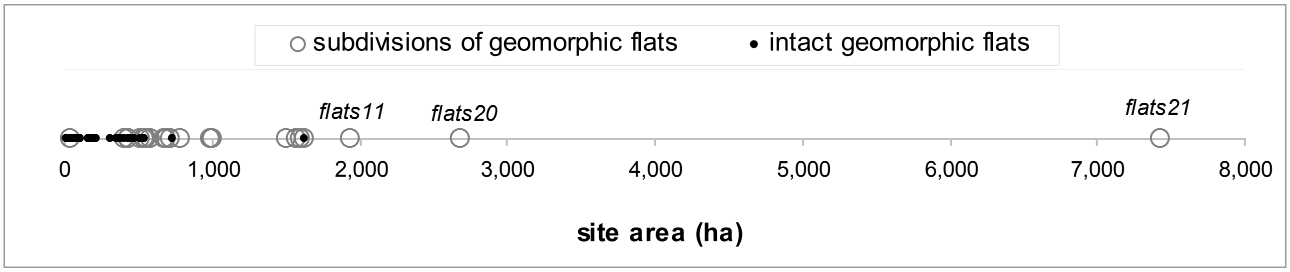


Figure 1-8. Distribution of flat stratum site areas in the 2000-03 sampling frame. Flat sites that are subdivisions of larger geomorphic flats are shown with open circles. In the three cases where a geomorphic flat was flagged for subdivision but had not yet been divided in the 2000-03 sampling frame, the overall area was divided into nearly equal units for the purpose of this figure only.

Note that the number of subdivisions varies by site and does not clearly follow site area (Figure 1-7), e.g., the largest site has three subdivisions while a smaller site has four. Also, a threshold for determining which sites are to be divided does not appear to be uniformly applied since Fidalgo Bay was flagged for division but a larger site, Drayton Harbor, was not.

site codes	flat site names	number of sampling sites per geomorphic flat	region	geomorphic flat area (ha)
flats19 20 21	Pull and Be Damned Point & Skagit Bay N S	3	swh	11,100
flats22	Port Susan W M E	3	swh	4,793
flats11 & 12	Samish Bay N S	2	nps	3,454
flats26, 27, 28	Snohomish Delta N M S	3	swh	2,614
flats07,08,09,10	Portage Bay N S & Nooksack Delta E W	4	nps	2,351
flats04	Lummi Flats N M S	3	nps	1,679
flats34 & 35	Nisqually Delta E W	2	cps	1,269
flats01	Drayton Harbor	1	nps	1,010
flats15	Fidalgo Bay N S	2	nps	841
flats03	Birch Bay	1	nps	781
flats14	March Pt.	1	nps	684
flats50	Dungeness Bay	1	fuc	546
flats18	Similk Bay	1	swh	544
flats30	Cultus Bay	1	cps	445
flats32	Dugualla Bay	1	swh	429
flats49	Old Town	1	fuc	404
flats42	Quilcene Bay	1	hdc	385
flats31	Oak Harbor	1	swh	363
flats39	Liberty Bay	1	cps	357
flats41	Dosewallips	1	hdc	316
flats59	Mud Bay, Lopez	1	sji	221
flats33	Quartermaster Harbor	1	cps	200
flats46	Kilisut Harbor	1	cps	200
flats47	Travis Spit	1	fuc	193
flats57	Fisherman's Bay, Lopez	1	sji	171
flats25	Tulalip Bay	1	swh	166
flats43	Dabob Bay	1	hdc	163
flats53	Westcott Bay, San Juan	1	sji	159
flats40	Miller Bay	1	cps	111
flats56	False Bay, San Juan	1	sji	111
flats45	Hood Head	1	hdc	101
flats62	Swifts Bay, Lopez	1	sji	96
flats38	Port Madison	1	cps	84
flats61	Shoal Bay, Lopez	1	sji	83
flats54	Garrison Bay, San Juan	1	sji	81
flats48	Sequim Bay	1	fuc	79
flats58	Barlow Bay, Lopez	1	sji	76
flats63	Blind Bay, Shaw	1	sji	74
flats52	Nelson Bay, Henry Island	1	sji	69
flats55	Mitchell Bay, San Juan	1	sji	67
flats60	Hunter Bay, Lopez	1	sji	66
flats29	Coronet Bay	1	swh	56
flats44	Case Shoal	1	hdc	55
flats37	Wing Point	1	cps	49
flats36	Eagle Harbor	1	cps	43
flats17	Bowman Bay	1	sji	37
flats67	Fossil Bay, Sucia	1	sji	36
flats68	Secret Harbor, Cypress	1	sji	30
flats51	Provost Harbor	1	sji	29
flats65	Thatcher Bay, Decater	1	sji	28
flats66	Shallow Bay, Sucia	1	sji	23
flats64	Squaw Bay	1	sji	18
flats69	Eagle Cove, Cypress	1	sji	13

Table 1-1. Geomorphic flats that encompass the 2000-03 flat stratum and designations for subdivision of these flats into multiple sampling sites in the 2000-03 sampling frame. In three cases, even though flats were designated for subdivision, this was never completed in the 2000-03 sampling frame (Fidalgo Bay, Lummi Bay, Port Susan).

The distribution of flat site areas (Figure 1-8) is used here as the basis for evaluating the flats sampling frame (i.e. the delineation of sampling units, or sites). The distribution of site *Z. marina* area (presented in section 1.3.1) could also be used for this purpose, but it is used here only for stratification of the flat sites. This choice is explained by the following rationale.

Z. marina area is a variable over the time scale of the monitoring program. Indeed, measurement of this variable is the primary monitoring objective. In contrast, the geomorphic controls on flat areas can generally be assumed to be relatively more stable. This is not necessarily true at all sites in all instances–for instance, flood events with high levels of sediment transport or tectonic events may alter embayment bathymetry and the area of potential *Z. marina* habitat–but will generally be true. Ideally, the delineation of the sampling frame is fixed for the duration of the monitoring program. In contrast, a given sampling site can be moved between different strata as conditions change.

The question of flats subdivision directly impacts the sampling frame. Since the sampling frame should remain fixed, it is better to base flats subdivision on fixed flats areas rather than variable *Z. marina* areas. The question of flats stratification is more flexible and, in addition to other reasons discussed in section 1.3, is more appropriately based on *Z. marina* area.

1.2.3 Recommendations on Flat Division

The recommendations given here pertain to the 2000-03 flats sampling frame. These recommendations were implemented as part of the preparation for the 2004 sampling season. The result is denoted here as the 2004 sampling frame.

The SVMP should maintain the current policy of subdividing large geomorphic flats with minor revision. It is not recommended to fundamentally change the policy and redesign the sampling frame to contain intact geomorphic flats.

The rationale for this recommendation is as follows. The primary trade-off here is between maintaining the ecological integrity of sampling units versus retaining as much value as possible in the results from the first four years of SVMP effort (the latter issue is a more compelling argument for status quo than the practicality of sampling).

There certainly may be situations where weaknesses in the sampling frame justify a complete redesign with concomitant break in the data record. In this case, however, the argument of ecological integrity is insufficient and may not even be applicable.

The SVMP as it currently exists is monitoring pattern and not function, although pattern and function are clearly interdependent. In this context, any supposed integrity of ecological pattern in geomorphic flat areas is more relevant than the supposed integrity of a suite of ecological functions that may be present. If flats had a consistent pattern with respect to within-flat *Z. marina* distribution, then this would support the sampling of flats as whole units. This would ensure the sampling of the characteristic pattern within flats and not just a component of the pattern. A hypothetical pattern might include: an upper delta area near freshwater input devoid of *Z. marina*, a mid-depth area with *Z. marina* and a deep area with no *Z. marina*.

However, a cursory review of a number of geomorphic flats suggests that there is no consistent characteristic pattern and that the pattern in a given flat is more a reflection of site-specific controls that vary from site to site. Without additional information to guide the sampling frame design that would capture these site-specific patterns, the SVMP must rely on the somewhat rough delineation provided by mean high water (DNR's water level line) and the -20ft (-6.1m) contour (from WDFW/Dale Gombert).

In the absence of consistent patterns across flats that would suggest flats are coherent units, and in the absence of data to delineate zones within flats, it is more accurate to say that the SVMP is sampling flats area that happens to be organized into discrete geomorphic units rather than sampling geomorphic flats that have varying amounts of flat area.

This recommendation to maintain the existing policy of dividing large flats, of course, must be revisited if new opposing arguments are presented or existing arguments are further articulated.

Four recommendations are made for minor refinements to the existing division of flat sites:

1. Fidalgo Bay should not be subdivided because it is smaller than Drayton Harbor–a site that was not flagged for subdivision in the 2000-03 sampling frame. Fidalgo Bay was designated for subdivision in the 2000-03 sampling frame into flats15 and flats16. Fidalgo Bay has not yet been sampled so this change will not directly affect existing data. However, the sample pool used for random site rotation will differ from previous years.
2. Skagit Bay south (flats21 of the 2000-03 frame) should be further divided into three sampling sites since this site is over twice as big as the next biggest site. Fortunately, the shape and orientation of this site make further subdivision rather straightforward by following the depth gradient (Figure 1-9).
3. Port Susan should be subdivided into two sampling sites (flats22 and flats23) rather than three as in the 2000-03 sampling frame. This site is difficult to divide using sensible geomorphological criteria and is not large enough to require subdivision into three sites (Figure 1-10)
4. Lummi flats should be subdivided into two sites (flats04 and flats05) rather than three as in the 2000-03 sampling frame. Lummi flats is at the lower end of the large flats that are subdivided but there is one smaller flat that has already been subdivided and sampled (Nisqually). For consistency, since Lummi is larger than Nisqually it should be subdivided.

The proposed new divisions for Skagit Bay, Port Susan and Lummi flats are shown in Figure 1-9, Figure 1-10 and Figure 1-11. The revised flats sampling frame that incorporates these recommendations includes 70 sites, including flat sites in the core stratum (Table 1-2).

	number of flat sites in core stratum	number of flat sites outside core stratum	total number of flat sites
2000-03 flat sampling frame	4	67	71
2004 flat sampling frame	4	66	70

Table 1-2. Comparison of 2000-03 and 2004 flat sampling frames.

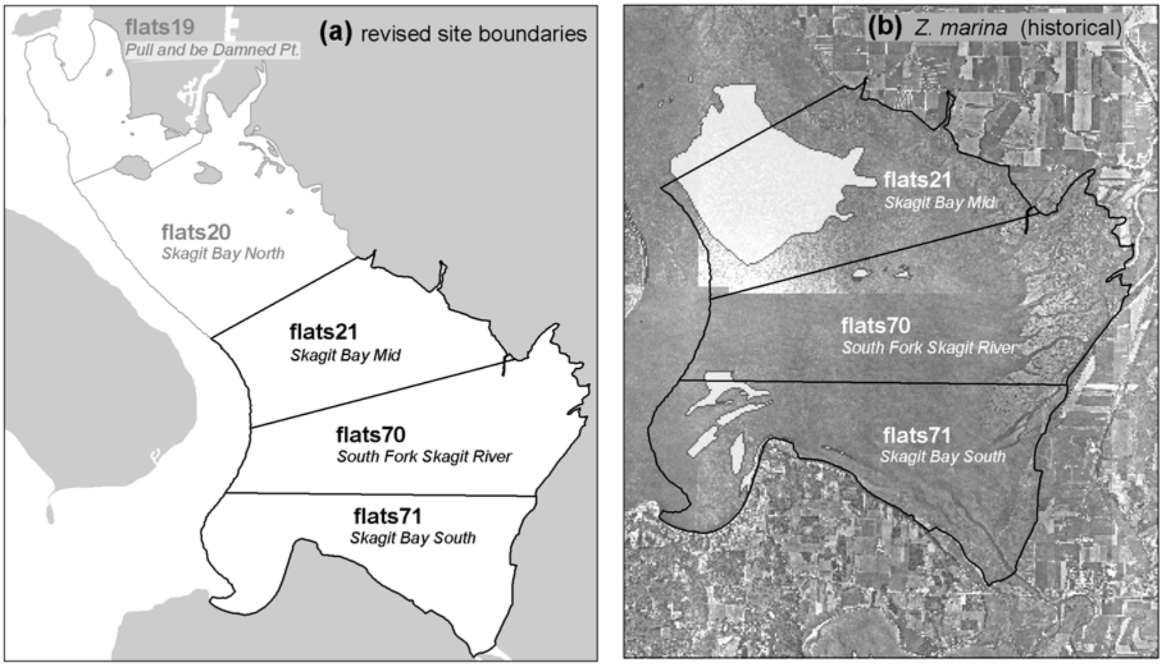


Figure 1-9. Revised subdivision of the flats21 as delineated for 2000-03 sampling into the smaller flats21 shown, flats70 and flats71 (a). Before subdivision, this was the largest site in the flats stratum. It was more than twice as large as the next largest site, flats20, Skagit Bay north. *Z. marina* distribution (b) depicts status in 1970s (Uhrich and McGrath 1997).

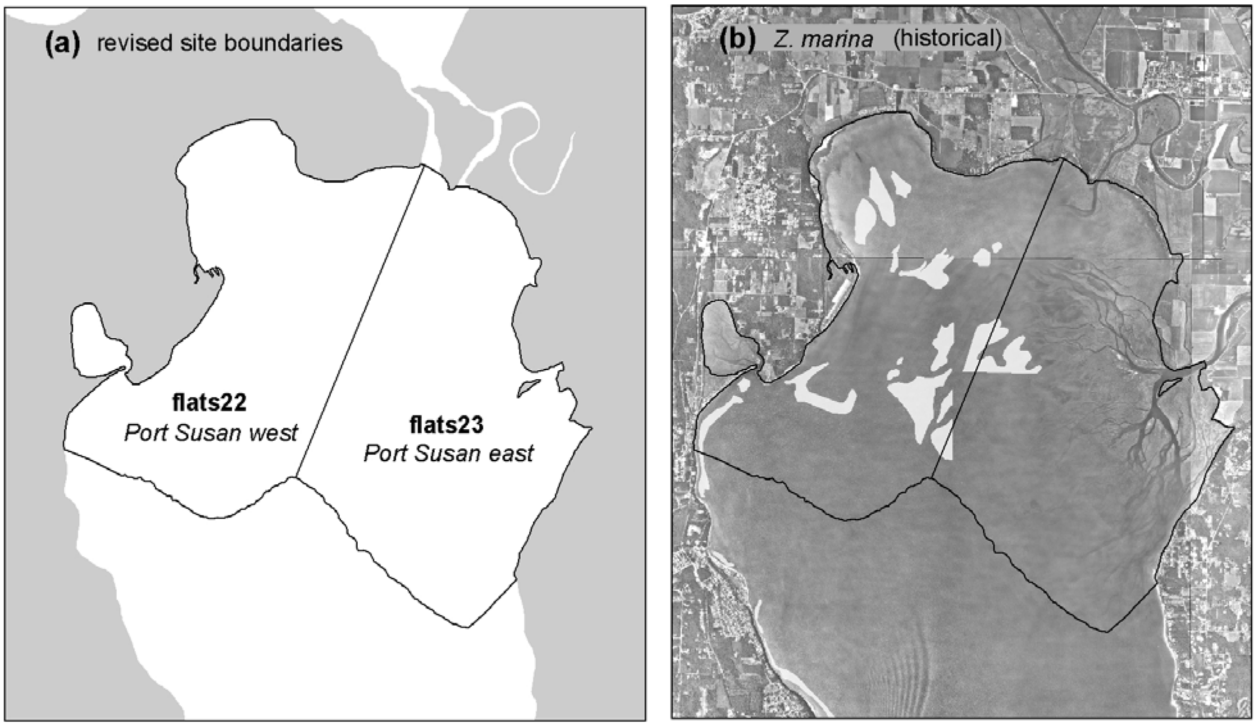


Figure 1-10. Revised subdivision of Port Susan into flats22 and flats23 (a) and historical (1970s) *Z. marina* distribution (Uhrich and McGrath 1998).

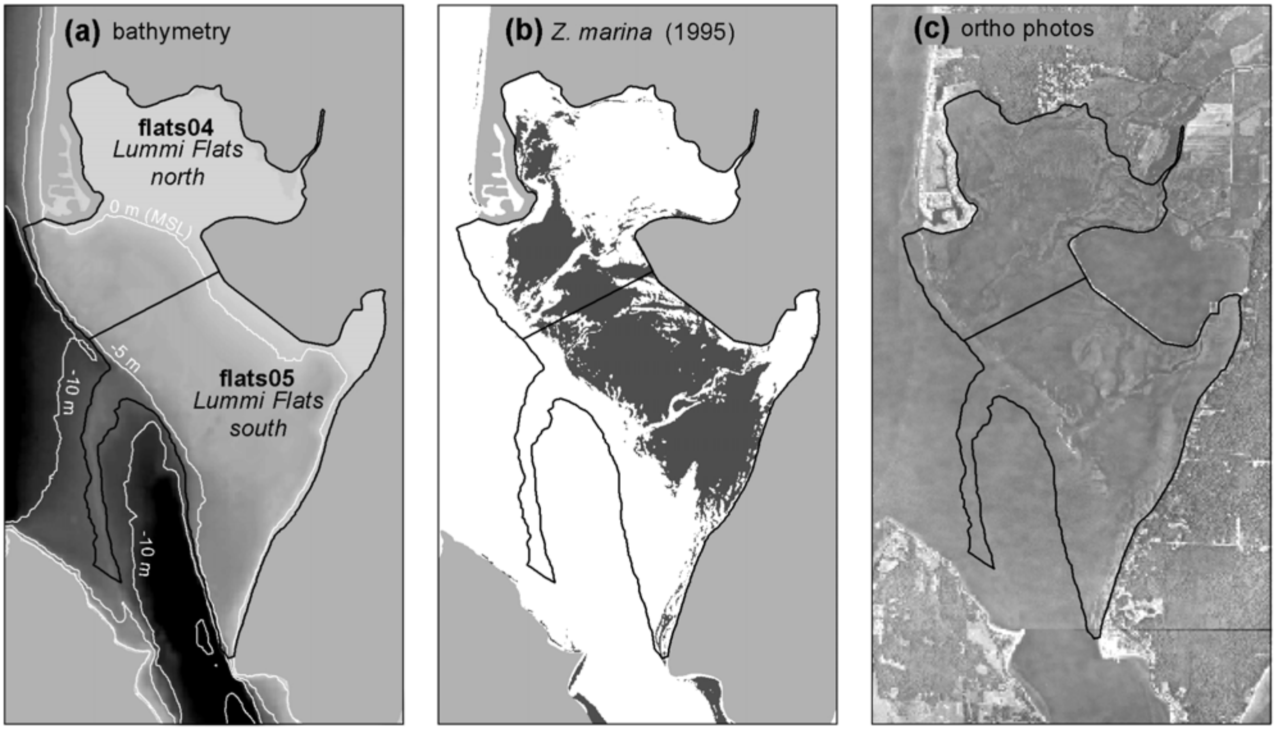


Figure 1-11. Revised subdivision of Lummi flats into flats04 and flats05 with (a) bathymetry (Finlayson et al. 2000), (b) *Z. marina* distribution (Berry and Ritter 1997) and (c) orthophotos. Lummi flats was originally designated for division into three sites.

1.3 Development of an Alternate Flats Stratification

It is clear from section 1.1 that flats11 is an outlier and that moving this site to another stratum would improve the precision of flats stratum results. However, it is important to devise an objective procedure for changing the stratification so that this is not a haphazard process.

A model flats dataset is presented in this section that encompasses all flat sites and is intended to serve as a basis for assessing stratification. This model dataset is useful because it allows us to complete quantitative analyses that could not be completed with the limited existing data (only 14 flat stratum sites and four flats in the core stratum have been sampled through 2003). Specifically, the model dataset allows us to evaluate influence of each flat site on the overall stratum estimates. This will systematically identify highly influential sites that are candidates for transfer to a separate stratum in order to improve precision.

While the model flats dataset serves a very useful purpose and was designed to be consistent with patterns seen in existing SVMP data, it is impossible to know its overall accuracy. It is therefore not useful and is not intended to be used for direct estimates of stratum-level *Z. marina* area or variance. It represents a dataset that statistically resembles the 2000-2003 flats data but has the advantage of being complete and can be used to drive studies that simulate the effects of the SVMP sampling and analysis.

As new flat sites are sampled each year and the preliminary estimates in the model flats dataset are revised, the data may suggest that there are other unanticipated outliers. The influence indices presented in this section provide a mechanism to determine if these sites should be shifted out of the main flats stratum. Ideally, the stratification of flats will remain fixed once the changes recommended later in this section are made. Future changes should only be made for very compelling reasons.

The recommendations for changes in flats stratification presented here emerged from an early draft of this report and discussions with J.R. Skalski. They were first implemented in the 2004 field season.

1.3.1 <u>Model Flats Dataset–Point Estimates of *Z. marina* Area</u>

The *Z. marina* area values in the model flats dataset are based on 2000-2003 data, site areas and estimates of *Z. marina* area ratio for unsampled sites derived from independent data and expert opinion.

Z. marina area has been sampled at only 14 of the 66 sites in the 2000-03 flats stratum over the first four years of SVMP sampling. For these 14 sites, the most recent *Z. marina* area estimates were used in the model flats dataset.

Rudimentary estimates of *Z. marina* area were made for the remaining 52 unsampled sites. The approach was to assign each unsampled site larger than 425 ha to a category of high, moderate or low *Z. marina* area ratio, *R*, the fractional cover of *Z. marina* at a site. These categories were characterized by reference to existing SVMP data. Figure 1-12 shows the distribution of *R* in 2000-2003 data and the high, moderate and low *R* categories that were manually delineated. Nominal *R* values for these categories are 0.10 (low), 0.30 (moderate) and 0.55 (high). These are the average category values rounded to the nearest 0.05. The four flat sites in the core stratum were included in the determination of these nominal values in order to increase the number of data points and better describe the distribution. This is considered legitimate since the criteria for isolating these sites in the core stratum were not based on *Z. marina* area ratio. Consequently, there is no reason to expect them to differ as a group from the main flats stratum in this respect, i.e. they are part of the same *R* distribution.

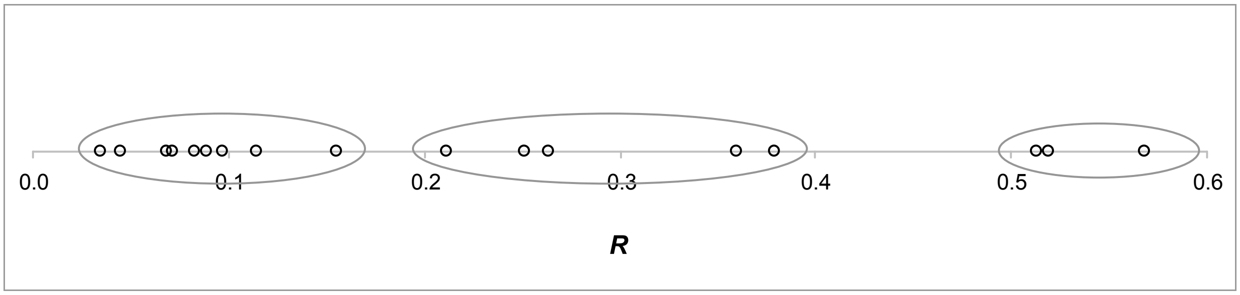


Figure 1-12. The distribution of *Z. marina* area ratio, *R*, for flat sites sampled over 2000-2003, including core sites. The ellipses indicate the manual grouping used to obtain nominal *R* values for the high (0.55), moderate (0.3) and low (0.1) *R* categories. Most recent values are shown for sites sampled in multiple years.

The unsampled flat sites larger than 425 ha were assigned to an *R* category based on the collective field experience of the DNR Aquatics Division science support staff[2] and examination of bathymetry and *Z. marina* distributions from independent sources. These other sources include the Puget Sound Environmental Atlas (Uhrich and McGrath 1997), the 1995 and 1996 Intertidal Habitat Inventories (Berry and Ritter 1997; Ritter et al. 1999) and Skagit System Cooperative results from 2001 for Skagit Bay.

Unsampled sites smaller than 425 ha were simply assigned to the moderate *R* category. The purpose of this cutoff was to focus effort only on those sites that are large enough to have potential to be influential sites. Figure 1-13 shows the location of the 425 ha cutoff relative to the distribution of flat site areas. Only 44% of the sites are above the 425 ha cutoff, but these comprise 76% of the area in the flat sites.

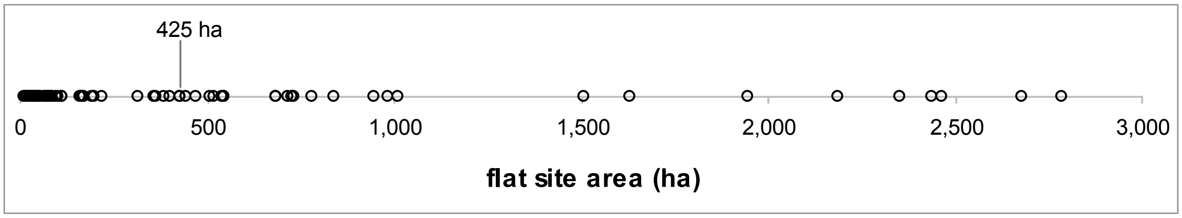


Figure 1-13. Location of the 425 ha threshold relative to the distribution of flat site areas of all flat sites, excluding core sites, in the 2004 flat sampling frame.

---

[2] Amy Sewell and Betty Bookheim led this exercise.

Site *Z. marina* area was calculated as the product of site area and the *R* value, whether based on 2000-2003 data or the nominal value for the assigned *R* category. The results of this exercise are shown in Table 1-3 and Figure 1-14.

Flats number	Flats Name	region	Area (ha)	*R* (sampled)	*R* category	*R*	ZM Area (ha)
flats11	Samish Bay N	nps	1,944	0.568		0.57	1104.1
flats70	S. Fork Skagit River	swh	2,786		moderate	0.30	835.8
flats12	Samish Bay S	nps	1,509		high	0.55	830.1
flats21	Skagit Bay mid	swh	2,186		moderate	0.30	655.7
flats01	Drayton Harbor	nps	1,010		high	0.55	555.4
flats05	Lummi Flats S	nps	948		high	0.55	521.5
flats15	Fidalgo Bay	nps	841		high	0.55	462.5
flats14	March Pt.	nps	684		high	0.55	376.0
flats07	Portage Bay S	nps	517		high	0.55	284.4
flats71	Skagit Bay south	swh	2,466		low	0.10	246.6
flats23	Port Susan E	swh	2,441		low	0.10	244.1
flats22	Port Susan W	swh	2,352		low	0.10	235.2
flats20	Skagit Bay north	swh	2,680	0.083		0.08	221.9
flats04	Lummi Flats N	nps	731		moderate	0.30	219.3
flats50	Dungeness Bay	fuc	546		moderate	0.30	163.7
flats27	Snohomish Delta mid	swh	1,630		low	0.10	163.0
flats19	Pull and Be Damned Point	swh	983	0.155		0.15	152.4
flats32	Dugualla Bay	swh	429		moderate	0.30	128.6
flats49	Old Town	fuc	404		moderate	0.30	121.1
flats42	Quilcene Bay	hdc	385		moderate	0.30	115.4
flats31	Oak Harbor	swh	363		moderate	0.30	109.0
flats39	Liberty Bay	cps	357		moderate	0.30	107.2
flats28	Snohomish Delta south	swh	473	0.212		0.21	100.2
flats41	Dosewallips	hdc	316		moderate	0.30	94.9
flats03	Birch Bay	nps	781		low	0.10	78.1
flats34	Nisqually Delta W	cps	727		low	0.10	72.7
flats09	Nooksack Delta W	nps	687		low	0.10	68.7
flats59	Mud Bay, Lopez	sji	221		moderate	0.30	66.3
flats33	Quartermaster Harbor	cps	200		moderate	0.30	60.1
flats46	Kilisut Harbor	cps	200		moderate	0.30	60.0
flats57	Fisherman's Bay, Lopez	sji	171		moderate	0.30	51.4
flats26	Snohomish Delta N	swh	511		low	0.10	51.1
flats25	Tulalip Bay	swh	166		moderate	0.30	49.8
flats47	Travis Spit	fuc	193	0.252		0.25	48.6
flats30	Cultus Bay	cps	445		low	0.10	44.5
flats08	Portage Bay N	nps	429	0.097		0.10	41.8
flats18	Similk Bay	swh	544	0.072		0.07	39.0
flats40	Miller Bay	cps	111		moderate	0.30	33.2
flats56	False Bay, San Juan	sji	111		moderate	0.30	33.2
flats45	Hood Head	hdc	101		moderate	0.30	30.3
flats38	Port Madison	cps	84		moderate	0.30	25.3
flats61	Shoal Bay, Lopez	sji	83		moderate	0.30	24.9
flats35	Nisqually Delta E	cps	543	0.045		0.04	24.3
flats54	Garrison Bay, San Juan	sji	81		moderate	0.30	24.2
flats48	Sequim Bay	fuc	79		moderate	0.30	23.6
flats58	Barlow Bay, Lopez	sji	76		moderate	0.30	22.8
flats63	Blind Bay, Shaw	sji	74		moderate	0.30	22.3
flats52	Nelson Bay, Henry Island	sji	69		moderate	0.30	20.7
flats55	Mitchell Bay, San Juan	sji	67		moderate	0.30	20.1
flats37	Wing Point	cps	49	0.360		0.36	17.6
flats29	Coronet Bay	swh	56		moderate	0.30	16.8
flats44	Case Shoal	hdc	55		moderate	0.30	16.4
flats53	Westcott Bay, San Juan	sji	159	0.089		0.09	14.1
flats36	Eagle Harbor	cps	43		moderate	0.30	12.9
flats17	Bowman Bay	sji	37		moderate	0.30	11.2
flats43	Dabob Bay	hdc	163	0.068		0.07	11.1
flats62	Swifts Bay, Lopez	sji	96	0.114		0.11	11.0
flats67	Fossil Bay, Sucia	sji	36		moderate	0.30	10.7
flats68	Secret Harbor, Cypress	sji	30		moderate	0.30	8.9
flats65	Thatcher Bay, Decater	sji	28		moderate	0.30	8.4
flats66	Shallow Bay, Sucia	sji	23		moderate	0.30	7.0
flats64	Squaw Bay	sji	18		moderate	0.30	5.3
flats69	Eagle Cove, Cypress	sji	13		moderate	0.30	3.8
flats51	Provost Harbor	sji	29		low	0.10	2.9
flats60	Hunter Bay, Lopez	sji	66	0.035		0.03	2.3
flats10	Nooksack Delta E	nps	719	0.000		0.00	0.0

Table 1-3. *Z. marina* area ratio, *R*, and *Z. marina* area for each site in the 2004 flats sampling frame, excluding flat sites in the core stratum. The *R* values were either determined from existing SVMP data ("*R* sampled"), or assigned *R* category. Large unsampled sites were assigned to an *R* category based on examination of existing inventory data and expert opinion. Small unsampled sites were simply assigned to the moderate category.

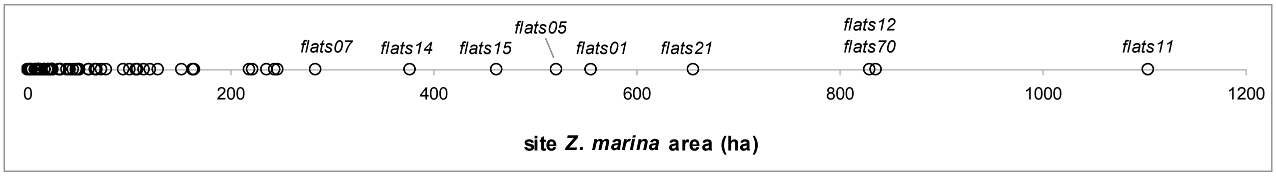


Figure 1-14. Distribution of estimated *Z. marina* area in the model flats dataset. This includes all flat sites in the 2004 flats sampling frame, excluding flat sites in the core stratum.

1.3.2 Model Flats Dataset–Variance Estimates

The wide range in site-level variance in the 2000-2003 SVMP results reflect both a range in relative precision as well as the wide range in *Z. marina* area across sites. The distribution of coefficient of variation (CV) in the 2000-2003 data was used as the basis for generating CV values for unsampled sites. Variance values in the model flats dataset were then derived from these site CV values. This approach ensured that the distribution of variance in the model flats dataset resembled the actual distribution of precision in the 2000-03 SVMP data.

Characteristics of CV Values in the 2000-2003 Data

Figure 1-15 shows the distribution of site-level CV for all 18 flat sites sampled during 2000-2003, including the four flats in the core stratum. This distribution appears to be fairly uniform between values of 0.038 and 0.174 if flats10 is excluded where no *Z. marina* has been observed (CV=0).

However there does seem to be some structure to the pattern of site-level CV when viewed relative to *Z. marina* area and *R*, the *Z. marina* area ratio (Figure 1-16). The relationship with *Z. marina* area appears to be strongly non-linear while the relationship with *R* appears more linear although with one outlier (flats37).

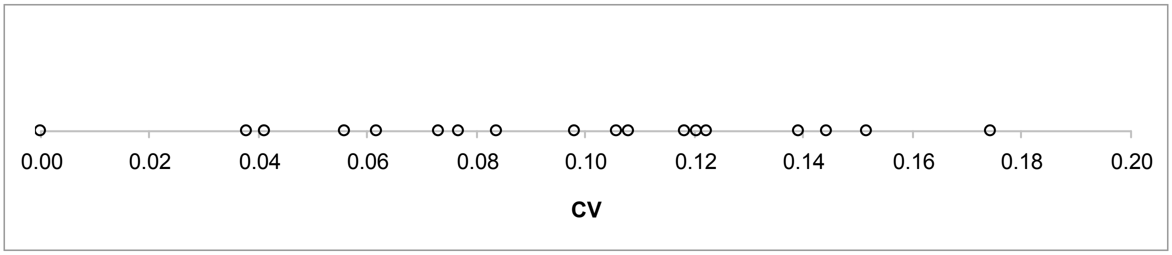


Figure 1-15. Distribution of site-level CV for all flat sites sampled during 2000-2003 including the four flat sites in the core stratum. The most recent estimates are shown for each site.

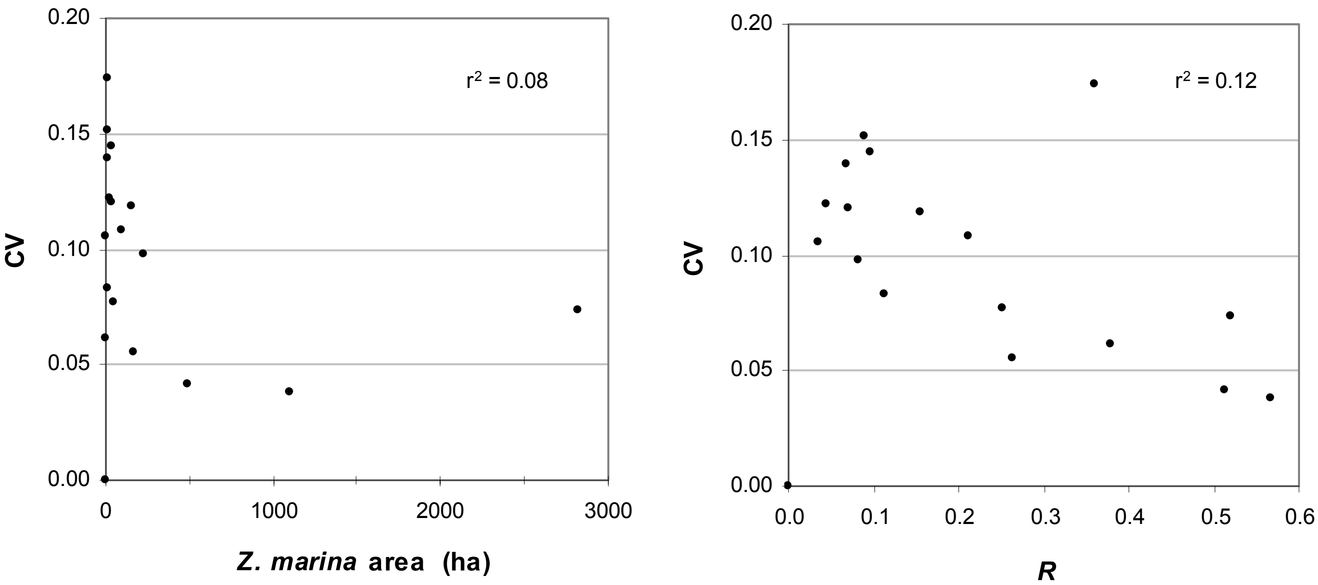


Figure 1-16. Relationship of site-level CV to *Z. marina* area and *R* for all flat sites sampled during 2000-2003. This includes the four flat sites in the core stratum. The most recent estimates are shown for each site.

Figure 1-17 shows the inverse curve that was developed to represent the relationship between flat site CV and *Z. marina* area in existing 2000-2003 data for the purposes of generating the model flats dataset. This curve has the general form

$$y = \frac{k}{x + x_0} + y_0$$

Equation 1-1

where the parameters have the values k=5 ha, x_0=40 ha and y_0=0.05 in the curve shown in Figure 1-17. This curve was used to generate CV values for unsampled flat sites.

The mean square error, MSE_{CV}, was calculated from the sum of squares, SS_{CV}, of 17 of the 18 flat sites (including core stratum flat sites) sampled during 2000-2003. Flats10 was withheld from the calculation of SS_{CV}, and MSE_{CV} since it does not have *Z. marina* present and is therefore an outlier.

MSE_{CV} was calculated as

$$MSE_{CV} = \frac{SS_{CV}}{n-3} = \frac{\sum_{i=1}^{n}\left(CV_{obs} - CV_{pred}\right)^2}{n-3}$$

Equation 1-2

where

$n = 17$

CV_{obs} = CV of 17 flats sites sampled in 2000-2003 (flats10 withheld)

CV_{pred} = CV predicted by Equation 1-1.

The resulting values are

$SS_{CV} = 0.02720,$
$MSE_{CV} = 0.001943.$

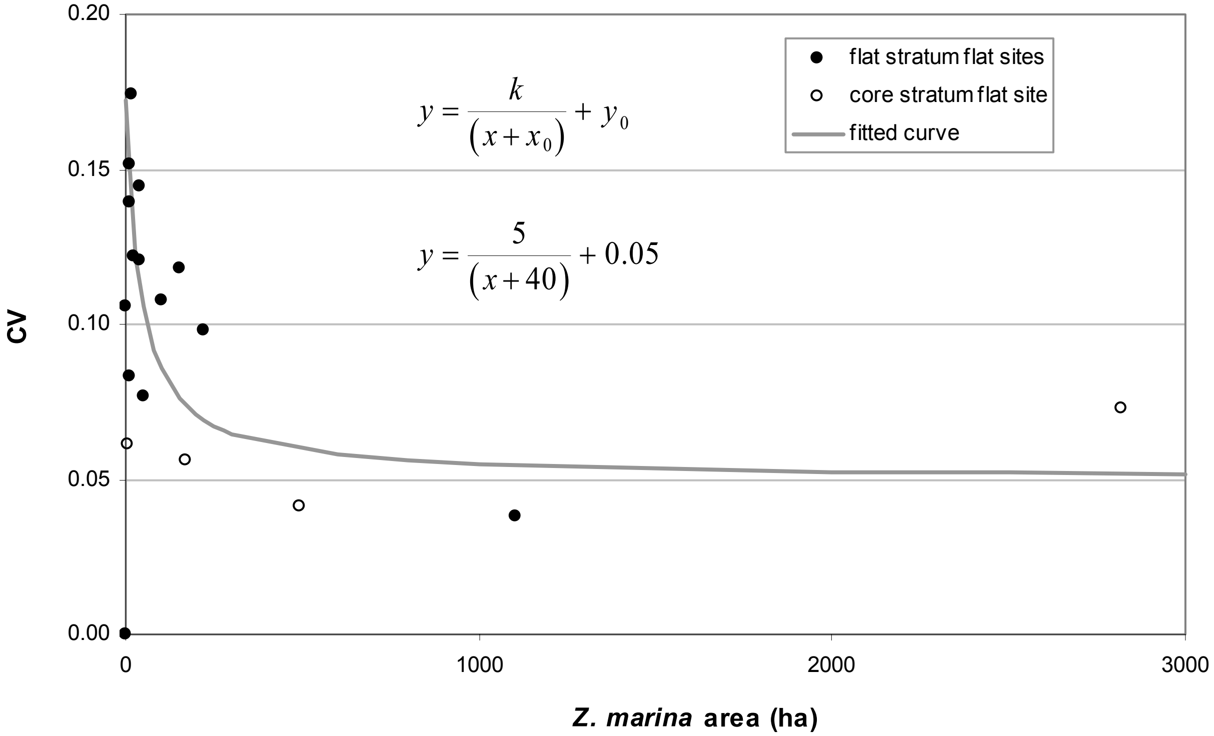


Figure 1-17. Inverse curve used to represent relationship between flat site CV and *Z. marina* area. Points are most recent estimates for sites sampled in 2000-2003.

CV Values in the Model Flats Dataset

The CV values in the model flats dataset include values from the 2000-2003 SVMP data and simulated values for unsampled flats. The simulated values were constrained to have the same dispersion characteristics about the CV-*Z. marina* area curve as the 2000-2003 data (Figure 1-17). Specifically, the residuals in the simulated data were drawn from a normal distribution with the same variance as the residuals of the 2000-2003 data, and a mean of zero.

The CV values for unsampled flat sites were generated with Gaussian stochasticity using

$$CV_i = \frac{k}{X_i + x_0} + y_0 + N(0, MSE_{CV})$$

Equation 1-3

where

CV_i = site-level CV for flat site i,
X_i = site-level *Z. marina* area for flat site i in hectares,
$N(0,MSE_{CV})$ = a random number drawn from a normal distribution with mean 0 and variance MSE_{CV},

MSE_{CV} = mean square error of the fit of the curve (Figure 1-17) to 2000-2003 flat site data,
x_0 = 40 ha,
y_0 = 0.05,
k = 5 ha.

This approach produced one negative simulated CV value that was clearly not physically meaningful. In addition, there were positive but extremely small CV values that were probably not realistic.

To address these problems, a low-value filter was applied. Simulated values of CV that were less than 50% of the minimum CV observed in the 2000-2003 data were rejected. Values of $N(0,MSE_{CV})$ that produced simulated CV values that did not meet this requirement were rejected and replaced by new random numbers until this requirement was met. Three values were rejected out of the initial 52 simulated values.

Figure 1-18 shows the scatter in the final simulated CV values for the 52 unsampled flat sites with respect to site *Z. marina* area.

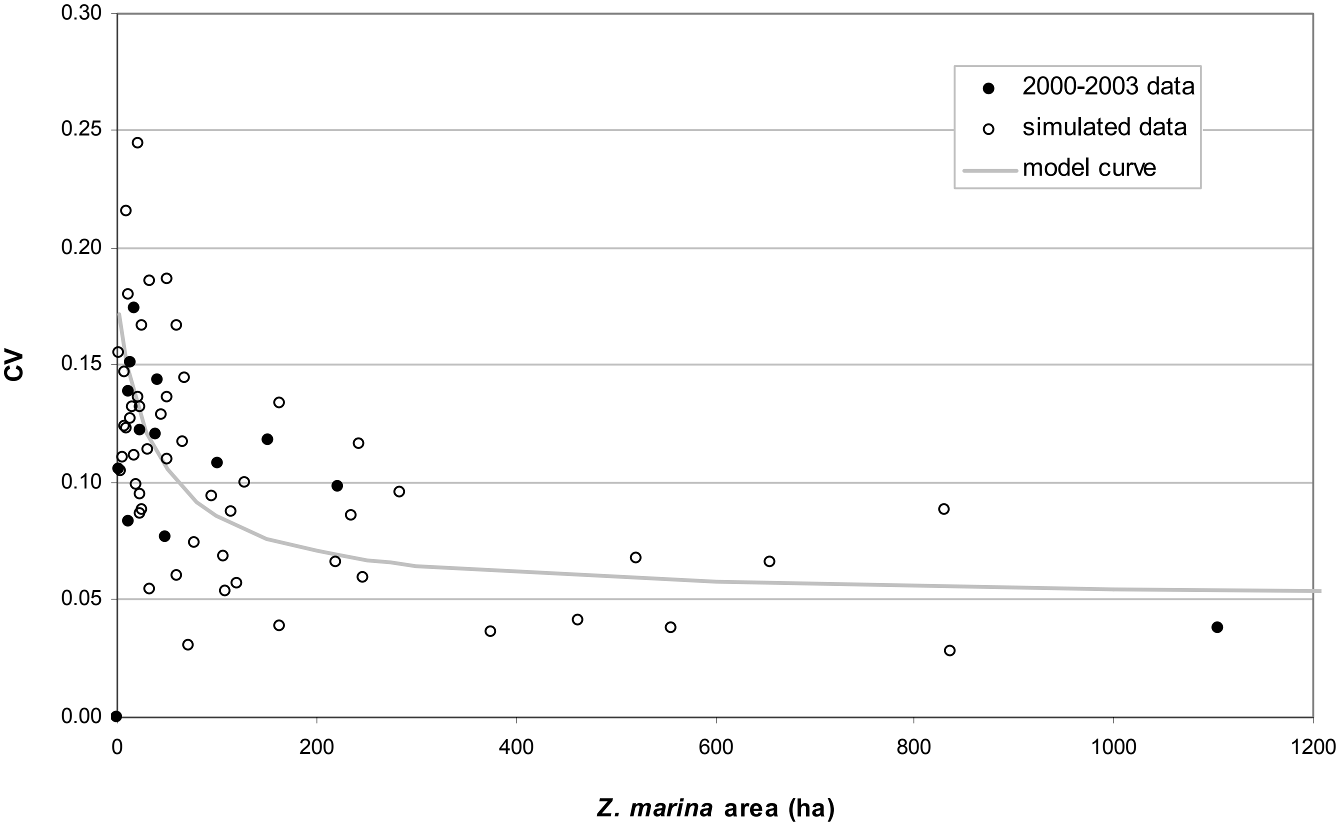


Figure 1-18. The relationship between CV and *Z. marina* area in the model flats dataset. Simulated values were derived from Equation 1-3 and a low-value filter. Values of $N(0,MSE_{CV})$ that did not meet this requirement were rejected and replaced by new random numbers.

Figure 1-19 shows the simple distribution of CV and compares this to the separate components of the distribution—sampled and simulated CV values. Although the total range in the simulated values is greater then the 2000-03 estimates, this is to be expected with the greater number of values. In general the simulated values closely resemble the 2000-03 data both in terms of central tendency and dispersion.

To confirm that the central tendencies were the same, a *t*-test was used to test for significant difference in the means of the two sets of residuals. The Smith-Satterthwaite procedure was used to estimate the degrees of freedom in the distribution of the test statistic (Milton and Arnold 1990, p.320; Zar 1999, p.129). This produced an estimate of 24 degrees of freedom. Results of the *t*-test indicate that we are not able to reject the null hypothesis of no difference (p=0.52).

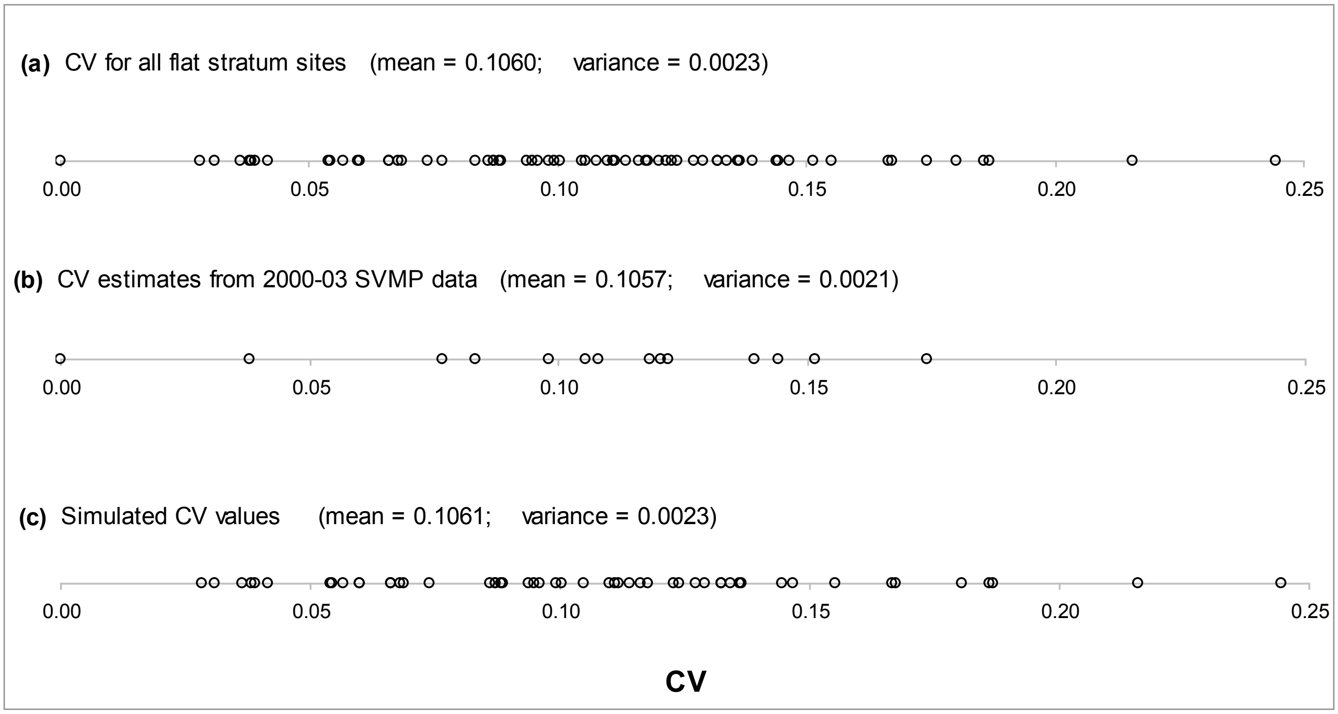


Figure 1-19. Distribution of CV in the model flats dataset; (a) includes values for all flat stratum sites, (b) includes only the 14 values based on 2000-2003 sampling; (c) includes only the 52 CV values simulated with Equation 1-3 and the low-value filter.

As a test of the stipulation that the residuals are normally distributed, three tests for normality were applied to the residuals of the 2000-03 CV values with respect to the model curve. The three tests provide no evidence to reject the null hypothesis of normality (Table 1-4).

Values of variance for unsampled sites were calculated as the product of simulated CV and the site *Z. marina* area value in the model flats dataset. The complete dataset is shown in Table 1-5.

test	statistic	significance
Kolmogorov-Smirnov	d=0.12526	p > 0.20
Lilliefors	-	p > 0.20
Shapiro-Wilk	W=0.91608	p = 0.22196

Table 1-4. Normal distribution tests for the distribution of residuals for the 2000-2003 CV-*Z.marina* area data points.

code	flats name	region	Site Area (ha)	*Z. marina* Area (ha)	Variance (ha^2)
flats01	Drayton Harbor	nps	1009.8	555.4	451.7
flats03	Birch Bay	nps	780.9	78.1	33.4
flats04	Lummi Flats N	nps	730.9	219.3	210.7
flats05	Lummi Flats S	nps	948.3	521.5	1254.5
flats07	Portage Bay S	nps	517.1	284.4	745.0
flats08	Portage Bay N	nps	428.7	41.8	36.3
flats09	Nooksack Delta W	nps	686.9	68.7	98.5
flats10	Nooksack Delta E	nps	718.6	0.0	0.0
flats11	Samish Bay N	nps	1944.5	1104.1	1758.5
flats12	Samish Bay S	nps	1509.3	830.1	5394.6
flats14	March Pt.	nps	683.7	376.0	186.5
flats15	Fidalgo Bay	nps	840.9	462.5	372.3
flats17	Bowman Bay	sji	37.3	11.2	4.1
flats18	Similk Bay	swh	544.4	39.0	22.1
flats19	Pull and Be Damned Point	swh	983.2	152.4	324.9
flats20	Skagit Bay north	swh	2679.7	221.9	473.5
flats21	Skagit Bay mid	swh	2185.6	655.7	1875.8
flats22	Port Susan W	swh	2352.2	235.2	408.8
flats23	Port Susan E	swh	2440.7	244.1	804.2
flats25	Tulalip Bay	swh	166.1	49.8	46.0
flats26	Snohomish Delta N	swh	510.6	51.1	91.0
flats27	Snohomish Delta mid	swh	1630.3	163.0	40.6
flats28	Snohomish Delta south	swh	473.0	100.2	117.0
flats29	Coronet Bay	swh	55.9	16.8	3.5
flats30	Cultus Bay	cps	444.9	44.5	33.0
flats31	Oak Harbor	swh	363.5	109.0	34.7
flats32	Dugualla Bay	swh	428.5	128.6	166.3
flats33	Quartermaster Harbor	cps	200.4	60.1	101.1
flats34	Nisqually Delta W	cps	726.9	72.7	5.1
flats35	Nisqually Delta E	cps	542.6	24.3	8.8
flats36	Eagle Harbor	cps	43.0	12.9	2.7
flats37	Wing Point	cps	49.0	17.6	9.4
flats38	Port Madison	cps	84.3	25.3	5.0
flats39	Liberty Bay	cps	357.5	107.2	54.1
flats40	Miller Bay	cps	110.8	33.2	38.1
flats41	Dosewallips	hdc	316.2	94.9	79.2
flats42	Quilcene Bay	hdc	384.8	115.4	101.6
flats43	Dabob Bay	hdc	163.2	11.1	2.4
flats44	Case Shoal	hdc	54.6	16.4	4.7
flats45	Hood Head	hdc	101.1	30.3	11.9
flats46	Kilisut Harbor	cps	200.0	60.0	13.0
flats47	Travis Spit	fuc	193.0	48.6	13.9
flats48	Sequim Bay	fuc	78.8	23.6	4.2
flats49	Old Town	fuc	403.5	121.1	47.2
flats50	Dungeness Bay	fuc	545.7	163.7	481.4
flats51	Provost Harbor	sji	28.6	2.9	0.2
flats52	Nelson Bay, Henry Island	sji	69.0	20.7	25.6
flats53	Westcott Bay, San Juan	sji	158.8	14.1	4.6
flats54	Garrison Bay, San Juan	sji	80.8	24.2	10.3
flats55	Mitchell Bay, San Juan	sji	66.8	20.1	4.0
flats56	False Bay, San Juan	sji	110.6	33.2	3.3
flats57	Fisherman's Bay, Lopez	sji	171.4	51.4	32.0
flats58	Barlow Bay, Lopez	sji	76.0	22.8	4.7
flats59	Mud Bay, Lopez	sji	220.8	66.3	60.8
flats60	Hunter Bay, Lopez	sji	65.9	2.3	0.1
flats61	Shoal Bay, Lopez	sji	83.1	24.9	17.2
flats62	Swifts Bay, Lopez	sji	96.4	11.0	0.8
flats63	Blind Bay, Shaw	sji	74.2	22.3	9.2
flats64	Squaw Bay	sji	17.7	5.3	0.3
flats65	Thatcher Bay, Decater	sji	28.1	8.4	1.5
flats66	Shallow Bay, Sucia	sji	23.4	7.0	0.8
flats67	Fossil Bay, Sucia	sji	35.5	10.7	5.3
flats68	Secret Harbor, Cypress	sji	29.6	8.9	1.2
flats69	Eagle Cove, Cypress	sji	12.7	3.8	0.2
flats70	S. Fork Skagit River	swh	2786.1	835.8	556.8
flats71	Skagit Bay south	swh	2465.5	246.6	217.2

Table 1-5. The complete model flats dataset.

1.3.3 Influence of Flat Sites on Point Estimates–the ζ index

This section presents a quantitative approach that used the model flats dataset to identify outlier sites in the flats stratum for potential removal and placement in a separate stratum.

Z. marina area is a reasonable parameter to use for separating flats into two strata (see Skalski 2003, Equation 9). The flats extrapolation calculation can be represented as the product of a weighted mean eelgrass ratio, $\overline{R}$, and total area of the flats stratum, *A*. The value of $\overline{R}$ is simply the mean of eelgrass ratios of sampled flats weighted by site area:

$$\overline{R} = \frac{\sum \left(R_i \bullet a_i \right)}{\sum a_i}$$

Equation 1-4

where a_i is the site area for site *i*. Large sites have more influence on the result. Sites that should be removed from the existing flats stratum and put into a separate stratum will be outliers with respect to their influence on $\overline{R}$. These sites can be identified by comparison of individual terms in the numerator of Equation 1-4. Each of these terms reduces simply to site *Z. marina* area, as seen in Skalski (2003, equation 9), thereby supporting the use of this parameter as the basis for revising the stratification.

The use of *Z. marina* area alone for stratification, however, ignores the effect of each site in the denominator of Equation 1-4. Consequently a site such as flats20, which is a very large site with low *R*, would not be identified by a *Z. marina*-based criterion as an influential site. However, flats20 was shown in section 1.1 (e.g. Figure 1-3, p.9) to have a strong influence on stratum results.

Consequently, it is preferable to have an approach that could identify low-*R* outliers such as flats20 as well as high-*R* outliers such as flats11. The calculations used in section 1.1 to isolate the influence of individual sites on the stratum estimates can be adapted for this purpose. In this case, the flat stratum *Z. marina* area equation (Skalski 2003, Equation 9) is applied as each of the 66 flat sites in the model flats dataset is in turn withheld from the calculation. The difference obtained when site *i* is withheld and when it is not withheld is a measure of influence of site *i* on the *Z. marina* area estimate.

The following equations develop a quantitative index designed to measure the influence of each flat site on stratum *Z. marina* area estimates.

For each of the 66 sites in the 2004 flats stratum, an estimate of flats stratum *Z. marina* area is estimated with that site withheld. Only the 65 remaining sites are used in the calculation. The estimate with site *i* withheld is denoted $\hat{B}_{(i)}$, and is calculated as

$$\hat{B}_{(i)} = \frac{\sum_{j \neq i}^{N} X_j}{\sum_{j \neq i}^{N} a_j} \cdot A$$

Equation 1-5

where

X_j = *Z. marina* area of site j
a_j = site area of site j
A = total area of all sites in the flats stratum

and the sums are over all sites except site *i*. The resulting distribution of $\hat{B}_{(i)}$ is shown in Figure 1-20.

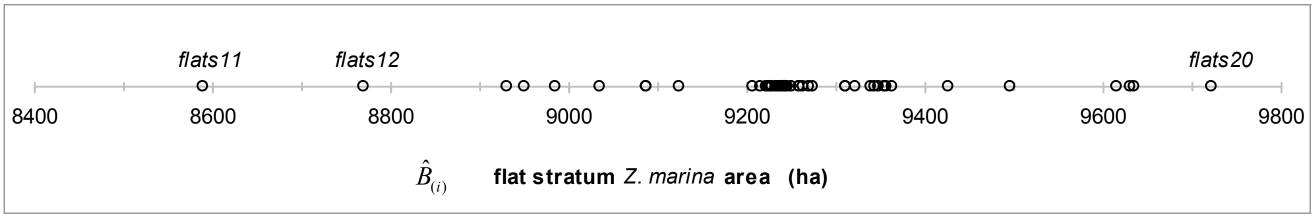


Figure 1-20. Estimates of *Z. marina* area for the flats stratum where one site has been withheld from each calculation, $\hat{B}_{(i)}$. Numbers above data points indicate which site was withheld for that calculation. Based on the site *Z. marina* estimates in Table 1-5.

The calculations using Equation 1-5 result in 66 estimates of flat stratum *Z. marina* area, each based on a sample size of 65. For each site *i*, the difference

$$\delta_i = \hat{B} - \hat{B}_i$$

is a measure of the influence of site *i*. The mean and variance of the 66 values of δ_i are used to produce the normalized index

$$\zeta_i = \frac{\delta_i - \overline{\delta}}{\sqrt{Var(\delta)}}$$

Equation 1-6

The index ζ_i essentially translates the distribution of $\hat{B}_{(i)}$, shown in Figure 1-20, so that the deviation from zero can be interpreted as a direct measure of influence of site *i*. Positive values indicate that the site tends to raise the estimate of stratum *Z. marina* area and negative values indicate that the site tends to drop the estimate. The distribution of the ζ index is shown in Figure 1-21.

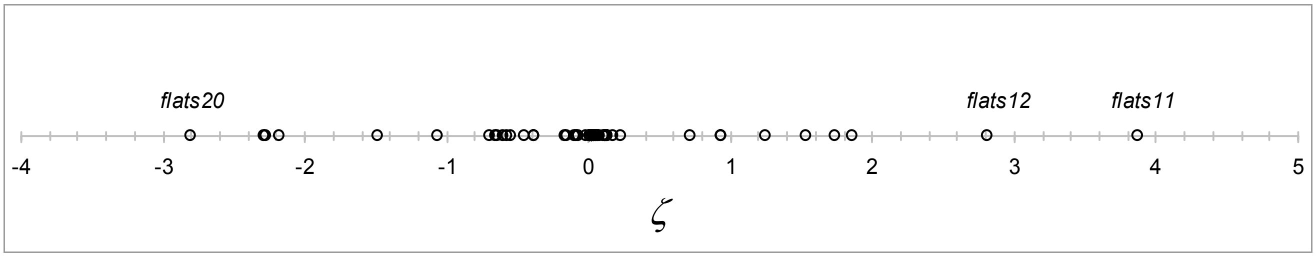


Figure 1-21. Distribution of the influence index, ζ, calculated from Equation 1-6 for the 66 sites in the 2004 flat sampling frame, excluding core stratum flat sites. Deviation from zero is a direct measure of influence of the site on flat stratum *Z. marina* area estimates.

The results in Figure 1-21 show that flats11 remains the most influential site when the entire stratum is considered and that it tends to raise the stratum estimate of *Z. marina* area. Flats20 remains the most influential site with the opposite effect—its presence drops the stratum estimate.

The absolute value of the index ζ is useful in identifying the most influential sites regardless of the sign of the influence. The 16 most influential sites are shown in ranked order in Figure 1-22 and Table 1-6.

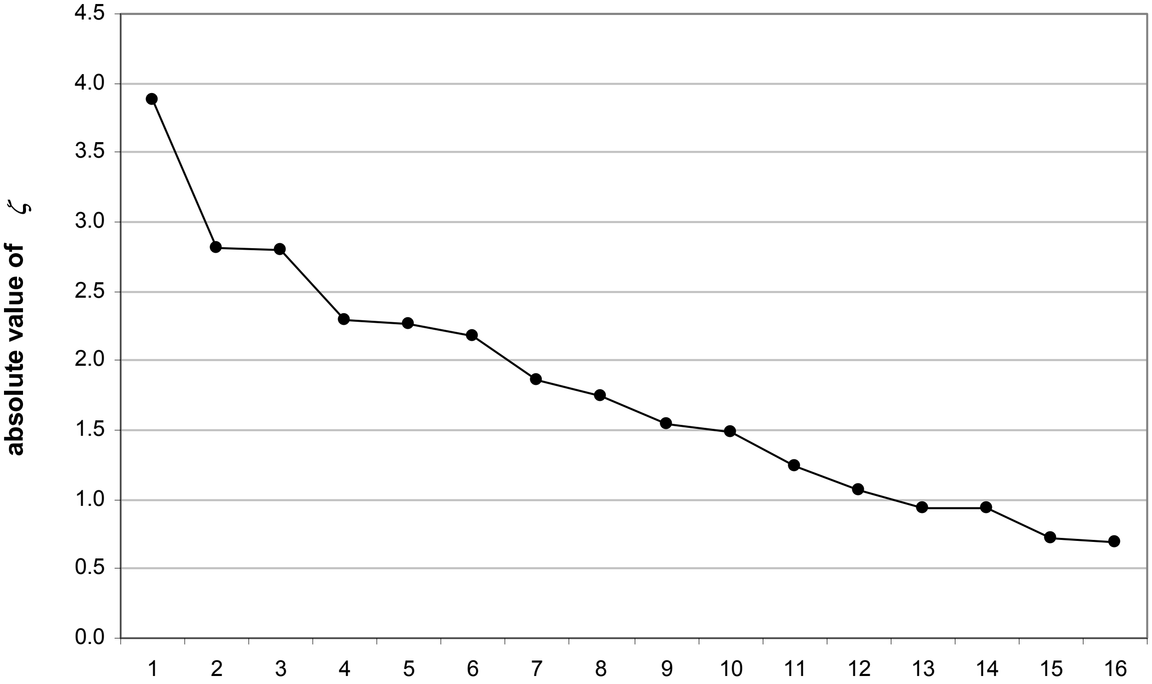


Figure 1-22. The absolute value of influence index ζ for the 16 most influential flat sites ranked in order of influence. This index measures influence of individual sites on estimates of flat stratum *Z. marina* area.

rank	site code	site name	region	ζ
1	flats11	Samish Bay north	nps	3.88
2	flats12	Samish Bay south	nps	2.81
3	flats20	Skagit Bay north	swh	-2.80
4	flats71	Skagit Bay south	swh	-2.29
5	flats23	Port Susan east	swh	-2.27
6	flats22	Port Susan west	swh	-2.18
7	flats01	Drayton Harbor	nps	1.85
8	flats05	Lummi Flats south	nps	1.74
9	flats15	Fidalgo Bay	nps	1.54
10	flats27	Snohomish Delta mid	swh	-1.48
11	flats14	March Pt.	nps	1.25
12	flats10	Nooksack Delta east	nps	-1.07
13	flats07	Portage Bay south	nps	0.94
14	flats70	S. Fork Skagit River	swh	0.93
15	flats21	Skagit Bay mid	swh	0.72
16	flats03	Birch Bay	nps	-0.69

Table 1-6. The 16 most influential flat sites on stratum *Z. marina* area estimation based on the absolute value of the index ζ.

1.3.4 Influence of Flat Sites on Variance Estimates–the ξ index

The influence of flat sites on stratum variance estimates was calculated for comparison with the results of the previous section.

The approach here also assesses the influence of each flat site by withholding it from the stratum calculation, in this case the calculation of the variance estimator. However, two changes are made to the variance estimator for these calculations. The flat stratum variance estimator is given by (Skalski 2003, Equation 11)

$$\widehat{Var}\left(\hat{B}\right) = N^2\left(1-\frac{n}{N}\right)\frac{\sum_{j=1}^{n}\left(\hat{X}_j - a_j\hat{R}\right)^2}{n\left(n-1\right)} + \frac{N\sum_{j=1}^{n}\widehat{Var}\left(\hat{X}_j \middle| x_j\right)}{n}.$$

Equation 1-7

Since these influence calculations estimate variance using almost the entire population (65 of 66 flat sites), the finite population correction

$$fpc = \left(1-\frac{n}{N}\right)$$

exerts a strong effect that is not useful for this purpose. It is therefore ignored for the influence calculations. Furthermore, the effect of withholding individual sites on the estimate of *Z. marina* area ratio, $\hat{R}$, is ignored to simplify the calculations.

Therefore, the effect of withholding a site from the calculation is simply to ignore one term from each of the sums in Equation 1-7. The estimate with site *i* withheld is denoted $\widehat{Var}_{(i)}\left(\hat{B}\right)$, and is calculated as

$$\widehat{Var}_{(i)}\left(\hat{B}\right) = N^2 \frac{\sum_{j \neq i}^{n}\left(\hat{X}_j - a_j\hat{R}\right)^2}{n\,(n-1)} + \frac{N\sum_{j \neq i}^{n}\widehat{Var}\left(\hat{X}_j \middle| x_j\right)}{n}.$$

Equation 1-8

The application of Equation 1-8 results in 66 estimates of flat stratum variance, each based on a sample size of 65. For each site *i*, the difference

$$\delta_i = \widehat{Var}\left(\hat{B}\right) - \widehat{Var}_{(i)}\left(\hat{B}\right)$$

is a measure of the influence of site *i*, where $\widehat{Var}\left(\hat{B}\right)$ is the variance estimate based on all flat sites (none withheld). The mean and variance of the 66 values of δ_i are used to produce the normalized index

$$\xi_i = \frac{\delta_i - \overline{\delta}}{\sqrt{Var(\delta)}}.$$

Equation 1-9

As before, deviation from zero can be interpreted as a direct measure of influence of site *i* and positive values indicate that the site tends to raise the variance estimate of the stratum. The distribution of the ξ index is shown in Figure 1-23.

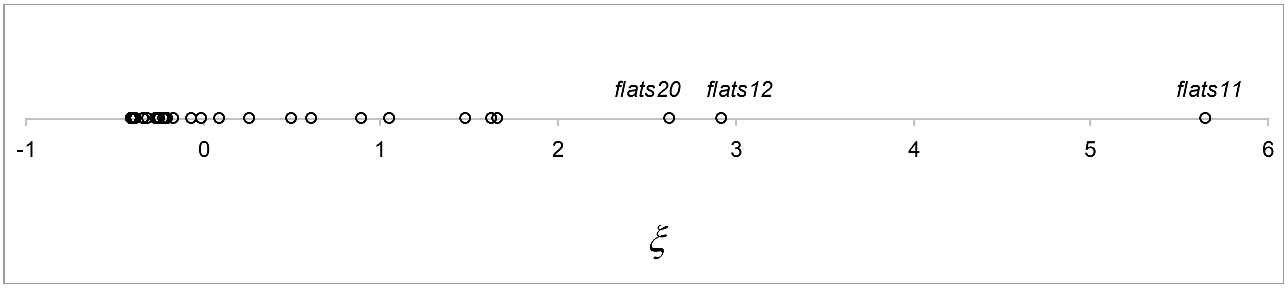


Figure 1-23. Distribution of the influence index, ξ, calculated from Equation 1-9 for the 66 sites in the 2004 flat sampling frame, excluding core stratum flat sites. Deviation from zero is a direct measure of influence of the site on flat stratum variance estimates.

The absolute value of the index ξ is useful in identifying the most influential sites regardless of the sign of the influence. The 15 most influential sites are shown in ranked order in Figure 1-24 and Table 1-7.

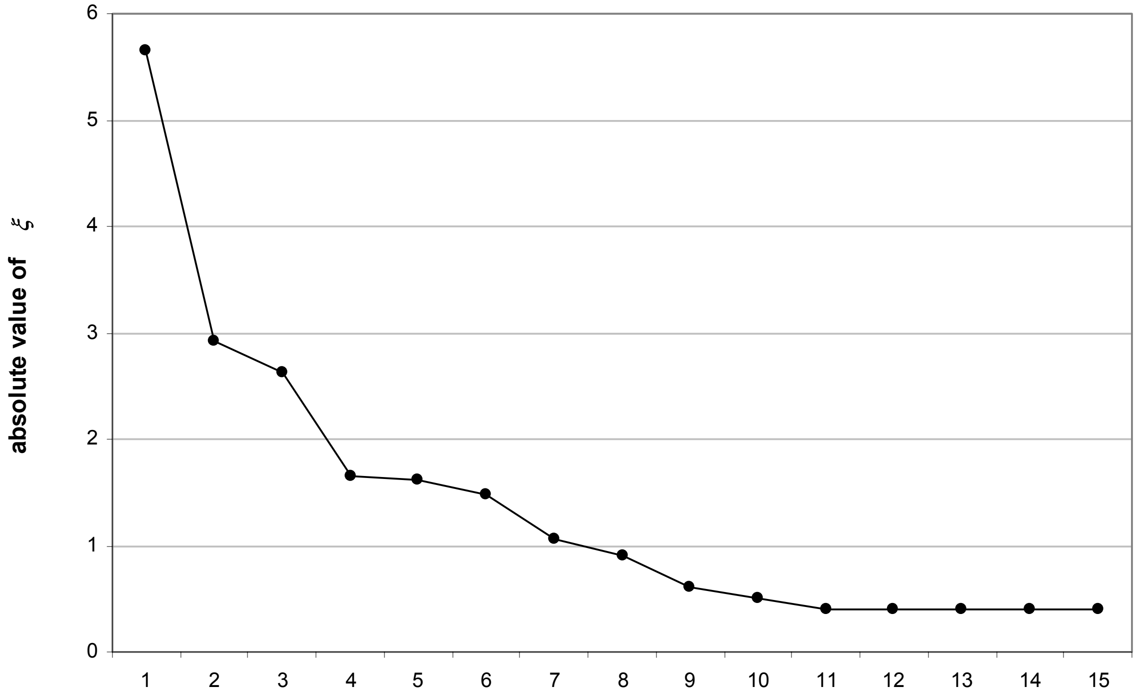


Figure 1-24. The absolute value of influence index ξ for the 15 most influential flat sites ranked in order of influence. This index measures influence of individual sites on estimates of flat stratum variance.

rank	code	flats name	region	ξ
1	flats11	Samish Bay N	nps	5.66
2	flats12	Samish Bay S	nps	2.92
3	flats20	Skagit Bay N	swh	2.63
4	flats71	Skagit Bay S	swh	1.66
5	flats23	Port Susan E	swh	1.63
6	flats22	Port Susan W	swh	1.48
7	flats01	Drayton Harbor	nps	1.06
8	flats05	Lummi Flats S	nps	0.90
9	flats15	Fidalgo Bay	nps	0.61
10	flats27	Snohomish Delta mid	swh	0.50
11	flats69	Eagle Cove, Cypress	sjs	-0.40
12	flats64	Squaw Bay	sjs	-0.40
13	flats66	Shallow Bay, Sucia	sjs	-0.40
14	flats68	Secret Harbor, Cypress	sjs	-0.40
15	flats65	Thatcher Bay, Decater	sjs	-0.40

Table 1-7. The 15 most influential flat sites on flat stratum variance based on the absolute value of the index ξ.

The ranking of flat sites based on the influence on variance estimates (Table 1-7) is identical to the ranking based on influence on *Z. marina* area estimates (Table 1-6) through the top ten sites.

1.3.5 Recommendations on Flats Stratification

It is clear from the results presented in this chapter that there is a consistent ranking of flat sites that are most influential on both *Z. marina* area estimates and variance estimates for the 2000-03 flat stratum. It is also clear that the precision of the stratum estimates would improve if the flats stratification were revised and outlier sites were removed.

There are two general approaches to modifying the stratification. If the 66 flat sites are divided into two subsets of comparable size, then rotational sampling could be applied in each new stratum. An alternate approach is to remove just a few sites from the existing flats stratum and move them to a new stratum that, like the core stratum, is completely surveyed each sample season. The latter approach is recommended here because of two reasons.

1. Results suggest that just a few sites are strongly influential on stratum estimates and, if removed, would improve the precision of these estimates.
2. The SVMP is currently at or near capacity in terms of number of sites sampled given current resources and it would not be possible to adequately sample a new stratum under rotational sampling.

Figure 1-22 and Figure 1-24 indicate that there are two breakpoints in the ranking of most influential sites. The first is after flats11, the most influential site. The second is after the third site. It is recommended that the second breakpoint be used as the basis for removing the top three sites from the flats stratum and placing them in a separate stratum where they are sampled each year. These top three sites to be removed are flats11, flats12 and flats20.

This recommendation was initially made in an early draft of this report and has been implemented in the 2004 sampling season. Flats11 and flats20 were in the 2003 sample pool and were already scheduled to be sampled in the 2004 sample season. Under this recommendation, flats12 was deliberately added to the 2004 sample pool.

These changes do not affect the existing SVMP estimates based on 2000-03 data. The 2000-03 estimates are still based on the 2000-03 flats stratification and sampling frame with 67 flat sites.

The remainder of this report will denote the main flats stratum that is subject to rotational sampling as the “rotational flats” stratum. The new flats stratum that contains the three most strongly influential sites and is subject to complete census every season will be denoted as the “persistent flats” stratum.

Table 1-8 summarizes the recommendations for changes to the flats sampling frame that were implemented in the 2004 sampling season. Table 1-9 gives details of which flat sites have been sampled and shows the stratum changes in the context of flat site rotation over 2000-04.

flats sampling frame	total number of sites	stratum	number of sites in stratum
2000-03	71	core	4
		flats	67
2004	70	core	4
		persistent flats	3
		rotational flats	63

Table 1-8. Summary of changes to the flats sampling frame and flats strata implemented in the 2004 sampling season.

As new flats rotate into the sample pool in future years, data from these sites should be compared to the values in the model flats dataset presented in this chapter. Very large discrepancies may suggest that the 2004 flats stratification should be reevaluated. However, no changes should be made unless another flat site is shown to have influence similar to the three sites in the persistent flats stratum. In general, future changes should be avoided unless there is strong evidence to support the change.

If flat sites are removed from the rotational flats stratum, they should only be removed after they have completed their random rotation cycle (five years in the rotational flats stratum). The creation of the persistent flats stratum in 2004 removes flats11 and flats20 from the rotational flats stratum before the end of their rotation cycle, but in this case the benefits of improved precision justify this change.

	2000	2001	2002	2003	2004
rotational flats stratum	flats47	flats47			
	flats53	flats53			
	flats60	flats60	flats60		
	flats28	flats28	flats28		
	flats18	flats18	flats18	flats11	
	flats20	flats20	flats20	flats20	
	flats35	flats35	flats35	flats35	flats35
	flats43	flats43	flats43	flats43	flats43
	flats62	flats62	flats62	flats62	flats62
		flats11	flats11	flats18	flats18
			flats10	flats10	flats10
			flats37	flats37	flats37
				flats08	flats08
				flats19	flats19
					flats41
					flats70
persistent flats stratum					flats11
					flats20
					flats12

Table 1-9. Summary of flats sites sampled in the first five years of the SVMP. Dark shading indicates where flat sites rotate into the sample pool. Light shading indicates sites that rotate out of the sample pool in the following year. The sites that will rotate out after 2004 had not been selected at the time this report was written.

1.4 Effect of Change in Flats Strata on Precision

The change to the flats stratification described in the previous section was motivated by the clear identification of outlier sites and the tremendous gain in precision when outliers are removed. However, this gain in precision was demonstrated in section 1.1 only with respect to the initial stratum estimates. It is possible that the precision of the retrospectively adjusted estimates could decline when outlier sites are removed from the stratum. This would represent an argument against the changes recommended in the previous section. This section investigates this possibility.

As noted by Cochran (1977), the gain in precision associated with the adjustment of estimates is dependent on high correlation in estimates between sampling occasions. If the correlation is lowered, then the gain in precision declines. The presence of outliers in the flats stratum ensures a high correlation, as shown in Figure 1-25. As the outlier sites are shifted to the persistent flats stratum, we would expect the correlation in the residual rotational flats stratum to decline and the gain in precision associated with retrospective adjustment to decline.

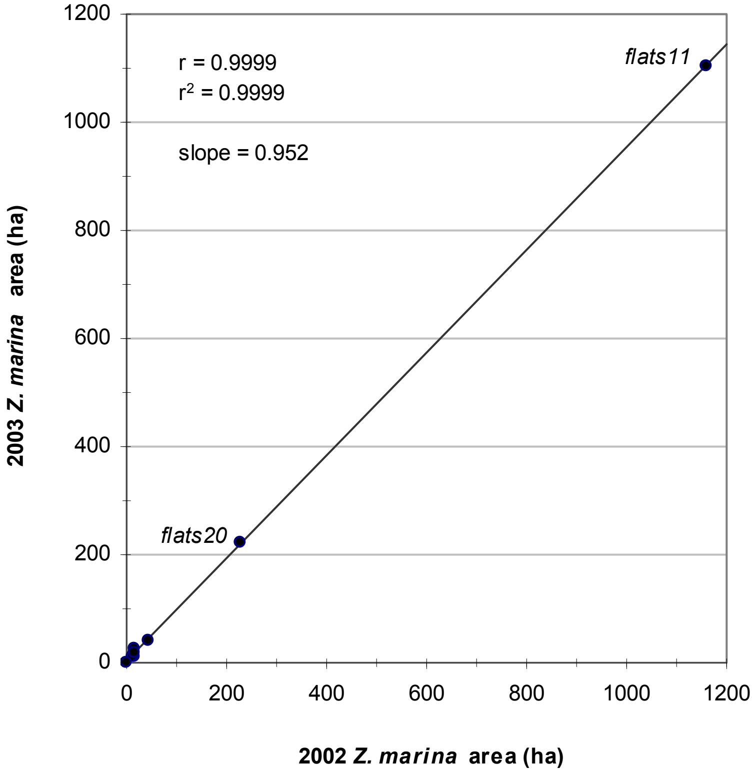


Figure 1-25. Regression of 2003 *Z. marina* area estimates vs. 2002 estimates for flat stratum sites. This illustrates the high correlation in these data and the role of an outlier site such as flats11 in enhancing this correlation.

Table 1-10 confirms that when flats11 is withheld the correlations decline but only modestly–the correlations remain high.

	original correlation	correlation with flats11 withheld
2001-02	0.9971	0.9309
2002-03	0.9999	0.9987

Table 1-10. The effect of an outlier site, flats11, on year-to-year correlations in matching flats stratum sites.

To see how this decline affects the gain in precision associated with adjustment, the 2001 and 2002 estimates were recalculated with flats11 withheld. The results of this exercise are shown in Figure 1-26 and Table 1-11. Two key points emerge from these results:

1. Withholding flats11 from the calculation strongly diminishes the magnitude of the improvement in variance associated with adjustment (e.g. -98% decline in variance with flats11 and only -49% decline without flats11 in the 2001 results).

2. The initial variance is much less when flats11 is withheld so that even with a weaker improvement the adjusted precision is greater than when flats11 is included.

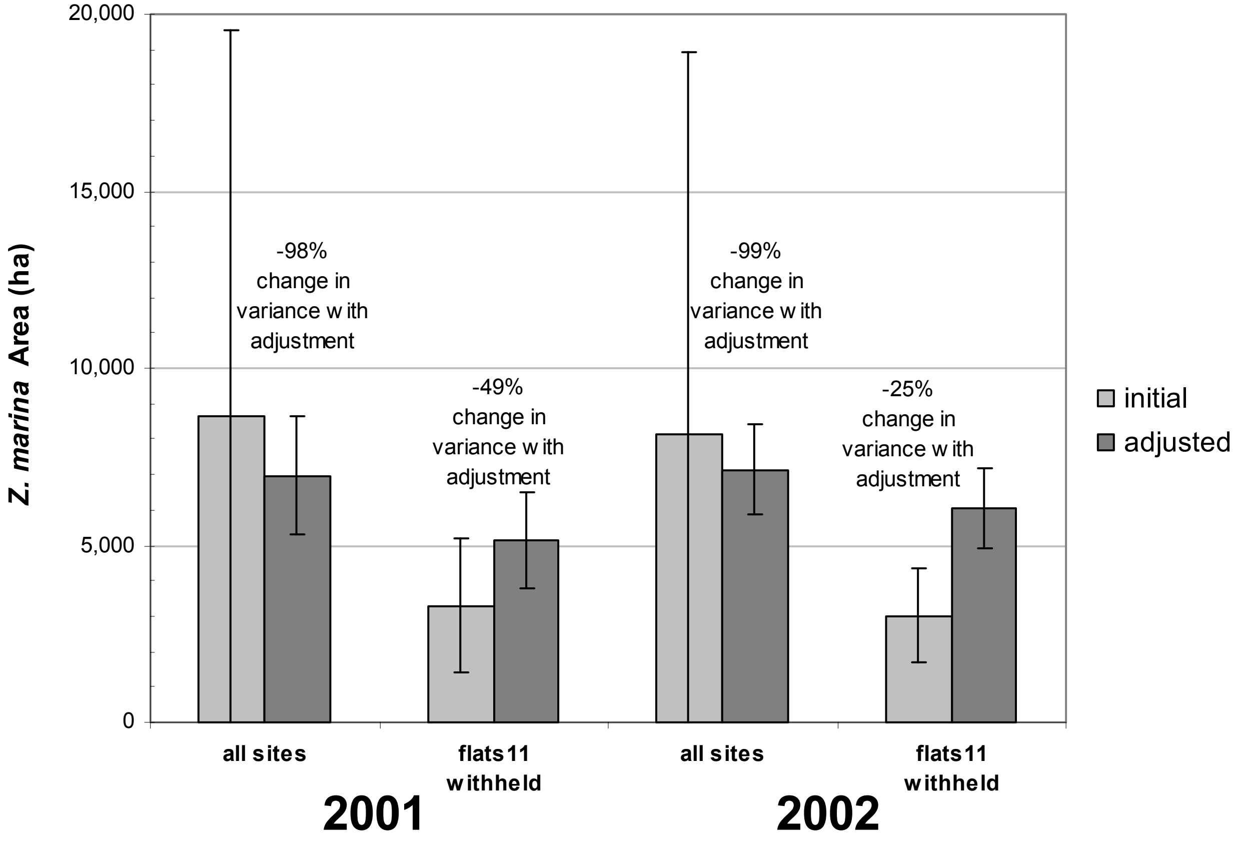


Figure 1-26. 2001 and 2002 initial and adjusted estimates of flat stratum *Z. marina* area calculated with all sites and with flats11 withheld from the sample. Error bars indicate 95% confidence intervals. The percent change in variance associated with adjustment (relative to initial variance) is also indicated. Values are shown in Table 1-11.

		2001		**2002**	
		all sites	**flats11 withheld**	**all sites**	**flats11 withheld**
initial	***Z. marina* area** (ha)	8,628	3,302	8,150	3,006
	variance (ha^2)	31,034,593	939,613	30,092,020	454,592
adjusted	***Z. marina* area** (ha)	6,970	5,129	7,130	6,036
	variance (ha^2)	729,006	480,727	419,624	339,424

Table 1-11. 2001 and 2002 initial and adjusted estimates of flat stratum *Z. marina* area calculated with all sites and with flats11 withheld from the sample.

The second point suggests that overall, it is beneficial to remove the outlier sites from the flats stratum, even though the adjustment calculation is then less effective in improving precision. However, a thorough assessment of the new 2004 flats stratification must account for the variance of the outlier sites in the new persistent flats stratum. The appropriate comparison is between the 2000-03 flats variance and the combined variance of the rotational flats and persistent flats strata. This combined variance of the initial *Z. marina* area estimates is given by

$$\widehat{Var}\left(\hat{B}\ \right)=\widehat{Var}\left(\hat{B}_{SPR}\right)+\sum_{j=1}^{3}\widehat{Var}\left(\hat{B}_{\zeta}\middle|B_{\zeta}\right),$$

Equation 1-10

where

$\widehat{Var}\left(\hat{B}_{SPR}\right)$ = the estimated variance of the rotational flats stratum (calculated using the flats stratum variance estimator of Skalski 2003–Equation 11),

$\sum_{j=1}^{3}\widehat{Var}\left(\hat{B}_{\zeta}\middle|B_{\zeta}\right)$ = the sum of site-level variances for each flat in the persistent flats stratum.

The combined variance of the adjusted estimates is the same as Equation 1-10 except the variance of the rotational flats is replaced with the adjusted variance estimate, $\widehat{Var}\left(\tilde{B}_{SPR}\right)$, giving

$$\widehat{Var}\left(\tilde{B}\ \right)=\widehat{Var}\left(\tilde{B}_{SPR}\right)+\sum_{j=1}^{3}\widehat{Var}\left(\hat{B}_{\zeta}\middle|B_{\zeta}\right).$$

We use 2002 SVMP data to produce a rough comparison of total flats variance (excluding the core flats) under the 2000-03 and 2004 flats stratification scenarios. To produce estimates for the rotational flats under the 2004 scenario we simply remove flats11 and flats20 from the sample. To produce an estimate for the persistent flats stratum we use the flats11 and flats20 site-level results and assume that flats12 is identical to flats11. The results of this exercise are shown in Table 1-12.

These results demonstrate that even when both flats strata are taken into account, we can expect the overall flats variance to be lower under the 2004 stratification scenario. Once the outlier flats are removed, the effectiveness of the retrospective adjustment is much lower in the residual rotational flats stratum (559,725 ha^2 improves to 365,455 ha^2) when compared to the 2000-03 flats stratum (30,092,020 ha^2 improves to 419,624 ha^2), but the final variance is lower (371,964 ha^2 compared to 419,624 ha^2). Note also that the change in stratification dramatically improves the precision of the initial estimates, before the effectiveness of the adjustment comes into play.

2002	2000-03 Flats Stratum	2004 Flats Strata		
	1 $\widehat{Var}\left(\hat{B}\right)$ $\widehat{Var}\left(\tilde{B}\right)$	2 $\widehat{Var}\left(\hat{B}_{SPR}\right)$ $\widehat{Var}\left(\tilde{B}_{SPR}\right)$	3 $\sum_{j=1}^{3}\widehat{Var}\left(\hat{B}_{\zeta} \mid B_{\zeta}\right)$	4 $\widehat{Var}\left(\hat{B}\right)$ $\widehat{Var}\left(\tilde{B}\right)$
initial estimates (ha^2)	30,092,020	559,725	6,510	566,235
adjusted estimates (ha^2)	419,624	365,455	6,510	371,964

Table 1-12. Comparison of variance in the 2000-03 flats stratum (column 1) with the variance in the 2004 flats strata (column 4). The latter is a sum of variance of the rotational flats stratum (column 2) and variance of the persistent flats stratum (column 3). Results are based on 2002 data with adjustment using 2003 data. Flats11 and flats20 were simply withheld to simulate the rotational flats stratum. The site-level variance of flats12 was assumed to be identical to flats11 in order to calculate the persistent flats variance.

The precision of the rotational flats stratum under the 2004 stratification scenario is expected to be slightly better in actual application than estimated here because the sample size was artificially reduced in this example by the removal of flats11 and flats20.

1.5 Summary

The primary results produced in this chapter are recommended changes to the flats sampling frame and to the stratification of flat sites in this frame. The bulk of the chapter presents analysis that weights the pros and cons of these changes and develops the rationale for making the changes.

These changes have been implemented for the first time in the 2004 SVMP sampling season.

The changes to the flats sampling frame improve the consistency of the delineation of flat sites. It is anticipated that the changes to the flats stratification will improve the precision of estimates of *Z. marina* area in the flat sites thereby improving the ability of the SVMP to detect change in *Z. marina* area over Puget Sound.

2 Anomalous Retrospective Adjustment Results

The retrospective adjustment of the 2001 and 2002 status results produced modest changes to the *Z. marina* area estimates for Puget Sound (-9% and -6% respectively, relative to initial estimates) and dramatic improvements in variance estimates (-94% and -95% declines in variance respectively) (Berry et al. 2003[3]).

The SVMP statistical framework (Skalski 2003, pp.27-29) shows that this strong improvement in variance is necessary for the project to meet its multi-year trend detection goal with a reasonable level of power (see also Chapter 3). However, this strong improvement is much greater than anticipated from theoretical considerations. These results therefore deserve further attention to ensure they are reasonable and consistent with the assumptions of the technique.

The previous chapter suggests that changes to the flats stratification implemented in 2004 will profoundly reduce the magnitude of the improvement in precision associated with retrospective adjustment. Once the analysis of 2004 data is completed, it will be clear whether this is accurate. Nevertheless, it will be useful to examine the retrospective adjustment calculations to ensure that the estimates based on the 2001-2003 SVMP data are robust.

This chapter first compares observed SVMP retrospective adjustment results with theoretical calculations. Then an examination of the adjustment calculations for the flats stratum traces the origin of the unexpectedly strong improvements in precision. Results in the literature addressing bias in variance estimates of ratio and regression estimators are then reviewed. The feasibility of this bias as an explanation for the strong improvements in SVMP results is explored.

2.1 Theoretical Improvement with Adjusted Estimates

Cochran (1977) presents the improvement in precision to be expected with adjustment for a number of specific cases. The example that is most analogous to the SVMP design includes sampling on two occasions (h=2), high correlation in matched sites (ρ=0.95) and high level of matched sites (75%). He estimates a 22% gain in efficiency for the status value upon adjustment. This example differs from the SVMP case, most obviously by using fixed weights (see further discussion of this example in section 4.3.1 on p.84).

---

[3] The results in Berry et al. (2003) produce slightly different values than these due to slight refinement of estimates following publication of that report. The revised estimates will be published in Dowty et al. (in prep.).

In section 4.3.2 (p.86), improvements in status estimates are calculated for a case more analogous to the SVMP design. In this case, when matched site correlation is ρ=0.98 and 80% of sites are matched, the ratio of adjusted to initial variance for the status estimate is 0.84. The weights were not fixed for this calculation as is the case for actual SVMP analyses.

In order to compare the observed SVMP improvements with these theoretical results, we must use the same measure of improvement. If the ratio of adjusted to initial variance is denoted g_h for sampling occasion h, then the percent change relative to the initial estimates is given by

$$\%\text{ change} = 100 \cdot \frac{Var\left(\tilde{B}_{ih}\right) - Var\left(B_{ih}\right)}{Var\left(B_{ih}\right)} \quad = \quad 100\left(g_h - 1\right)$$

where $Var\left(\tilde{B}_{ih}\right)$ is the variance of the adjusted *Z. marina* area estimate for stratum *i* on occasion *h* and $Var\left(B_{ih}\right)$ is the corresponding initial estimate and

$$g_h = \frac{Var\left(\tilde{B}_{ih}\right)}{Var\left(B_{ih}\right)}.$$

The percent gain in efficiency as defined by Cochran (1977, p.353) is given by

$$\%\,\text{gain in efficiency} = 100\left(\frac{1-g_h}{g_h}\right) = 100\left[\frac{Var\left(B_{ih}\right) - Var\left(\tilde{B}_{ih}\right)}{Var\left(\tilde{B}_{ih}\right)}\right].$$

A comparison of improvements associated with adjustment of the status estimates is given in Table 2-1 for the three measures discussed above. It is clear that the improvement from adjustment of the overall SVMP estimates is much greater than expected based on the theoretical calculations. While the adjustment of each individual stratum is greater than expected, it is the improvement in the flats stratum that dominates the overall improvement. It is the magnitude of the discrepancy between observed and expected improvement in the flats stratum that provided the motivation for the work reported in this chapter.

Following the retrospective adjustment of 2001 results using 2002 data, it was determined that “the dramatic improvement in the flats variance is largely due to chance associated with the random rotation of sites” (Berry et al. 2003, p.44). This is further attributed more specifically to the unmatched 2001 sites. The fact that the second application of the retrospective adjustment with a new set of unmatched sites (2002) produced a very similar result suggests that there may be another explanation.

		Variance Ratio g_h	Percent Change in Variance	Percent Gain in Efficiency
Cochran example		0.821	-18%	22%
Section 4.3.2 calculation		0.840	-16%	19%
SVMP 2001 adjusted results	overall	0.058	-94%	1640%
	flats	0.023	-98%	4160%
	narrow fringe	0.676	-32%	48%
	wide fringe	0.719	-28%	39%
SVMP 2002 adjusted results	overall	0.046	-95%	2070%
	flats	0.014	-99%	7070%
	narrow fringe	0.584	-42%	71%

Table 2-1. Comparison of theoretical and observed improvements in adjusted status estimates using three measures. The Cochran example is from Cochran (1977, Table 12.5, p.353). The section 4.3.2 calculation is a closer analogue to the SVMP design. The 2002 wide fringe stratum was not adjusted due to an insufficient number of unmatched sites, *u*=1.

2.2 Breakdown of the Adjustment Calculations for the Flats Stratum

The variance estimate for adjusted *Z. marina* area of the flats stratum is given by (Skalski 2003, Equation 29)

$$\widehat{Var}\left(\tilde{B}\right)=W^2\widehat{Var}\left(\hat{B}'_{U1}\right)+\left(1-W\right)^2\widehat{Var}\left(\hat{B}'_{M1}\right).$$

Equation 2-1

This is a weighted combination of independent estimates from the adjusted matched (subscript *M*) and unmatched (subscript *U*) sites. The calculated values for these terms in the SVMP analyses are shown in Table 2-2.

	W^2	$(1-W)^2$	$\widehat{Var}\left(\hat{B}'_{U1}\right)$ (ha^2)	Unmatched Variance Ratio	$\widehat{Var}\left(\hat{B}'_{M1}\right)$ (ha^2)	Matched Variance Ratio
2001 adjustment	0.69	0.03	875,592	0.03	4,354,544	0.14
2002 adjustment	0.83	0.01	461,340	0.02	4,640,550	0.15

Table 2-2. Terms used in Equation 2-1 to calculate variance of adjusted *Z. marina* area for the flats stratum. The estimated variances of the adjusted *Z. marina* area estimates are shown for both unmatched (subscript *U*) and matched (subscript *M*) sites. Also shown are the ratios of variance of adjusted estimates to variance of initial estimates.

The variance ratios shown in Table 2-2 indicate that the adjusted variance estimates for both matched and unmatched sites reflect improvements much greater than expected. This is true for both years. Specifically, the matched site variance ratios (0.14 and 0.15) are more than a factor of five less than the overall theoretical values (0.82 and 0.84) in Table 2-1. The unmatched cases are much more extreme with variance ratios (0.03 and 0.02) more than a factor of 30 less than the theoretical values. The difference between adjusted variances of matched and unmatched sites is compounded when the weights are considered. The unmatched result is weighted roughly 30 times more in the combined result for the flats stratum (0.69 and 0.83 for unmatched weights versus 0.03 and 0.01 for matched weights).

It is significant that the adjusted variance estimates for both matched and unmatched sites for both years are far lower than anticipated. This suggests that the reason for the strong improvements in adjusted estimates is not simply due to the chance selection of sites for rotation and the small sample size (*n*=2) for unmatched site estimates. The fact that two years of adjusted estimates produced similarly dramatic improvements suggests that while the chance selection of sites may be responsible, this chance is not negligible and must play a major role in the interpretation of results. The fact that improvements in the matched estimates, with a larger sample size (*n*=8), were also greater than anticipated suggests that other factors are important.

2.3 Assumptions Underlying Ratio and Regression Estimators

The estimate of adjusted variance for the flats stratum (Equation 2-1) is based on the estimate of adjusted *Z. marina* area, $\tilde{B}$, given by (Skalski 2003, Equation 23) as

$$\tilde{B} = W \cdot \hat{B}'_{U1} + \left(1 - W\right) \hat{B}'_{M1} .$$

This is simply the weighted sum of two independent estimates of *Z. marina* area in the flats stratum. The estimate based on the unmatched sites, $\hat{B}'_{U1}$, is a ratio estimator and the estimate based on the matched sites, $\hat{B}'_{M1}$, is a form of regression estimator.

Cochran (1977, p.153) notes that the variance estimate for ratio estimators is biased but acceptable if the sample size is at least moderately large (*n*>30) and the ratio is nearly normally distributed. The flats stratum calculations violate both of these conditions. The sample size is very small for both unmatched site (*n*=2) and matched site (*n*=8) calculations. Also, the distribution of *Z. marina* area ratio for the 14 flats sampled over 2000-2003 clearly deviates from normality when examined directly (Figure 2-1) or in the form of a stem-and-leaf diagram (Figure 2-2). The Shapiro-Wilk W statistic is significant (p=0.00599) leading to a rejection of a hypothesis of normality.

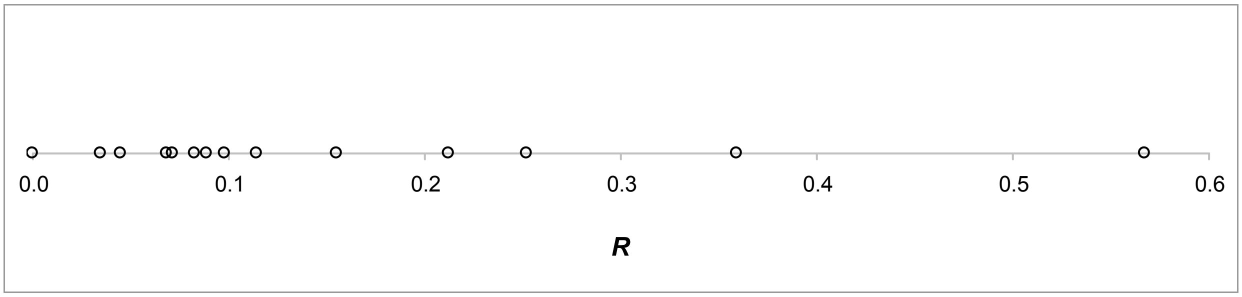


Figure 2-1. Distribution of *Z. marina* area ratio for the flat stratum sites sampled in 2000-2003. Estimates from 14 sites are depicted out of 67 in the current flats sampling frame. The most recent estimates available were used for sites sampled in multiple years.

0	034
0	67889
1	1
1	5
2	1
2	5
3	
3	5
4	
4	
5	
5	6

Figure 2-2. Stem-and-leaf diagram for the existing *Z. marina* area ratio estimates in the flats stratum.

An existing Monte Carlo study of eight small natural populations provides an indication of bias of ratio estimators with small sample size (Rao 1968; cited in Cochran 1977, p.163-164). These populations do not necessarily resemble the skewness of the flats stratum data but are generally illustrative nonetheless. The biases in the variances of ratio estimators are "much more serious in small samples" than the biases in the ratio estimators themselves and Cochran concludes from the Rao study that variance estimates are unsatisfactory at least up to n=12 (Cochran 1977, p.163; note that this contradicts Cochran's guidance discussed on p.82). The mean percent underestimation of the variance of the ratio estimators (mean of the eight study populations) found in the Rao (1968) study is shown in Table 2-3 for small sample sizes.

n	Percent Error
4	-31%
6	-23%
8	-21%
12	-18%

Table 2-3. The percent error in calculated variance of a ratio estimator derived from a Monte Carlo study of eight natural populations using small sample sizes (Rao 1968; cited in Cochran 1968, pp.163-164). Percent error is calculated as (variance estimate – true variance) / true variance.

The errors in Table 2-3 increase markedly at *n*=4 relative to *n* =6 and presumably would be significantly higher at the *n* =2 used in the unmatched site estimates. Caution must be used in applying these errors directly to the SVMP since the population distributions used by Rao may differ from the distribution of *Z. marina* area ratio.

In the case of regression estimators, Cochran (1977, p.190) similarly indicates that these estimators are generally biased but that the bias is small when the sample is large. The Monte Carlo study of Rao (1968) also presents the mean percent underestimation of variance of regression estimators for small sample sizes. These are shown in Table 2-4. These errors are significantly greater than those for ratio estimators.

n	Percent Error
6	-48%
8	-42%
12	-33%

Table 2-4. The percent error in calculated variance of a regression estimator derived from a Monte Carlo study of eight natural populations using small sample sizes (Rao 1968; cited in Cochran 1968, pp.197-198). Percent error is calculated as (variance estimate – true variance) / true variance.

2.4 Potential Effect of Bias on Flats Adjustment

Percent errors of the magnitude reported by Rao (1968) when calculating ratio and regression estimators from small sample sizes, cannot by themselves account for the discrepancy between anticipated and observed variance improvements in the flats stratum. To confirm this, a correction was applied to the adjusted variance estimates for matched and unmatched sites and then the overall flats stratum variance was recalculated.

The unmatched site variance was corrected under the assumption that it represented a 45% underestimate (rudimentary extrapolation for *n*=2 using Table 2-3). The matched site variance was corrected using the -42% value for *n*=8 in Table 2-4. Using these corrections, the variance ratios for the adjusted flat stratum variance were 0.04 (2001 adjusted) and 0.03 (2002 adjusted)–still dramatically less than the variance ratio of approximately 0.82 – 0.84 suggested by the theoretical calculations (Table 2-1). In fact, the unmatched site variance would have to be an underestimate on the order of >95% for this to completely account for the discrepancy between anticipated and observed variance improvements.

As noted above, caution is in order when applying the bias results of Rao (1968) directly to SVMP calculations. The primary concern is that the eight populations studied by Rao may have distributions that differ markedly from the distributions of *Z. marina* area and *Z. marina* area ratio in the flats stratum. The distribution of sites in the flat stratum is skewed in terms of *Z. marina* area ratio (Figure 2-1) but much more strongly skewed in terms of *Z. marina* area. Also, in the case of ratio estimators, Rao did not give bias results for the unmatched site sample size, *n*=2. It is entirely possible that the actual bias in SVMP calculations is much larger than Rao's estimates given the very low sample size and the very skewed distributions. A Monte Carlo study designed specifically to address the SVMP case could be useful in resolving the role of bias in the adjustment calculations.

2.5 A Limitation in the Theoretical Model – Measurement Error

This section explores another potential contributing factor to the discrepancy between observed and expected improvements in precision associated with retrospective adjustment. This factor has to do with the fact that there are two distinct sources of uncertainty in the observed improvements but the expected improvements are based on a theoretical model that considers only one of these sources (J. Skalski, pers. comm.).

One source of uncertainty is associated with using a sample of sites to represent the whole population of sites within the stratum. We will refer to this uncertainty as sampling error. It is explicitly considered in the model used to derive the expected improvement with retrospective adjustment.

The second source of uncertainty arises from the fact that the value used for *Z. marina* at each sample unit (or site) is a measurement based on line intercept sampling. There is uncertainty associated with this measurement that we will refer to as measurement error. This is not considered in the model used to derive the expected improvement. This is because the model emerged from the survey sampling literature (e.g. Cochran 1977) where the information used for each sample unit is frequently known without error. Cochran discusses only briefly the issue of measurement error and its implications (Cochran 1977, p.377)

It is entirely possible from a theoretical perspective that greater variance in initial estimates provides an opportunity for greater improvement in precision in the adjusted estimates. Section 1.4 demonstrated this phenomenon in an analogous situation where the dramatically higher initial variance when flats11 was present in the sample pool led to dramatically higher improvements in precision with adjustment. This flats11 case, of course, highlights two scenarios with contrasting sampling error rather than measurement error, but the principal may still hold.

This limitation in the underlying theoretical model suggests that the discrepancy between observed and predicted improvements in precision may be due, at least in part, to artificially low predictions. If the underlying model took into account measurement error, it is possible that the predicted improvement in precision would be more in line with the SVMP observations.

The relative contributions of sampling error and measurement error to the overall flats stratum variance may indicate the role that measurement error plays in the discrepancy between observed and predicted improvements in precision. The estimates of flats variance in the 2000-03 SVMP results are broken down into these two components.

For the flats stratum, the estimate of stratum *Z. marina* area is a ratio estimate given by (Skalski 2003, Equation 9)

$$\hat{B}\big|x_i = \left[\frac{\sum_{j=1}^{n}\hat{X}_j}{\sum_{j=1}^{n_2} a_j}\right]\cdot A$$

Equation 2-2

where

$\hat{X}_j$ = estimate of *Z. marina* area at flats site *j*,

a_j = site area of flats site *j*,

A = total area in the flats stratum, i.e. the sum of all site areas in the stratum

n = number of sites sampled in the flats stratum.

Equation 2-2 can be inserted into the total variance equation (see Appendix E) to look at components of variance of the flats stratum *Z. marina* area estimator. Skalski (2003, Appendix B2) has shown that the result can be reduced to

$$\widehat{Var}\left(\hat{B}\right) = N^2\left(1-\frac{n}{N}\right)\frac{\sum_{j=1}^{n}\left(\hat{X}_j - a_j\hat{R}\right)^2}{n\left(n-1\right)} + \frac{N\sum_{j=1}^{n}\widehat{Var}\left(\hat{X}_j\big|x_j\right)}{n}$$

Equation 2-3

where

N = total number of sites in the flats sampling frame.

Equation 2-3 can be rearranged to a form that emphasizes the two components, giving

$$\widehat{Var}\left(\hat{B}\right) = \frac{N^2}{n}\left(1-\frac{n}{N}\right)s^2_{\hat{X}_j} + \left(N\right)\overline{\widehat{Var}\left(\hat{X}_j\big|x_j\right)}$$

Equation 2-4

where

$$s^2_{\hat{X}_j} = \frac{\sum_{j=1}^{n}\left(\hat{X}_j - a_j\hat{R}\right)^2}{\left(n-1\right)}.$$

In the first term of Equation 2-4, $s^2_{\hat{X}_j}$ is the mean square deviation of the flat sites from the mean and this term represents sampling error. In the second term, $\overline{\widehat{Var}\left(\hat{X}_j\big|x_j\right)}$ is the mean site-level variance and this term represents measurement error.

If measurement error is assumed to be zero, Equation 2-4 reduces to the analogous ratio estimator given in Cochran (1977, Equation 6.9) as expected, since Cochran's estimators are based on a model with no measurement error.

Table 2-5 shows the breakdown of the 2000-03 flats stratum variance estimates based on Equation 2-4. These results indicate that sampling error dominates the overall variance. In 2000, before flats11 entered the sample pool, sampling error was more than an order of magnitude greater than measurement error. After flats11 entered the sample pool, sampling error was consistently three orders of magnitude greater than measurement error.

These results do not preclude the possibility that the lack of measurement error in the underlying model contributes to the discrepancy between observed and predicted improvement in precision with adjustment. However, the small contribution of measurement error to the overall variance suggests that this effect is at most a minor consideration.

A more in depth Monte Carlo study based on the model flats dataset would provide a more definitive characterization of the role of measurement error in the anomalous retrospective adjustment results.

	2000	2001	2002	2003
sampling error (ha^2)	495,246	31,009,689	30,066,407	27,488,936
measurement error (ha^2)	15,858	24,905	25,613	17,666
total variance (ha^2)	511,104	31,034,593	30,092,020	27,506,602

Table 2-5. Breakdown of 2000-03 flats stratum variance into sampling error and measurement error components.

2.6 Discussion and Summary

A very strong improvement in variance is seen in 2001-2003 SVMP results associated with the retrospective adjustment. This improvement is inconsistent with the theoretical predictions of Cochran (1977). This discrepancy is particularly troubling because Skalski (2003) has shown that this strong improvement in variance is necessary for the SVMP to reach its *Z. marina* trend detection limit target. As seen in the next chapter, the new flats stratification introduced in 2004 strongly improves precision, perhaps even to the point where it can be seen as an alternative to the retrospective adjustment calculation. Nevertheless, it is important to examine the retrospective adjustment calculations to ensure that the estimates based on the 2001-2003 SVMP data are robust.

This chapter has reviewed the theoretical calculations of improvements in variance and compared these to the SVMP observations of improvements in 2001 and 2002 *Z. marina* area results. The observed improvements in overall *Z. marina* area estimates are far greater than predicted and this is almost entirely due to the improvement in the flats stratum.

The extreme improvement in flats stratum estimates occurs in both 2001 and 2002 adjustments and in both unmatched site and matched site components of the estimate, although the unmatched estimate is more extreme. The fact that this has occurred in estimates from two years suggests that this was not a rare occurrence of low probability and needs to be more fully explained. The fact that the matched site estimates on their own represent an extreme improvement suggests that the n=2 sample size for the unmatched sites is not the sole explanation.

The ratio and regression estimators used in the retrospective adjustment are known to be biased—strongly so at low sample sizes. An existing Monte Carlo study has quantified these biases for specific datasets but the reported values cannot fully explain the extreme improvements seen in the flats stratum (Rao 1968). However, the Monte Carlo study did not consider sample sizes as low as the n=2 case for unmatched site estimates. Also, the distribution of *Z. marina* area and area ratio in the flats stratum are strongly skewed and it is not clear if the Rao study considered such departures from normality.

The presence of measurement error in the SVMP study may also contribute to the discrepancy between observed and predicted improvements in precision. The predictions are based on an underlying model that does not consider measurement error. In the 2000-03 results, measurement error is only a minor component of the overall variance, but it may still be a contributing factor to the discrepancy.

The extreme improvements in flats stratum variance associated with adjustment are not definitively explained, relative to the theoretical predictions. Three possible explanations have been reviewed:

- low sample size for regression and ratio estimators, which has been demonstrated to introduce significant bias;
- strong departures from normality—although there is no underlying assumption of normality (except in the construction of confidence intervals) very skewed distributions may further bias interval estimates for small sample sizes;
- presence of measurement error, which is not considered in the theoretical model on which the predictions are based.

Future work is planned to perform a Monte Carlo study to evaluate bias in SVMP estimates.

3 Ability to Detect Long-Term Change

3.1 Introduction

One of the strengths of the Submerged Vegetation Monitoring Project (SVMP) is that an overall performance measure is explicitly integrated into the design. It is a statistical measure of the ability to detect long-term change in *Z. marina* area over Puget Sound. The ability to detect a 20% decline over a 10-year period is the broad target set by SVMP staff for this performance measure (Berry et al. 2003, p.2). Skalski (2003) presents a detailed methodology for evaluating this performance measure in the form of the statistical power of 5-year and 10-year tests for decline in *Z. marina* area. This chapter builds upon that work to develop a more thorough characterization of the power of these tests.

The concepts of statistical power and significance are central to this chapter. In simple terms, the significance is a measure of the confidence in a test result that indicates there has been change. Power is a measure of the ability of a test to detect a specific level of change. These two characteristics of a statistical test are inversely related. For example, if a test is conducted with a high level of significance, we are very confident in a result that indicates a change has taken place (low probability of false positives), but at the same time our ability to detect change is reduced because it takes a bigger change to trigger a positive result.

The significance level of a test is denoted α, where $\alpha=0.05$ indicates there is a 5% chance that a test will indicate change when there really is none (false positive result). The quantity $100(1-\alpha)$ is the level of confidence (e.g. 95% confidence level) that, assuming there is no change, the test will correctly give a result of no change.

The power of a test is denoted $1-\beta$ and is always given for a specific level of real change. For example, if there is a real change of 20%, a power of $1-\beta=0.80$ would indicate that there is an 80% chance that the test would successfully detect that change.

The statistical power associated with the SVMP tests for long-term change is a complex function of many parameters. These parameters include details of the *Z. marina* population, details of the sampling design, the magnitude of the decline to be detected and the significance level, α, used in testing. However, a performance measure must be presented in a simple and easily understood form in order to be useful in communicating to a broader audience. Typically, this means presenting defensible generalizations while not revealing the complexity of the analyses. This in itself requires a solid understanding of the analyses underlying performance measures.

Long-term change analysis will become more important in the SVMP as the data record is extended in the coming years. It will be critical for SVMP staff to understand the power associated with these analyses in order to help interpret results in a resource management, or decision-making context.

This chapter has two main objectives:

1. Develop a richer understanding of the statistical power of the SVMP tests for long-term decline. Specifically, this chapter evaluates power under different scenarios of *Z. marina* decline and sampling design parameters.
2. Reevaluate the central performance measure of the SVMP – the ability of the monitoring project to detect a 20% decline in *Z. marina* area in Puget Sound over a 10-year period.

While the SVMP target detection limit is associated with a 10-year test for decline, the 5-year test will be the first to be implemented and it will be implemented many times before application of the 10-year test is possible. This chapter starts with a comparison of power associated with the 5- and 10-year tests under scenarios of steady decline in *Z. marina*. Then, the sensitivity of power to the temporal pattern of decline is presented. Finally, the potential of additional project resources to improve power through the addition of sampling sites is explored.

This chapter does not address the ability to detect change using the paired site techniques developed by Skalski (2003). This includes testing for year-to-year change at both the sound-wide scale and the regional scale. Change detection at the site-level is also not addressed.

It is also important to note that the tests for change described here do not discriminate between ‘natural’ change and change associated with anthropogenic activities although this distinction is clearly important from a resource management perspective. The relative role of natural and anthropogenic factors must be discerned from consideration of SVMP results in conjunction with a broader analysis that considers climatic signals and patterns of anthropogenic activities in marine waters and adjoining watersheds.

3.2 Comparison of 5- and 10-Year Tests Under Steady Decline

Skalski (2003) presented one-tailed *t*-tests for detecting declines in *Z. marina* area over 5- and 10-year records for Puget Sound. These test the null hypothesis of no decline by testing whether the slope, *b*, of a regression line on the data points is zero. These are one-tailed tests that only test for decline:

$$H_0 : b \geq 0$$

$$H_a : b < 0$$

Skalski (2003) presents a method for explicitly calculating the power of these tests. The method assumes that the test statistic follows a noncentral-F distribution under the alternate hypothesis of decline in *Z. marina* (Skalski and Dobson 1992, pp.134-135, pp.217-221; Zar 1999, pp.135-136, p.190). The noncentral-F distribution is defined by two values for degrees of freedom, ν_1 and ν_2, and the noncentrality parameter, Φ_{ν_1,ν_2}. In this application $\nu_1 = 1$.

The noncentrality parameter is calculated as

$$\Phi_{1,3} = \sqrt{5} \cdot \frac{|\Delta|}{\sqrt{\mathrm{CV}^2}}$$

Equation 3-1

for the test for 5-year decline (Skalski, 2003, equation 48) and

$$\Phi_{1,8} = \sqrt{41.25} \cdot \frac{|\Delta|}{\sqrt{\mathrm{CV}^2}}$$

Equation 3-2

for the test for 10-year decline (Skalski, 2003, equation 49) where

Δ = annual fractional decline in *Z. marina* area over Puget Sound,
CV = coefficient of variation of the estimates of *Z. marina* area (assumed constant across years).

Once the noncentrality parameter has been calculated for a particular scenario, the associated statistical power is read directly from a noncentral F table (Skalski and Dobson, 1992). The resulting power applies to scenarios with steady decline over the 5- and 10-year test periods.

Figure 3-1 presents results that show how power varies as a function of the level of *Z. marina* decline and as a function of the level of significance, α, used in the test. These results are based on a nominal value of CV=0.07. This value is generally representative of the CV associated with the adjusted estimates of *Z. marina* area in Puget Sound in results from 2001 (CV=0.075) and 2002 (CV=0.063) as well as simulated results under the new 2004 flats sampling strata (CV=0.079, Table 3-3).

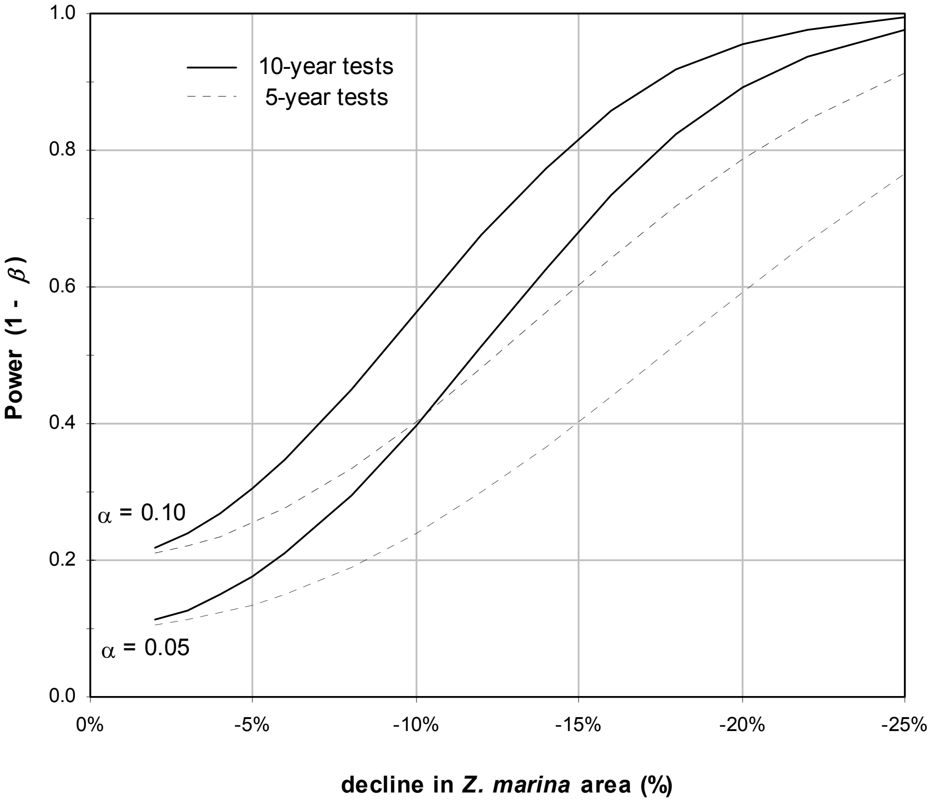


Figure 3-1. Power of 5- and 10-year tests to detect decline in *Z. marina* area over Puget Sound as a function of level of decline over the test period. Results are shown for two levels of significance α. These results are based on Equation 3-1 and Equation 3-2 and the noncentral F tables in Skalski and Dobson (1992) with CV=0.07 for all years.

The following key points emerge from Figure 3-1:

- The tests do not have sufficient power to detect small levels of decline. The decline must be approximately 14-17%[4] over 10 years or 20-27% over 5 years for the tests to have a power of $1-\beta=0.80$.
- The 10-year test is significantly more powerful at the levels of decline necessary (e.g., more severe than –15%) to achieve satisfactory levels of power.
- There is a generally a large gain in power associated with raising the significance from $\alpha=0.05$ to $\alpha=0.10$. For instance, in a scenario of a 20% decline, a 5-year test with $\alpha=0.05$ has moderately low power ($1-\beta \approx 0.6$) while a test with $\alpha=0.10$ has substantially higher power ($1-\beta \approx 0.8$).

Of course Figure 3-1 reflects a fundamental tradeoff between the objective of minimizing the probability α of false positive results (test indicates change where there is no true change) and the desire to maximize the probability, $1-\beta$, of correctly identifying a real change. Figure 3-1 suggests that it will be necessary to accept a substantial risk of false positives ($\alpha \geq 0.1$) in order to reach a meaningful level of power in the 5-year test, which will be the first test to be implemented.

---

[4] Minus signs on values of decline are dropped in some instances for simplicity. Negative values of change are implied for declines.

The comparison between the 5- and 10-year tests in Figure 3-1 is somewhat confounded by differences in spreading a given level of decline, say 20%, over different time intervals. Taking this effect into account further highlights the strength of the 10-year test. For example, a 20% decline over the test period corresponds to an annual decline of 5.4% over five years but only 2.4% over ten years. A given level of decline over the test period (as shown in Figure 3-1) provides for a more sensitive test in the 10-year case, in terms of detecting lower *rates* of decline. Figure 3-2 shows the same power curves of Figure 3-1 but with the x-axis transformed to annual rate of decline in *Z. marina* associated with the 5- and 10-year declines shown in Figure 3-1.

On the basis of annual decline in *Z. marina* area, the 10-year test clearly has much higher power then the 5-year test except for very low levels of decline (less than –0.5%).

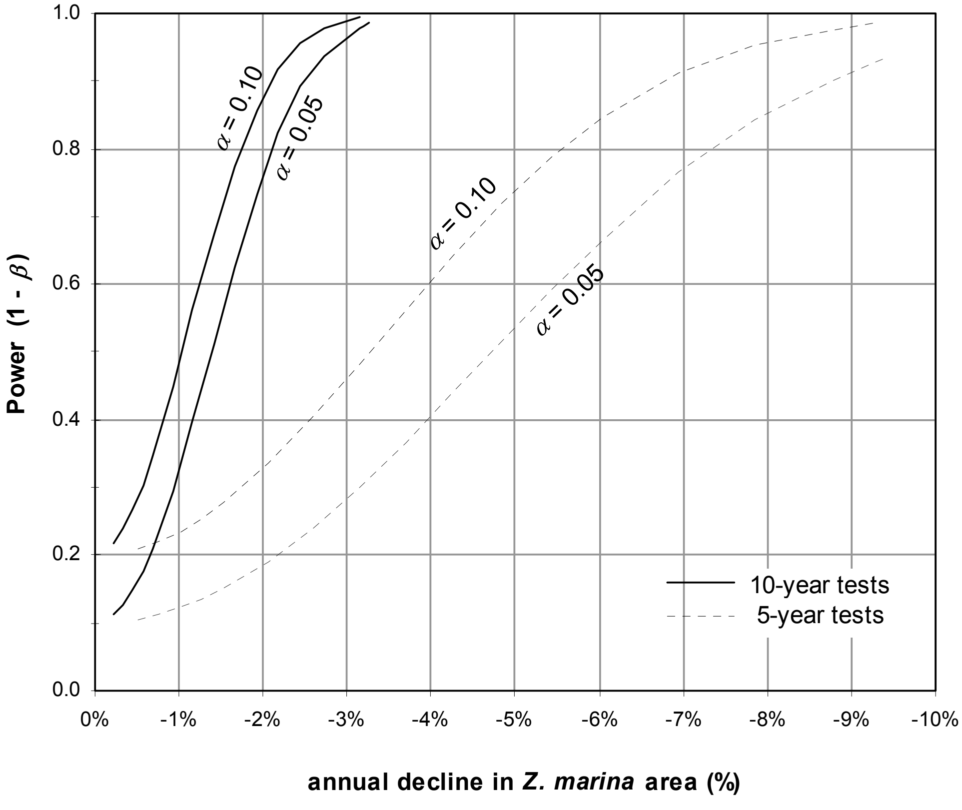


Figure 3-2. Power of 5- and 10-year tests with respect to annual rate of decline in *Z. marina* area over Puget Sound. The annual rate of decline is assumed to be constant over the five or ten year test period. Results are shown for two levels of significance α. These results are based on Equation 3-1 and Equation 3-2 and the noncentral F tables in Skalski and Dobson (1992) with CV=0.07 for all years.

The 5-year test is also more sensitive to the CV associated with the estimate of *Z. marina* area (Figure 3-3). As CV increases, both 5- and 10-year tests lose power rapidly but the 5-year test loses power more rapidly (Figure 3-3). Furthermore, the current estimate of CV ≈ 0.07 is at the steepest part of the power-CV curve. Any increases in CV estimates in the coming years will lead to dramatically lower power. For example, if the 5-year average

CV increases from the current estimate of ~0.07 to 0.1, the power of the 5-year test to detect a 20% decline drops from approximately 0.8 to 0.6.

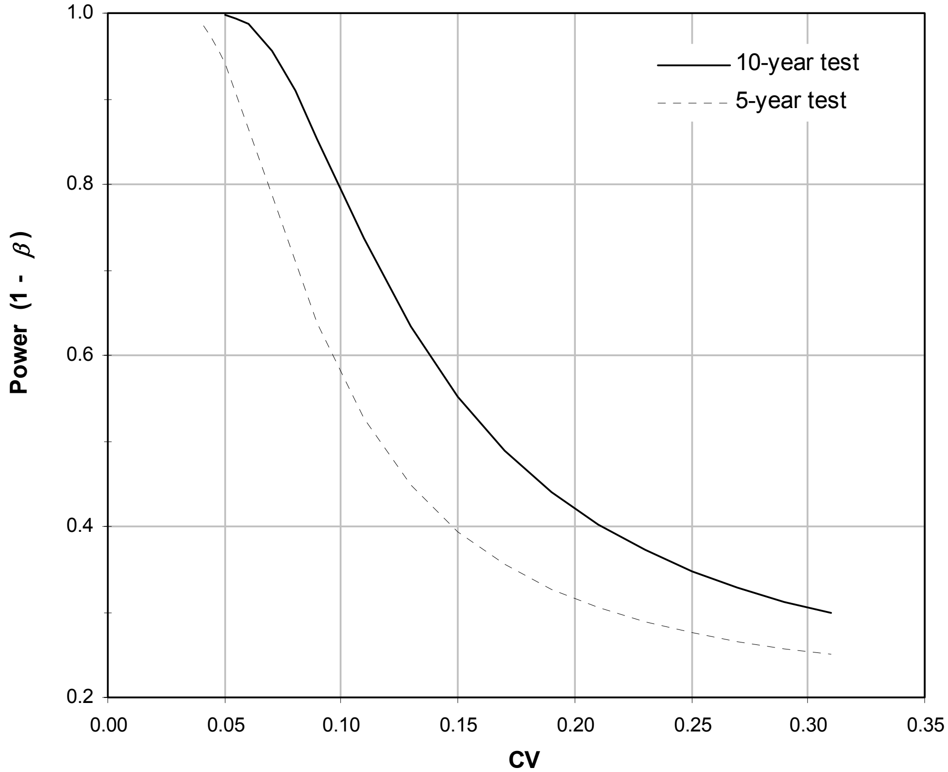


Figure 3-3. Power of 5- and 10-year tests to detect a 20% decline in *Z. marina* area over Puget Sound for a range of CV associated with the *Z. marina* area estimates. Calculated using Equation 3-1 and Equation 3-2 and the noncentral F tables in Skalski and Dobson (1992) with α=0.10.

In summary, these results show that the 10-year test for decline in *Z. marina* area is superior to the 5-year test in terms of

a) the statistical power associated with detecting a given level of decline,
b) its ability to detect lower *rates* of decline with a given level of statistical power,
c) the lower sensitivity of its statistical power to lower levels of precision.

The 5-year test of course has the important benefit of lower data requirements and delivery of early results. Power curves, such as those presented here, will be very useful in interpreting the results of the 5-year tests when the required data record becomes available.

3.3 Sensitivity of Tests to Temporal Pattern of Decline

The results in the previous section are based on the assumption of a uniform rate of decline in *Z. marina* area over the 5- or 10-year test period. While this is a plausible scenario in a general sense, the possibility of brief episodes of decline in an otherwise stable system is also of interest. Disturbances such as climatic or oceanographic anomalies or a pathogen outbreak could conceivably lead to discrete periods of relatively sharp declines.

Norris et al. (2001) discussed the sensitivity of the tests to the pattern of decline but the two scenarios investigated (exponential vs. linear decline) were similar and there was little difference in power to detect change under these scenarios. Here this concept is explored further by focusing specifically on the sensitivity of the 10-year test to a discrete three-year period of decline embedded within the 10-year data record.

The specific objective here is to assess how the power of the 10-year test changes with the timing of a three-year event that produces a 20% loss of *Z. marina* within Puget Sound. Since calculation of the noncentrality parameter as used in section 3.2 and 3.3 is problematic with such an irregular time series, we used a Monte Carlo approach.

3.3.1 Monte Carlo Approach

This approach was based on the simulation of a large number (1000) of 10-year data records with "known" *Z. marina* decline. The 10-year test was applied to each data record and power was estimated as the fraction of tests that successfully detected this known decline.

Two signals were superimposed to simulate each of the one thousand 10-year datasets used for a single Monte Carlo estimate:

a) A sequence of ten annual *Z. marina* area values with the specific pattern and level of overall decline to be tested. This signal represented "truth" and was identical for each of the 1000 datasets.
b) Stochastic noise that reflected sampling error (associated with random selection of sites) and measurement error (associated with line transect sampling at each site) as estimated from the 2000-2003 SVMP data.

A stochastic term was used to perturb each annual data point from the "true" value. Each stochastic term was an observed value of a normal random variable with mean zero and variance equal to the approximate variance of the 2000-2003 estimates of sound-wide *Z. marina* area.

Normal random numbers were generated by applying a Box-Muller transform (Wilkinson 2004) to uniform random numbers generated by the Mersenne Twister algorithm (Matsumoto and Nishimura 1998).

3.3.2 Test of Monte Carlo Approach

Before investigating alternate patterns of decline, the Monte Carlo approach was used to reproduce a portion of the analysis that is presented in Figure 3-1. This provided a direct comparison to power estimated with the noncentral F distribution and served as a test of the Monte Carlo calculations.

Figure 3-4 shows an example that illustrates the relative magnitudes of "true" decline and stochastic noise for a case of uniform decline totaling 20% over 10 years. The error bars represent the standard errors of each estimate and correspond to a CV of 0.07 associated with the initial *Z. marina* area. These standard errors were used to constrain the stochastic noise in the simulated datasets so that in approximately 68% of the 1000 datasets generated for this case, the data point a given year fell within the range of the error bars — in 42% of the datasets the data point was outside the range of the error bars.

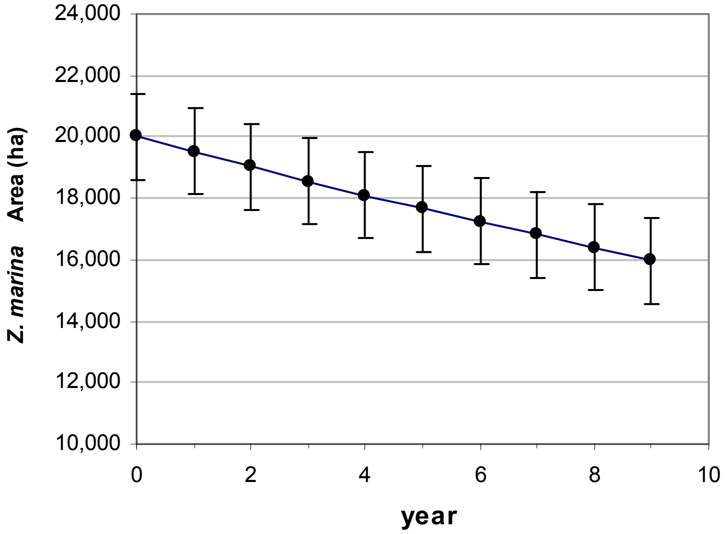


Figure 3-4. Hypothetical decline in sound-wide *Z. marina* area totaling 20% over 10 years. Error bars show standard error associated with CV=0.07 associated with the initial 20,000 ha value.

The stochastic noise, or uncertainty, is clearly of sufficient magnitude to produce simulated datasets whose trend deviates substantially from the "true" trend. Figure 3-5 shows the distribution of 10-year trends in the 1000 simulated datasets based on the slope of regressions lines fit to each of the datasets.

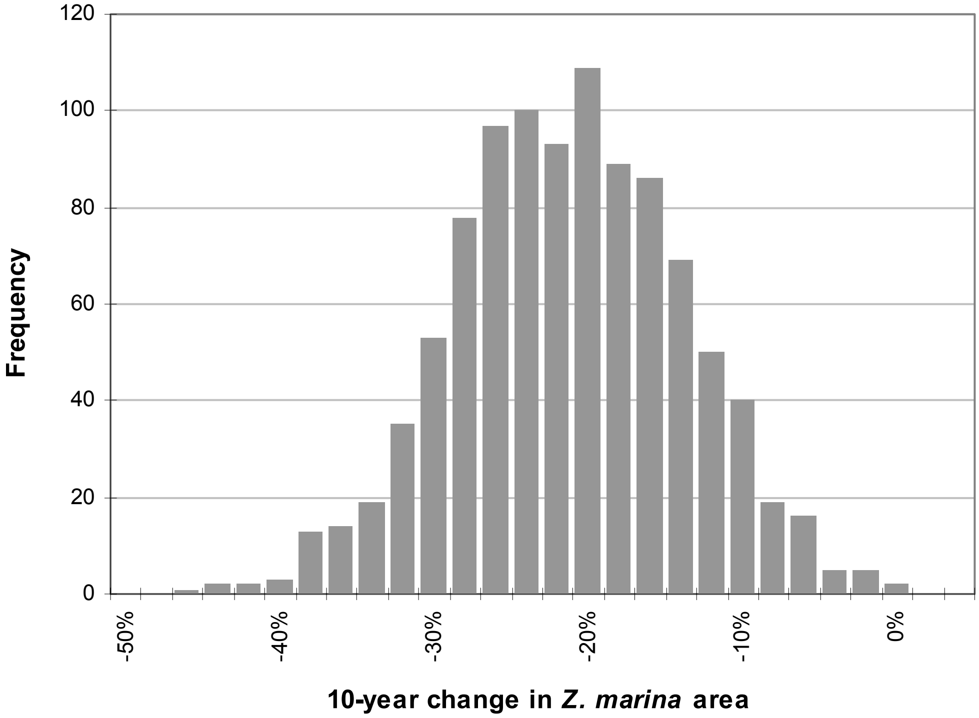


Figure 3-5. Frequency distribution of sound-wide 10-year *Z. marina* area change from 1000 Monte Carlo simulations. Change is based on slope of regression lines fit to each simulated dataset generated from the 20% decline scenario illustrated in Figure 3-4. Note that –20% is the mode of the distribution.

In order to test the power of the 10-year test to detect a particular scenario of "known" change, each of the 1000 datasets simulated under this scenario are subjected to a *t*-test using the statistic (Skalski 2003, Equation 46)

$$t_{m-2} = \frac{\left|\hat{\beta} - 0\right|}{\sqrt{\dfrac{\text{MSE}}{\sum_{i=1}^{m}\left(t_i - \overline{t}\right)^2}}}$$

where

$$\hat{\beta} = \text{slope of regression line}$$

$$\sum_{i=1}^{m}\left(t_i - \overline{t}\right)^2 = 82.5 \text{ for } t_i = \left(0,1,2,\ldots,9\right)$$

$$\text{MSE} = \text{mean square error of the regression.}$$

The Monte Carlo estimate of power for the test is simply the percentage of the 1000 tests that successfully detect the "known" decline.

This technique was applied to fourteen scenarios representing the range of *Z. marina* decline shown in Figure 3-1. The CV was set to 0.07 and $\alpha = 0.1$ was used for the testing so that results would be directly comparable to the corresponding results in Figure 3-1 derived with the noncentral F distribution. Both sets of results are shown in Figure 3-6.

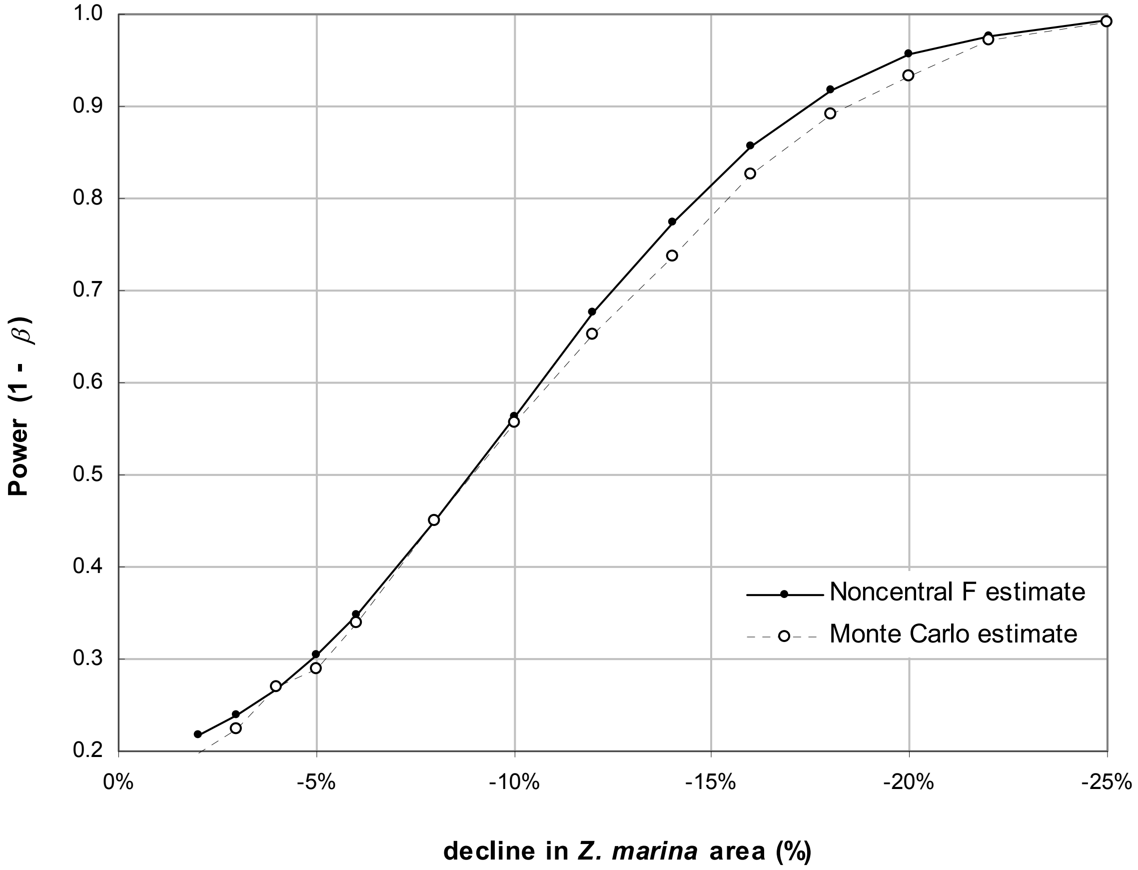
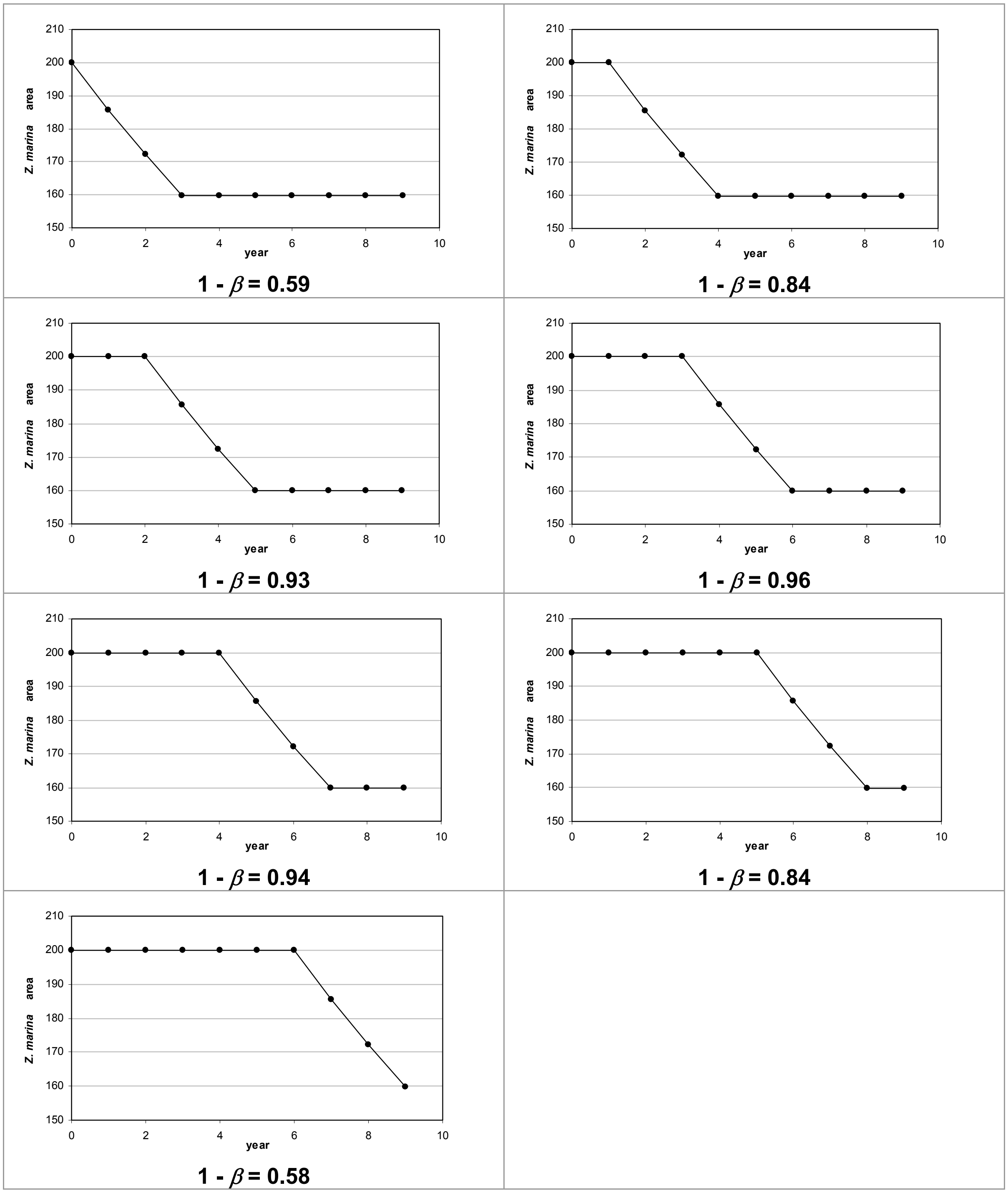


Figure 3-6. Comparison of power estimated by the Monte Carlo simulations and the noncentral F distribution for the 10-year test of *Z. marina* decline. All estimates were calculated with CV=0.07 and $\alpha = 0.10$.

The Monte Carlo and noncentral F estimates respond similarly to the level of decline in *Z. marina*. There is a consistent discrepancy in the range of 10-22% decline but the maximum discrepancy in power is only 0.038. This suggests that the two approaches can be considered equivalent for the purposes of characterizing gross patterns of response in power.

3.3.3 Results

The power of the 10-year test for decline conducted at α=0.05 was evaluated for seven hypothetical scenarios. In each scenario there is a 20% decline in *Z. marina* area over the 10-year data record, but the period of decline is restricted to three years. Except for these three years there is no change to *Z. marina* area. The six scenarios differ in the timing of the three years of decline within the 10-year record. Results of this study are shown in Figure 3-7.



1 - β = 0.59

1 - β = 0.84

1 - β = 0.93

1 - β = 0.96

1 - β = 0.94

1 - β = 0.84

1 - β = 0.58

Figure 3-7. Sensitivity of power of 10-year test to detect change to seven scenarios of 20% change. All test were conducted at a=0.05. Power results are based on Monte Carlo estimates. Each plot represents one scenario of *Z. marina* decline. Power, 1-β, is shown below each plot.

In the results shown in Figure 3-7 it is clear that the pattern of decline of *Z. marina* area has a significant effect on the power of the 10-year detection test, ranging from 0.58 to 0.96 for these seven scenarios. Power is highest when quick (3-year) decline is in the middle of the 10-year data record and much lower when the decline is at the beginning or end of the record.

These results suggest that given the role of chance and uncertainty, the chances of detecting a 20% decline that occurs in 3 years in an otherwise stable record, ranges from roughly 95% to roughly 60% depending the timing of the decline within the record.

3.4 Sensitivity of Tests to Number of Sites

There is a fundamental trade-off between the cost of adding additional sampling sites and the benefit of a greater ability to detect change when additional sites are sampled. In order to make informed decisions on setting the number of sites, it is necessary to have a clear understanding of the relationship between sample size and the statistical power of the 5- and 10-year tests to detect *Z. marina* decline.

This section presents the results of rudimentary calculations to characterize this relationship based on estimates of power (Equation 3-1 and Equation 3-2) that rely on an important simplifying assumption. The assumption is that sample variances in the 2000-2003 data are representative of all samples of all sizes.

The simplification is conceptually equivalent to using the sample variance from one sample under simple random sampling, s^2, to investigate the effect of sample size, n, on the standard error of the mean

$$s_{\bar{x}}^2 = \frac{s^2}{n}.$$

In this simple case, n is varied with fixed s^2 to estimate the effects of sample size on the variance of the mean $s_{\bar{x}}^2$. While the concept is the same, the actual calculations presented here are more complex because we have a stratified design with more complicated estimators.

3.4.1 Approach for Testing Effects on Adjusted Estimates

We will estimate statistical power to detect a linear 20% decline in *Z. marina* in Puget Sound under different scenarios of sample size when the adjusted estimators are used. We evaluate both 5-year and 10-year tests for decline. The 2000-2003 SVMP results are used to calculate various terms in these calculations. We use the new flats stratification described in section 1.3.5.

We first present the general framework for the calculations and then present detailed development of estimators for the flats, narrow fringe and wide fringe strata that are functions solely of sample size. Results are then given for various sample size scenarios.

We will use Equation 3-1 and Equation 3-2 to estimate the noncentrality parameter and noncentral F tables to estimate power under different sample size scenarios. We must first develop a relationship between the number of sites sampled, n, and the CV associated with the adjusted estimate of sound-wide *Z. marina* area.

The CV is given by

$$CV = \frac{\sqrt{Var\left(\tilde{B}_T\right)}}{\tilde{B}_T}$$

Equation 3-3

where $\tilde{B}_T$ is the adjusted estimate of total *Z. marina* area in the Puget Sound study area. Only the numerator will have a dependence on *n*. The 2002 estimate of $\tilde{B}_T$ is used here for the value of the denominator,

$$\tilde{B}_T = 20,389\,\text{ha}\,.$$

The variance of $\tilde{B}_T$ is given by (cf. Skalski 2003, Equation 13)

$$Var\left(\tilde{B}_T\right) = \widehat{Var}\left(\hat{B}_c\right) + \widehat{Var}\left(\hat{B}_\zeta\right) + \widehat{Var}\left(\tilde{B}_{SPR}\right) + \widehat{Var}\left(\tilde{B}_{fr}\right) + \widehat{Var}\left(\tilde{B}_{fw}\right),$$

Equation 3-4

where

$\hat{B}_c$ = initial estimate of *Z. marina* area for the core stratum,

$\hat{B}_\zeta$ = initial estimate of *Z. marina* area for the persistent flats stratum,

$\tilde{B}_{SPR}$ = adjusted estimate of *Z. marina* area for the rotational flats stratum,

$\tilde{B}_{fr}$ = adjusted estimate of *Z. marina* area for the narrow fringe stratum,

$\tilde{B}_{fw}$ = adjusted estimate of *Z. marina* area for the wide fringe stratum.

Only the latter three terms of Equation 3-4 will have a dependence on *n* since only these strata are subject to random sampling. The 2002 result is used for the estimate of variance in the core stratum,

$$\widehat{Var}\left(\hat{B}_c\right) = 24,567\,\text{ha}^2\,.$$

The variance in the ζ-flats stratum is simply the sum of the measurement errors of the three sites in this stratum (cf. Skalski 2003, Equation 6),

$$\widehat{Var}\left(\hat{B}_\zeta\right) = \widehat{Var}\left(\hat{B}_{11}\right) + \widehat{Var}\left(\hat{B}_{12}\right) + \widehat{Var}\left(\hat{B}_{20}\right),$$

where the three terms on the right represent variance estimates of site-level *Z. marina* area at flats11, flats12 and flats20, respectively. Since we do not have an estimate for flats12 in the 2000-2003 SVMP data, we will extrapolate the 2002 flats11 result to flats12 and use the 2002 flats20 result giving

$$\begin{aligned}\widehat{Var}\left(\hat{B}_\zeta\right) &= 2\left[\widehat{Var}\left(\hat{B}_{11}\right)\right] + \widehat{Var}\left(\hat{B}_{20}\right)\\ &= 2\left[2,837\,\text{ha}^2\right] + 836\,\text{ha}^2\\ &= 6,510\,\text{ha}^2\end{aligned}$$

Variance Estimator for Main Flats Stratum

The variance estimate in Equation 3-4 associated with the adjusted estimate for the main flats stratum is (Skalski 2003, Equation, 30)

$$\widehat{Var}\left(\tilde{B}_{SPR}\right)=\frac{\widehat{Var}\left(\hat{B}'_{U1}\right)\cdot\widehat{Var}\left(\hat{B}'_{M1}\right)}{\widehat{Var}\left(\hat{B}'_{U1}\right)+\widehat{Var}\left(\hat{B}'_{M1}\right)}$$

Equation 3-5

where

$\widehat{Var}\left(\hat{B}'_{U1}\right)$ = variance estimate associated with the estimate of *Z. marina* area in the main flats stratum based only on unmatched sites,

$\widehat{Var}\left(\hat{B}'_{M1}\right)$ = variance estimate associated with the estimate of *Z. marina* area in the main flats stratum based only on sites that are matched in the subsequent year.

The variance of $\hat{B}'_{U1}$ is estimated by the expression (Skalski 2003, Equation 25)

$$\widehat{Var}\left(\hat{B}'_{U1}\right)=N^2\left(1-\frac{u_1}{N}\right)\frac{\sum_{j=1}^{u_1}\left(\hat{X}_{j1}-a_{j1}\hat{R}_{U1}\right)^2}{u_1\left(u_1-1\right)}+\frac{N\sum_{j=1}^{u_1}\widehat{Var}\left(\hat{X}_{j1}\middle|x_{j1}\right)}{u_1}$$

Equation 3-6

where the sums are only over unmatched sites and

N = total number of sites in the main flats sampling frame,

u_1 = number of unmatched sites in the main flats stratum,

$\hat{X}_{j1}$ = estimate of *Z. marina* area at flat site *j*,

a_{j1} = area of flat site *j*,

$\hat{R}_{U1}$ = estimate of *Z. marina* area ratio based on unmatched sites.

Equation 3-6 can be rearranged to a form more appropriate for this application resulting in the expression

$$\widehat{Var}\left(\hat{B}'_{U1}\right)=N^2\left(1-\frac{u_1}{N}\right)\frac{s^2_{\hat{X}_{U1}}}{u_1}\quad+\quad N\ \overline{\widehat{Var}\left(\hat{X}_{U1}\middle|x_{U1}\right)}$$

Equation 3-7

where

$$s^2_{\hat{X}_{U1}}=\frac{\sum_{j=1}^{u_1}\left(\hat{X}_{j1}-a_{j1}\hat{R}_{U1}\right)^2}{\left(u_1-1\right)}.$$

Under our simplifying assumption that sample variances from 2000-2003 are representative of all samples and sample sizes, the sample estimate of sampling variance, $s^2_{\hat{X}_{U1}}$, is not a function of sample size. Similarly, the sample estimate of mean measurement error (i.e. mean variance of site *Z. marina* estimates associated with line transect sampling at the site-level), $\overline{\widehat{Var}\left(\hat{X}_{U1}\middle|x_{U1}\right)}$, is not a function of sample size, i.e. number of sites.

Consequently, the only dependence of $\widehat{Var}\left(\hat{B}'_{U1}\right)$ on sample size (specifically, the number of unmatched sites, u_1) is shown explicitly in Equation 3-7. There are no implicit dependencies.

We can use 2000-2003 SVMP data to calculate nominal values for the fixed terms $s^2_{\hat{X}_{U1}}$ and $\overline{\widehat{Var}\left(\hat{X}_{U1}\middle|x_{U1}\right)}$. To do this, however, we must construct a two-year flats dataset by selecting sites from the 2000-2003 data. This is necessary since we will need two years in order to calculate matched-site terms (next section) and the 2000-2003 data do not have two consecutive years that meet our criteria. Specifically we need a sample of 10 sites (eliminates 2000 data) that does not include flats11 or flats20 (eliminates 2001-2003) and has eight matching sites in the sample from the following year. Table 3-1 shows the sample units (sites) that were selected for this constructed dataset.

Year 1	Year 2
flats28 (2002)	
flats60 (2002)	
flats47 (2000)	flats47 (2001)
flats53 (2000)	flats53 (2001)
flats18 (2002)	flats18 (2003)
flats35 (2002)	flats35 (2003)
flats43 (2002)	flats43 (2003)
flats62 (2002)	flats62 (2003)
flats10 (2002)	flats10 (2003)
flats37 (2002)	flats37 (2003)

Table 3-1. Artificial two-year flats dataset constructed from actual 2000-2003 SVMP results. The actual year of sampling is shown in parentheses. This artificial dataset meets the following criteria: sample size of ten in year 1; eight sites matched in year 2; flats11 and flats20 are not included.

Based on this dataset, the desired results are

$$s^2_{\hat{X}_{U1}} = 210.1\,\text{ha}^2$$

$$\overline{\widehat{Var}\left(\hat{X}_{U1}\middle|x_{U1}\right)} = 58.5\,\text{ha}^2.$$

Equation 3-7 then becomes

$$\widehat{Var}\left(\hat{B}'_{U1}\right)=(63)^2\left(1-\frac{u_1}{63}\right)\frac{\left(210.1\,\text{ha}^2\right)}{u_1}\quad+\quad(63)\left(58.5\,\text{ha}^2\right).$$

Equation 3-8

We now turn our attention to the second term in Equation 3-5, $\widehat{Var}\left(\hat{B}'_{M1}\right)$, and evaluate its dependence on sample size. This term is estimated by the expression (Skalski, 2003, Equation 27)

$$\widehat{Var}\left(\hat{B}'_{M1}\right)=A_2^2\left[\frac{s^2_{\hat{R}_1\hat{R}_2}}{m}+\frac{s^2_{\hat{R}_{M1}}-s^2_{\hat{R}_1\hat{R}_2}}{n}-\frac{s^2_{\hat{R}_{M1}}}{N}\right]$$

Equation 3-9

where

$$s^2_{\hat{R}_{M1}}=\frac{\sum_{j=1}^{m}\left(\hat{R}_{j1}-\hat{\bar{R}}_{M1}\right)^2}{m-1}$$

where

$$\hat{\bar{R}}_{M1}=\frac{\sum_{j=1}^{m}\hat{R}_{j1}}{m}$$

$$\hat{R}_{j1}=\frac{\hat{X}_{j1}}{a_{j1}}\ \text{for}\ j=1,\ldots,m$$

$s^2_{\hat{R}_1\hat{R}_2}$ = MSE from the ANOVA for the regression analysis of year-one ratios, $\hat{R}_{j1}$, as dependent variable and year-two ratios, $\hat{R}_{j2}$, as independent variable for the matching sites,

and where the sums are over matched sites only.

Under our simplifying assumption that sample variances from 2000-2003 are representative of all samples and sample sizes, the only dependencies on sample size are the terms *m* and *n* in Equation 3-9. Again using the constructed dataset shown in Table 3-1, the remaining terms in Equation 3-9 are evaluated as

$$s^2_{\hat{R}_{M1}}=0.00888$$

$$s^2_{\hat{R}_1\hat{R}_2}=0.000534.$$

Equation 3-9 then becomes

$$\widehat{Var}\left(\hat{B}'_{M1}\right)=\left(31{,}218\,\text{ha}\right)^2\left[\frac{5.34\times10^{-4}}{m}+\frac{8.35\times10^{-3}}{n}-\frac{8.88\times10^{-4}}{63}\right].$$

Equation 3-10

Now, the variance estimated from matched sites (Equation 3-10) and the variance estimate from unmatched sites (Equation 3-8) can be substituted into Equation 3-5 and the adjusted variance for the flats stratum, $\widehat{Var}\left(\tilde{B}_{SPR}\right)$ has been reduced to a form that is a function only of sample size – specifically the total number of flat sites sampled, *n*, and the breakdown into unmatched (u_1) and matched (*m*) sites.

In these calculations of the effects of sample size on adjusted variance estimates, we will maintain the 20% rate of site rotation currently used in the SVMP. We will vary sample size *n* in increments of five so that u_1 will vary in increments of one and *m* will vary in increments of four.

Variance Estimator for the Narrow and Wide Fringe Strata

The only remaining terms in Equation 3-4 that need to be evaluated are the adjusted variance estimates for the narrow fringe stratum, $\widehat{Var}\left(\tilde{B}_{fr}\right)$ and the wide fringe stratum, $\widehat{Var}\left(\tilde{B}_{fw}\right)$. We will now evaluate these terms using our simplifying assumption to make explicit the dependency on sample size.

The form of the estimators is the same for both narrow and wide fringe. Using the narrow fringe for this development, the variance of the adjusted estimate *Z. marina* area in the narrow fringe stratum is estimated by (Skalski 2003, Equation, 20)

$$Var\left(\tilde{B}_{fr}\right)=\left(\frac{L_T}{L_N}\right)^2 N^2\left[\frac{\widehat{Var}\left(\hat{\bar{X}}'_{U1}\right)\cdot\widehat{Var}\left(\hat{\bar{X}}'_{M1}\right)}{\widehat{Var}\left(\hat{\bar{X}}'_{U1}\right)+\widehat{Var}\left(\hat{\bar{X}}'_{M1}\right)}\right],$$

Equation 3-11

where

N = total number of sites in the narrow fringe sampling frame,
L_T = total linear length of the narrow fringe in the study area,
L_N = total linear length of the narrow fringe sampling frame (N x 1000m).

Note that the variance estimates in Equation 3-11 are all at the site level (X_i) while the analogous variance estimates for the flats stratum (Equation 3-5) were at the stratum level (B_i).

The estimator for the unmatched sites is expressed by (Skalski 2003, Equation 16)

$$\widehat{Var}\left(\hat{\bar{X}}'_{U1}\right)=\frac{s^2_{\hat{X}_{U1}}\left(1-\frac{u_1}{N}\right)}{u_1},$$

Equation 3-12

where

$$s^2_{\hat{X}_{U1}} = \frac{\sum_{j=1}^{u_1}\left(\hat{X}_{1j} - \hat{\bar{X}}_{U1}\right)^2}{(u_1 - 1)},$$

and where the sum is over unmatched sites only.

Again, we apply our simplifying assumption that sampling variances from the 2000-2003 SVMP data are representative of all samples of all sizes. In this case, the sampling variance for the unmatched sites, $s^2_{\hat{X}_{U1}}$, does not vary with sample size. We used 2002 SVMP data to calculate a nominal values for this term resulting in

$$s^2_{\hat{X}_{U1}} = 1.38\,\text{ha}^2.$$

Substituting this value and N=2019 for the narrow fringe stratum into Equation 3-12 gives

$$\widehat{Var}\left(\hat{\bar{X}}'_{U1}\right) = \frac{\left(1.38\,\text{ha}^2\right)\left(1 - \frac{u_1}{2019}\right)}{u_1}.$$

Equation 3-13

We now turn our attention to the remaining unknown term in Equation 3-11, $\widehat{Var}\left(\hat{\bar{X}}'_{M1}\right)$, and evaluate its dependence on sample size. This term is estimated by the expression (Skalski 2003, Equation 17)

$$\widehat{Var}\left(\hat{\bar{X}}'_{M1}\right) = \frac{s^2_{\hat{X}_{M1}\cdot\hat{X}_{M2}}}{m} + \frac{s^2_{\hat{X}_{M1}} - s^2_{\hat{X}_{M1}\cdot\hat{X}_{M2}}}{n} - \frac{s^2_{\hat{X}_{M1}}}{N}$$

Equation 3-14

where

$$s^2_{\hat{X}_{M1}} = \frac{\sum_{j=1}^{m}\left(\hat{X}_{j1} - \hat{\bar{X}}_{M1}\right)^2}{m - 1},$$

$$\hat{\bar{X}}_{M1} = \frac{\sum_{j=1}^{m}\hat{X}_{j1}}{m},$$

$s^2_{\hat{X}_{M1}\cdot\hat{X}_{M2}} = \frac{\text{SSE}}{m-2}$ = MSE from the ANOVA for the regression analysis of year-one *Z. marina* area estimates, $\hat{X}_{j1}$, as dependent variable and year-two estimates, $\hat{X}_{j2}$, as independent variable for the matching sites,

and where the sums are over matched sites only.

As before, we assume that sample variances are independent of sample size and here we will use the matching sites in the 2002 and 2003 SVMP data to calculate the nominal values

$$s^2_{\hat{X}_{M1}} = 6.82\,\mathrm{ha}^2$$

$$s^2_{\hat{X}_{M1}\cdot\hat{X}_{M2}} = 0.0884\,\mathrm{ha}^2.$$

Substituting these values into Equation 3-14 gives

$$\widehat{Var}\left(\hat{\bar{X}}'_{M1}\right) = \frac{0.0884\,\mathrm{ha}^2}{m} + \frac{6.73\,\mathrm{ha}^2}{n} - \frac{6.82\,\mathrm{ha}^2}{2019}.$$

Equation 3-15

Now, the variance estimated from unmatched sites (Equation 3-13) and the variance estimate from matched sites (Equation 3-15) can be substituted into Equation 3-11. Then the adjusted variance for the narrow fringe stratum, $\widehat{Var}\left(\tilde{B}_{fr}\right)$ has been reduced to a form that is a function only of sample size – specifically the total number of narrow sites sampled, *n*, and the breakdown into unmatched (u_1) and matched (*m*) sites.

The same approach applied to the wide fringe stratum reduces the last term in Equation 3-4 to a form where dependency on sample size is explicit. The expression for this term is analogous to Equation 3-11 and has unmatched and matched site components which can be evaluated using 2001 and 2002 SVMP wide fringe data (2002-03 data cannot be used to calculate adjusted estimates since there is only one unmatched site). The resulting expressions are

$$\widehat{Var}\left(\hat{\bar{X}}'_{U1}\right) = \frac{\left(78.0\,\mathrm{ha}^2\right)\left(1-\dfrac{u_1}{351}\right)}{u_1},$$

Equation 3-16

$$\widehat{Var}\left(\hat{B}'_{M1}\right) = \frac{8.24\,\mathrm{ha}^2}{m} + \frac{92.3\,\mathrm{ha}^2}{n} - \frac{100.5\,\mathrm{ha}^2}{351}$$

Equation 3-17

using the value of *N*=351 for the total number of wide fringe sites.

Sample Size Scenarios

The variance estimators for the adjusted estimates have now been reduced to forms that depend only on sample size for the three strata subject to random sampling: flats (Equation 3-5, Equation 3-8, Equation 3-10), narrow fringe (Equation 3-11, Equation 3-12, Equation 3-15) and wide fringe (Equation 3-11, Equation 3-16, Equation 3-17). The data sources used are summarized in Table 3-2.

In order to assess the effects of sample size, the power of the 5-year and 10-year tests was evaluated as sample size was increased in increments of five from the base values shown in Table 3-2. For each increment of five, the number of matched sites was incremented by four and the number of unmatched sites incremented by one in order to maintain a 20% rate of site rotation.

stratum	data source for sample variances	*N*	*n*	*m*	*u*
core	2002	6	6	-	-
persistent flats	2002 (flats11)	3	3	-	-
rotational flats	constructed dataset drawn from 2000-03 SVMP data	63	10	8	2
narrow fringe	2002-03 SVMP data	2019	44	36	8
wide fringe	2001-02 SVMP data	351	12	10	2

Table 3-2. Data sources and sample sizes used to characterize sample variances that were in turn used to evaluate sensitivity to sample size. These data were used to calculate fixed terms in variance estimators.

The effects of sample size were evaluated independently for flats, narrow fringe and wide fringe. For each scenario, the total variance (Equation 3-4), CV (Equation 3-3) and noncentrality parameter (Equation 1-4, Equation 1-5) were calculated. Power estimates were interpolated from the noncentral F-tables of Skalski and Robson (1992).

3.4.2 Results and Discussion

The power results are shown in Figure 3-8 and Figure 3-9 for sample size increments from +0 to +40 sites added to the base levels in Table 3-2.

Several key points emerge from these results:

1. All tests based on adjusted estimates are reasonably powerful even without additional sites.
   $1\text{-}\beta = 0.870$ (5-year test, no additional sites)
   $1\text{-}\beta = 0.986$ (10-year test, no additional sites)
   This is generally consistent with the results presented by Skalski (2003, sections 6.2 and 6.3) of $1\text{-}\beta = 0.867$ and $1\text{-}\beta = 0.946$ for the 5- and 10-year tests, respectively, with no additional sites. Some discrepancy with Skalski (2003) is expected since he investigated a 25% decline, rather than 20%.
2. For both 5-year and 10-year tests, the addition of sites to the wide fringe stratum leads to the greatest increase in statistical power.
3. For both 5-year and 10-year tests, the addition of sites to the narrow fringe stratum has the least effect on statistical power.

This analysis does not consider the discrepancy in incremental field effort across the different strata. Flat sites are highly variable, but it is reasonable for this purpose to assume that on average a flat site requires twice the sampling effort as a narrow fringe site. In this case, the addition of five sites to the flats stratum would represent the same increment in effort as the addition of ten sites to the narrow fringe stratum. Even if this factor were considered explicitly in the analysis, the results would be qualitatively the same and the key results listed above would still hold.

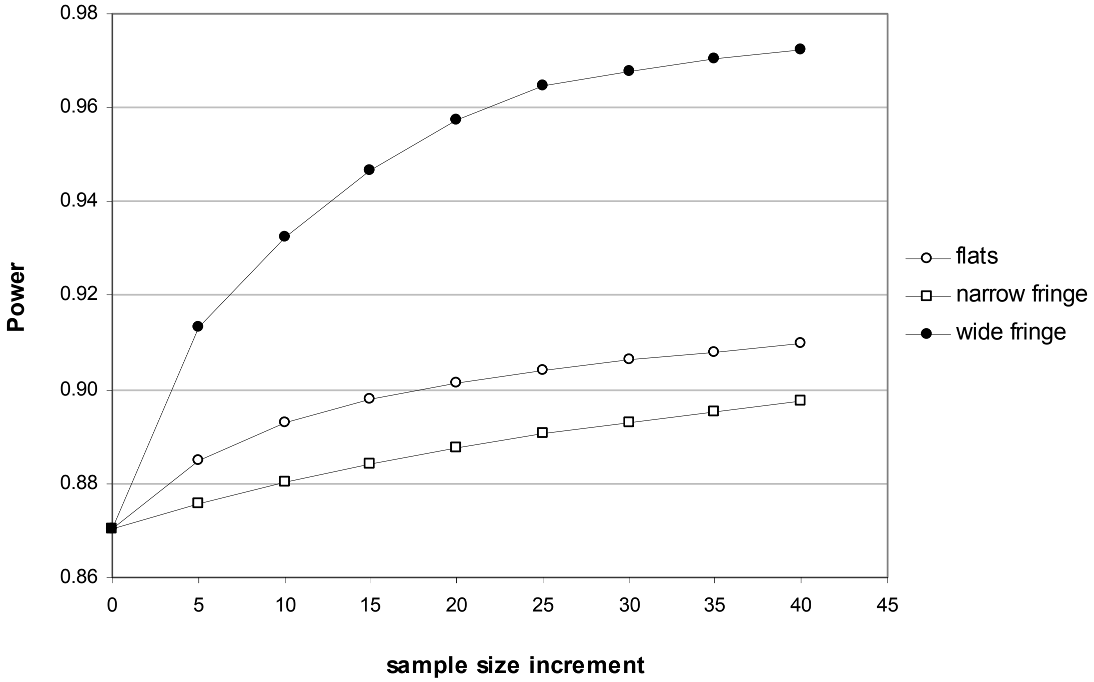


Figure 3-8. Effects of sample size on the power of the 5-year test using adjusted estimates to detect a steady decline in *Z. marina* area totaling 20% over five years. Sample size increment is the number of additional sites above the base levels shown in Table 3-2.

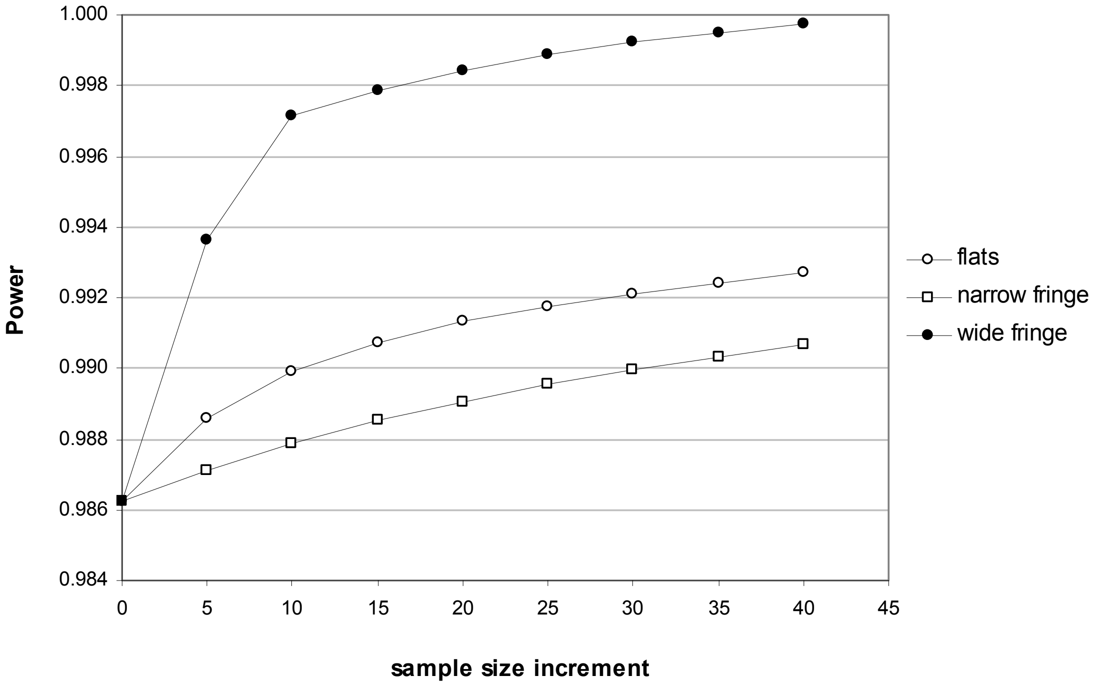


Figure 3-9. Effects of sample size on the power of the 10-year test using adjusted estimates to detect a steady decline in *Z. marina* area totaling 20% over ten years. Sample size increment is the number of additional sites above the base levels shown in Table 3-2.

3.4.3 Approach for Testing Effects on Initial Estimates

The power analysis presented in the previous section was based on the premise that all estimators used are unbiased. Section 2.4 presents an argument that the adjusted variance estimator for the flats stratum may be subject to bias because of the low sample size (*u*=2) and skewed distribution of *Z. marina* area ratio. A similar argument could be made for the wide fringe stratum, which also has a sample size of two for unmatched sites.

Section 2.4 only presents a plausibility argument and does not demonstrate the existence of bias in the adjusted estimators. Nevertheless, the power analysis presented in the previous section is applied here to the initial estimators in order to eliminate or at least minimize any potential bias as a factor in the analysis.

The variance of the initial estimate of total *Z. marina* area in Puget Sound will follow Equation 3-4 except initial estimates will replace adjusted estimates for the flats, narrow fringe and wide strata. Estimates of variance in the core and ζ-flats strata will remain the same.

The initial variance for the flats stratum has the same form as Equation 3-7 except u_1 is replaced by the total number of sites in the flats stratum and the sums are now over all flats sites (see Skalski 2003, Equation 11). Using the constructed Year 1 dataset from Table 3-1, this equation reduces to

$$\widehat{Var}\left(\hat{B}_{SPR}\right)=\frac{(63)^2}{n}\left(1-\frac{n}{63}\right)\left(975.9\,\mathrm{ha}^2\right)+(63)\left(18.9\,\mathrm{ha}^2\right).$$

Equation 3-18

The estimators for the initial variance of the narrow and wide fringe strata are given in Skalski (2003, Equation 8). Again using 2002 SVMP data for the narrow fringe and 2001 data for the wide fringe, these reduce to

$$\widehat{Var}\left(\hat{B}_{fr}\right)=\left(\frac{2066.8\,\mathrm{km}}{2019.0\,\mathrm{km}}\right)^2\left[\frac{2019^2\left(1-\frac{n}{2019}\right)\left(5.92\,\mathrm{ha}^2\right)}{n}+(2019)\left(0.0693\,\mathrm{ha}^2\right)\right],$$

Equation 3-19

$$\widehat{Var}\left(\hat{B}_{fw}\right)=\left(\frac{356.9\,\mathrm{km}}{351.0\,\mathrm{km}}\right)^2\left[\frac{351^2\left(1-\frac{n}{351}\right)\left(100.5\,\mathrm{ha}^2\right)}{n}+(351)\left(2.75\,\mathrm{ha}^2\right)\right].$$

Equation 3-20

As before, the effects of sample size are explored in increments of five sites and the flats, narrow fringe and wide fringe strata are investigated separately. For each scenario, the total variance, CV (Equation 3-3) and noncentrality parameter (Equation 3-1, Equation 3-2) were calculated. Power estimates were interpolated from the noncentral F-tables of Skalski and Robson (1992).

3.4.4 Results and Discussion

The power results for sample size increments from +0 to +40 sites are shown in Figure 3-10 and Figure 3-11 for the 5-year and 10-year tests, respectively.

Three key points emerge from these results using initial estimates:

1. For both 5- and 10-year tests the power under each sample size scenario is only slightly less than the power when using the adjusted estimates.
2. These results are not consistent with the single 5-year result presented in Skalski (2003, section 6.2.1) of $1-\beta \approx 0.30$ when using initial estimates.
2. Power is most sensitive to the addition of sites to the wide fringe stratum and least sensitive to the addition of sites to the narrow fringe stratum. This is the same pattern seen when using adjusted estimates.

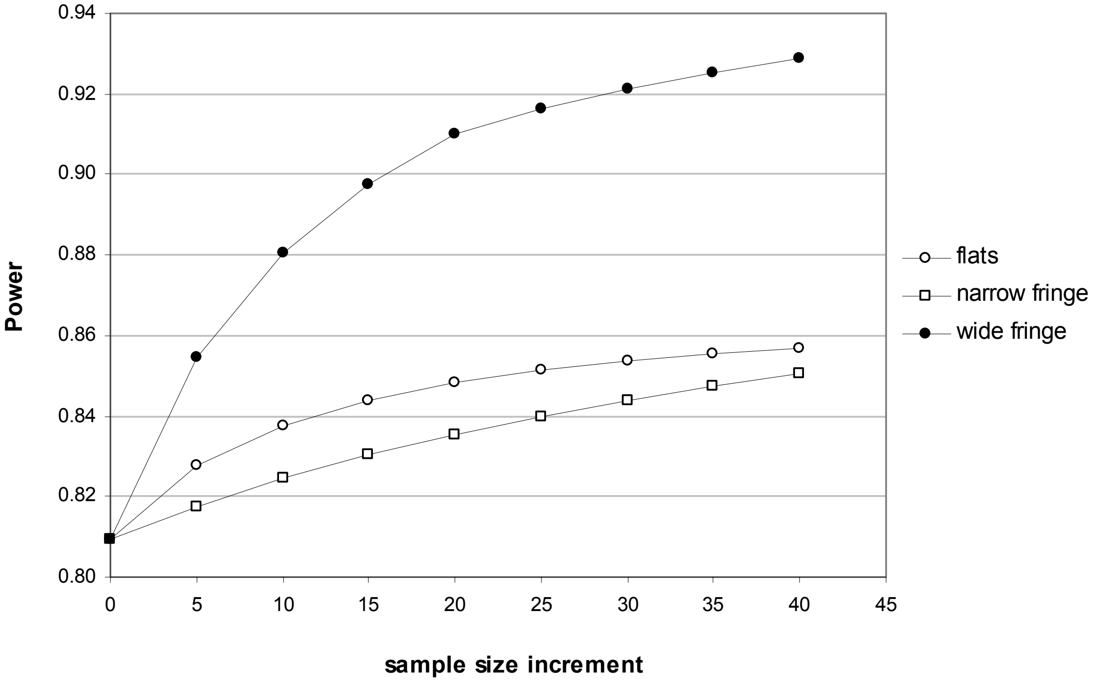


Figure 3-10. Effects of sample size on the power of the 5-year test using initial estimates to detect a steady decline in *Z. marina* area totaling 20% over five years. Sample size increment is the number of additional sites above the base levels shown in Table 3-2.

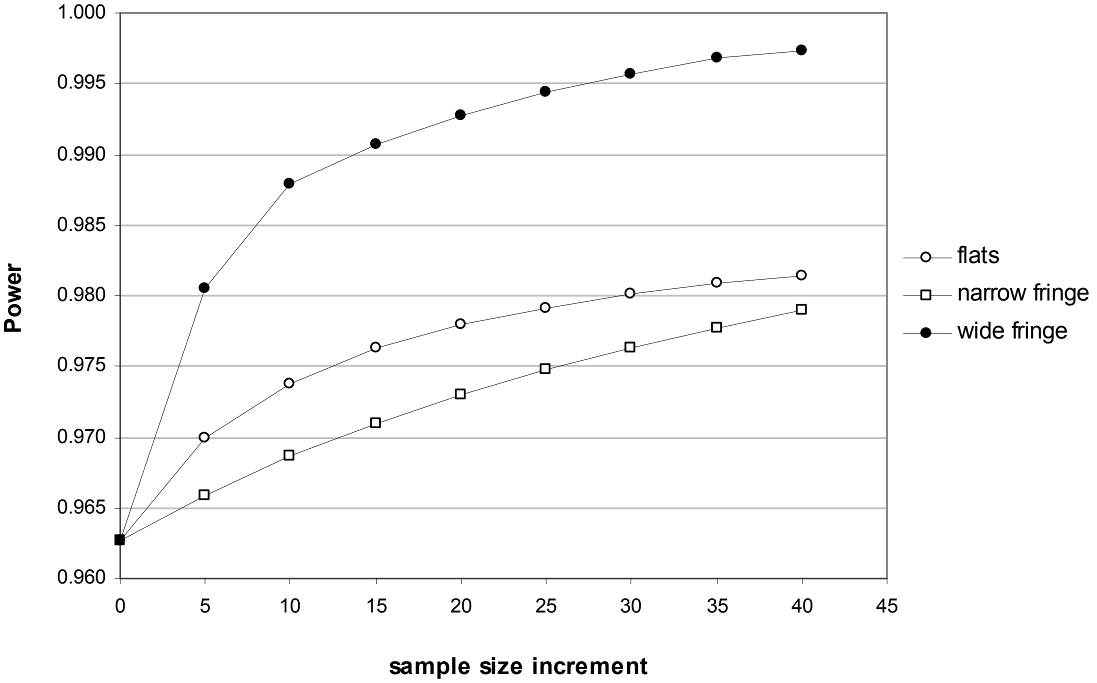


Figure 3-11. Effects of sample size on the power of the 10-year test using initial estimates to detect a steady decline in *Z. marina* area totaling 20% over five years. Sample size increment is the number of additional sites above the base levels shown in Table 3-2.

The similarity in power between tests with initial estimates (Figure 3-10, Figure 3-11) and tests with adjusted estimates (Figure 3-8, Figure 3-9) was unexpected. The previous power estimates by Skalski (2003, pp.27-29), based on 2001 results, showed a strong improvement in power associated with adjusted estimates ($1-\beta \approx 0.866$) relative to initial estimates ($1-\beta \approx 0.30$).

There are several possible explanations for the discrepancy between results presented here and the results of Skalski (2003). Here we tested for 20% decline in *Z. marina* area, while Skalski tested for 25% decline. Also, the estimates derived here were based on data drawn from the entire 2000-2003 SVMP dataset (Table 3-2), while the results of Skalski (2003) relied solely on 2001 results. There were also small adjustments to the 2001 results since they were originally published (Berry et al. 2003). It is far more likely, however, that the discrepancy originates from the introduction of the persistent flats stratum in the estimates presented here.

In order to test the plausibility of this explanation, the power of 5- and 10-year tests was calculated under two scenarios that differed only in the stratification of the flat sites. The 2001 SVMP data was used as originally collected to represent the 2000-03 stratification.

The 2001 SVMP data was also used to represent the 2004 stratification but two manipulations were necessary to make the data conform to the 2004 strata. First, flats11 and flats20 were removed from the rotational flats stratum and put in the new persistent flats stratum. This left the rotational flats stratum with a sample size of only n=8. Second, since there was no data to represent flats12 in the persistent flats stratum, it was assumed to be identical to flats11.

The results from these two scenarios, including stratum variances, CV and power, are shown in Table 3-3. These results clearly show that only by changing the stratification of the flats, the overall CV is dramatically reduced from 0.27 under the 2000-03 stratification to 0.08 under the 2003 stratificaiton. This decline in CV leads to dramatic increases in the power of the tests for *Z. marina* decline.

The power of the 5-year test with the 2000-03 stratification, 0.26, is also in closer agreement with the result of Skalski (2003), 0.30. This suggests that differences between the 2000-03 and 2004 stratification explain the discrepancy with Skalski (2003) discussed above (p.75).

	variance (ha^2)		CV	1-β 5-year test	1-β 10-year test
2000-03 Stratification (2001 data)	core	17,631	0.2725	0.2648	0.3262
	flats	31,034,593			
	narrow fringe	692,417			
	wide fringe	971,929			
	overall	32,716,570			
2004 Stratification (2001 data modified)	core	17,631	0.0785	0.7184	0.9188
	main flats	663,912			
	persistent flats	6,890			
	narrow fringe	692,417			
	wide fringe	971,929			
	overall	2,352,779			

Table 3-3. Effect of flats stratification on variance of *Z. marina* area estimates and power to detect *Z. marina* decline. Power results are for tests of steady 20% decline over either a 5-year or 10-year period. Results are based on 2001 data. These data were modified for the 2004 strata by moving flats11 and flats20 to the persistent stratum, leaving the rotational flats stratum with *n*=8. The third site in the persistent flats stratum, flats12, was assumed to be identical to flats11. All values for variance and CV are initial estimates.

The power of tests for decline is low (~ 0.30) when using initial estimates based on the 2000-03 strata. The improvement in power associated with changing to the new flats strata is roughly equivalent to the improvement associated with the use of retrospectively adjusted estimates. This is illustrated diagrammatically in Figure 3-12. The initial estimates based on the 2004 strata can also be retrospectively adjusted, but the improvement in power is modest.

This result is important because it suggests that even if future work reveals that adjusted estimates are biased due to small sample sizes and skewed distributions, the transition to the 2004 strata alone provides substantial improvement in the power of tests for *Z. marina* decline.

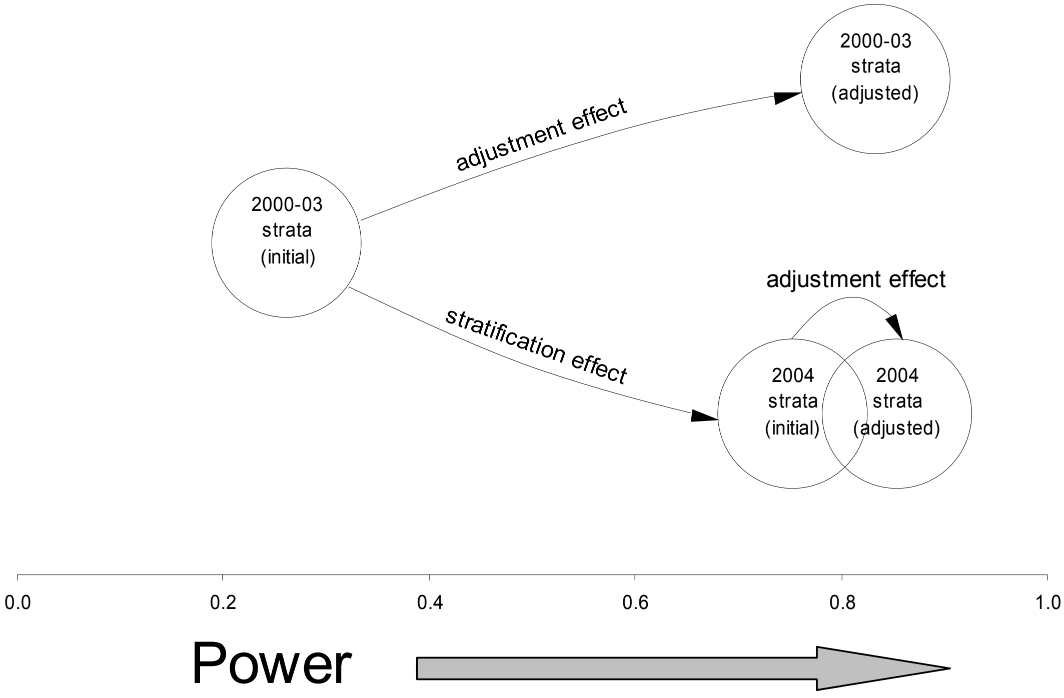


Figure 3-12. Conceptual diagram comparing the improvements in power achieved by retrospective adjustment and by changing the flats stratification. Testing for *Z. marina* decline has relatively low power when using initial estimates based on the 2000-03 strata. Similar large improvements in power are achieved by testing with either retrospectively adjusted estimates or initial estimates based on the new 2004 strata that include the persistent flats stratum. A modest additional improvement in power is achieved by using adjusted estimates based on the 2004 strata. Values of power are loosely based on results with the 5-year test for decline.

3.5 Summary and Recommendations

The SVMP was designed to identify decline in *Z. marina* in Puget Sound. It is based on the premise that *Z. marina* represents an important ecosystem resource and a detected decline of sufficient magnitude would warrant a coordinated management response. For SVMP results to be useful in a management context, they must be interpreted in light of statistical confidence and power associated with tests for decline.

This chapter emphasizes that statistical power of tests for *Z. marina* decline is a complex parameter that responds to level and pattern of decline as well as the number of sites and their allocation among strata.

The scope of this chapter was restricted to the 5- and 10-year tests for decline as described in Skalski (2003). The power of change detection using the paired site analysis was not addressed.

The following key findings emerged from the work described in this chapter:

1. The 5-year test has several disadvantages relative to the 10-year test when testing for linear decline in *Z. marina*:
   a. The 5-year test has substantially less power than the 10-year test to detect a given level of total decline over the test period.
   b. The 5-year test has even less power relative to the 10-year test to detect a given rate of annual decline.
   c. The 5-year test is more sensitive to CV of the estimates of *Z. marina* area so if precision decreases in future estimates, the power of the 5-year test will drop more than that of the 10-year test.

2. In contrast, the 5-year test clearly has the important advantage of lower data requirements and earlier delivery of results.

3. Given the current estimate of $\mathrm{CV} \approx 0.07$, the primary SVMP performance measure meets the target of achieving the ability to detect a 20% decline in Puget Sound *Z. marina* over 10 years under a scenario of linear decline. The 10-year test has a power of $1-\beta = 0.96$ when testing at $\alpha = 0.10$ and $1-\beta = 0.89$ when testing at $\alpha = 0.05$.

4. The 5-year test has unacceptably low levels of power when testing at $\alpha = 0.05$, e.g. $1-\beta = 0.59$ for 20% decline.

5. The current estimate of $\mathrm{CV} \approx 0.07$ is located on the steepest part of the power-CV curves so that small declines in precision in future estimates of *Z. marina* area will lead to strong declines in power for both 5- and 10-year tests.

6. The power of the tests can be very sensitive to the temporal pattern of decline. The power of the 10-year test to a sharp 3-year decline of 20% was shown to range from $1\text{-}\beta = 0.91$ to 0.42 depending on the timing of the decline within the 10-year data record.

7. The transition to the 2004 sampling strata (introduces the persistent flats stratum) leads to a strong improvement in power when using initial estimates of *Z. marina* area. The magnitude of this improvement is roughly equivalent to the dramatic improvement associated with the retrospective adjustment previously reported (Skalski 2003).

8. The addition of sample sites leads to the greatest improvement in power when added to the wide fringe stratum and the least improvement when added to the narrow fringe stratum.

The following specific recommendations are based on these findings:

- The 5-year test for *Z. marina* decline should only be applied with $\alpha \geq 0.10$ in order to achieve acceptable power.

- Additional power analyses should be completed if precision of *Z. marina* area estimates changes in future results or as other temporal patterns of decline become of particular interest.
- Future research should assess the benefit of shifting resources from the narrow fringe stratum to increase sample size in the wide fringe stratum in order to increase power.
- Future research should address the power of change detection using the paired site analyses.

4 Optimal Rate of Rotation

This chapter develops a framework for setting the rate of site rotation between sampling seasons. This effort was motivated by the observation that there are a range of guidelines given by Cochran (1977), some that are clearly violated by the SVMP rate of rotation over 2000-2003.

The SVMP statistical framework presents an equation to calculate an optimal rate (Skalski 2003, p.13) but it does not reflect the optimization criteria relevant to the SVMP. It produces optimal rates of rotation that are very different from what has actually been implemented during 2001-2003. It specifies that the fraction of total sites, n, that are matched, or retained, between years is given by

$$\frac{m}{n} = \frac{\sqrt{1-\rho^2}}{1+\sqrt{1-\rho^2}}$$

Equation 4-1

where m is the number of matched sites and ρ is the correlation coefficient for the matched sites between the two years[5].

The central tradeoff in determining the rate of site rotation is the optimization of status estimates versus the optimization of change estimates. Cochran (1977, pg. 345) sums this up in the following statements about sample replacement policy:

1. For estimating *change*, it is best to retain the same sample throughout all occasions.
2. For estimating the average *over all occasions*, it is best to draw a new sample on each occasion.
3. For *current* estimates, equal precision is obtained either by keeping the same sample or by changing it on every occasion. Replacement of part of the sample on each occasion may be better than these alternatives.

The SVMP aims to make estimates in all of these categories, particularly 1 and 3, but the relative importance of status versus change estimates has not been explicitly articulated.

Skalski (1990) echoes the statements of Cochran but also discusses a potential weakness of using the same sites (no rotation) even when concerned only in change. "Should a new and significant pollution source arise, the success of the monitoring program will depend more upon the chance the point source occurs near one of the earlier established [sites]

---

[5] In general, this chapter follows the notation of Cochran (1977).

than upon sampling efforts." He presents this as an argument for the partial replacement sampling design that is central in the SVMP statistical framework (Skalski, 2003).

This chapter first discusses the optimization of status estimates alone as presented by Skalski (2003) and Cochran (1977). Then details of rotation as implemented by SVMP are presented. The optimization of both status and change estimates is then discussed following the presentation of Cochran (1977). Also, a new framework is developed that is more applicable to the SVMP that is used for calculating hypothetical variance improvements. Finally, a recommendation for a change to SVMP analysis is made.

4.1 Optimal Rotation for Status Estimates

Optimization of site rotation with respect to the retrospectively adjusted status estimates corresponds to statement 3 above. This is the formulation presented in the statistical framework (Skalski 2003, p.13). It follows directly from Cochran (1977, Equation 12.75). In Cochran, this optimization is constrained to produce partial replacement – i.e. no replacement and total replacement are not allowed as solutions.

In this case, the goal is to optimize the retrospectively adjusted status estimate for the case of repeated sampling on two occasions. The optimization is based on minimizing the variance in the estimate. Note that this does not consider the performance of the adjustment over more than two occasions nor how rotation affects the change estimate (discussed in section 4.3).

Cochran (1977, p.347) points out that the optimal rate of rotation for status estimates using this method is never less than 50%.

Figure 4-1 shows the relationship between optimal rate of rotation and the correlation in matched sites (from Equation 4-1). As the correlation approaches one, the optimal rate of matching rapidly approaches zero, i.e. total rotation. Figure 4-1 also shows the gain in precision of the status estimate when sampling at the optimal rate of matching. As correlation approaches one, the gain in precision rapidly approaches 100%. At perfect correlation, $\rho = 1$, there is no reason to resample a site and in this case total rotation has the benefit of increasing the sample size by 100%.

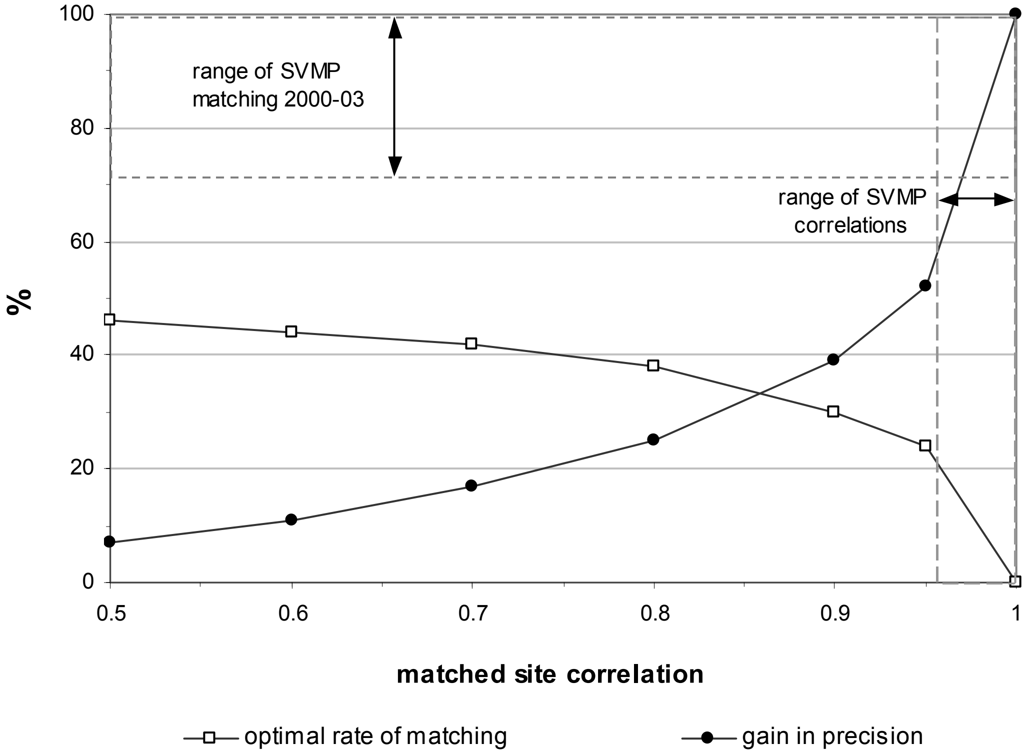


Figure 4-1. Results of optimizing status estimate only. The optimal rate of rotation varies with correlation between matched sites (Skalski, 2003, p.13; Cochran 1977, Equation 12.75, p.347). For the optimal rate of matching, the gain in precision of the status estimate is also given. The shaded regions indicate specific ranges for matching and correlations in the SVMP for 2000-2003. The values plotted are from Cochran (1977, Table 12.2, p.347).

The range of matched site correlations observed in the SVMP is also shown in Figure 4-1. This range is very high (>95%), leading to a very low optimal rate of matching (high optimal rate of rotation). This is not consistent with the high rate of matching that has been implemented in the SVMP (also shown in Figure 4-1). Clearly, the SVMP implementation does not optimize status estimates alone as suggested in the statistical framework (Skalski 2003, p.13).

Cochran (1977) makes a recommendation in his discussion (p.347) that is valuable in cases of low numbers of matched sites. In cases where the optimization points to total rotation or the absence of rotation, then the result is overruled since these solutions are not allowed in this particular optimization approach. Cochran's direction here is important:

> "When ρ=1, the formula suggests m=0, which lies outside the range of our assumptions, since m has been assumed reasonably large. The correct procedure in this case is to take m=2. The two matched units are sufficient to determine the regression line exactly."

This is notable because it points out that a minimum of two sites is required in the matched category and in this case Cochran deems a sample size of two to be sufficient.

In his discussion of bias in regression estimators, Cochran appears to contradict this advice by concluding sample size must be at least 12 (see discussion on p.45).

4.2 SVMP Rotation 2000-2003

Table 4-1 shows the actual rates of SVMP site rotation by stratum for 2000-03. Sites were not rotated between 2000 and 2001. The correlations and optimal rates resulting from the method in the Skalski (2003) are also shown. The following variables are used:

n = total number of samples (assumed the same in year i and year i+1)
m = number of matching sites in the two years
u = number of unmatched sites in year i (note that n = m + u)
ρ = correlation coefficient for the matched sites between the two consecutive years

u/n = rate of rotation
m/n = rate of matching [note that $m/n = (1 - u/n)$]

	2000-01						**2001-02**						**2002-03**					
stratum	n	u	m	u/n actual	ρ	u/n opt.	n	u	m	u/n actual	ρ	u/n opt.	n	u	m	u/n opt.	ρ	u/n opt.
core	6	0	6	0.00	0.9996	0.97	6	0	6	0.00	0.9999	0.99	6	0	6	0.00	0.9994	0.97
flat	9	0	9	0.00	0.9688	0.80	10	2	8	0.20	0.9978	0.94	10	2	8	0.20	0.9998	0.98
narrow fringe	42	0	42	0.00	0.9930	0.89	44	9	35	0.20	0.9916	0.89	44	8	36	0.18	0.9937	0.90
wide fringe	4	0	4	0.00	0.9658	0.79	13	3	10	0.23	0.9629	0.79	13	1	12	0.08	0.9920	0.89

Table 4-1. Actual rates of SVMP site rotation for 2000-03 and the optimal rates as calculated by Equation 4-1. The optimization is based on the method in the statistical framework which optimizes the retrospectively adjusted status based on year-to-year correlations, ρ, in matched sites. It does not consider optimization trade-offs between the status and change estimates.

Table 4-1 highlights three key points:

- The year-to-year correlations in matched site *Z. marina* area are very high (0.96 – 0.99).
- The optimal rates of site rotation for status estimates are high (a minimum of 79% for wide fringe in 2000-01).
- The actual rates of rotation implemented in the SVMP do not conform to the optimal rates calculated using this method. The actual rates are far less than the 50% limit noted by Cochran (1977, p.347).

The discrepancy between actual and optimal rotation indicates that this method of optimization was not used in the development of the sampling design. The method used to develop the 20% rotation guideline used in 2002 and 2003 is not reported but two refinements were possibly involved. The first and most important is the simultaneous optimization of the precision of the change estimate. As Cochran points out (1977, p.351), when estimates of change are also of interest, "this factor also points toward matching more than half the units from one occasion to the next" – i.e. rate of rotation less than 50%.

The second possible refinement involves the relative cost of sampling repeat sites (matched) versus new sites (unmatched). In the SVMP, new sites have higher reconnaissance requirements and perhaps a 3:2 or even 2:1 ratio of total effort requirement relative to repeat sites. If this is explicitly incorporated into the optimization process, the optimal solution shifts toward greater matching. Cochran presents an approach to including this relative cost in the optimization (Cochran 1977, Equation 12.78, p.348) but does not present a solution.

4.3 Optimization for Both Status and Change

4.3.1 Example from Cochran (1977)

Cochran (1977, section 12.13) develops an example of optimization for both status and change for sampling over more than two occasions. This example does not consider differential cost of new and repeat sites. It differs from the SVMP design in that it fixes the weights used to combine unmatched and matched estimates into an overall adjusted estimate. In the SVMP analysis, the weights are not fixed but rather, a function of the variances under consideration. Hence, the analysis here is only an approximation of the actual precision and optimality. Nevertheless it illustrates quallitative points relevant to the SVMP.

Cochran presents this as a practical example and assumes that in practice it is more convenient to have a constant rate of rotation rather than adjusting the rotation every year. He also assumes it is more convenient to keep the weights constant. These weights were previously (section 4.1) based on variance estimates from these pools and in practice would change from year to year.

A general feature of Cochran's approach to practical situations emerges in this example. In response to the goal of fixing weights and rate of rotation, Cochran sets these key parameters in a loose, imprecise manner rather than closely adhering to the exact calculated values specific to each case.

For example, Cochran derives an explicit formula (follows Equation. 12.86, p.352) for optimal weighting, ϕ_{opt}, for this multi-year retrospective adjustment method. For this method when sampling with 25% rotation, ϕ_{opt} has the following values depending on the correlation in the matched sites:

ρ	ϕ_{opt}
0.7	0.216
0.8	0.198
0.9	0.164

Table 4-2. Optimal weights for combining matched and unmatched estimates for different levels of correlation in matched sites (Cochran 1977, p. 352). This is not directly applicable to the SVMP because it is based on fixed weights over multiple occasions.

If data exhibits this range of correlation, Cochran suggests it would be adequate to fix ϕ_{opt} at the value of 0.2. The point here is not the specific value of the weighting, but the general manner in which it is set.

Similarly, the results from this practical example show that increases in the rate of matching, especially if ρ is high, produce substantial improvements in the efficiency of the change estimate at the expense of much smaller degradation in the current status estimate (Figure 4-2; note that gain in *efficiency* is a different measure than simple gain in precision, see section 2.1 on p.41). Cochran (p.354) concludes that "the results suggest that retention of 2/3, 3/4 or 4/5 [of samples] from one occasion to the next may be a good practical policy if current estimates and estimates of change are both important." It is reasonable that the SVMP rate of matching is at the upper end of this range (4/5) since correlations of matched sites are very high and because the higher cost of new sites compared to matched sites favors higher matching.

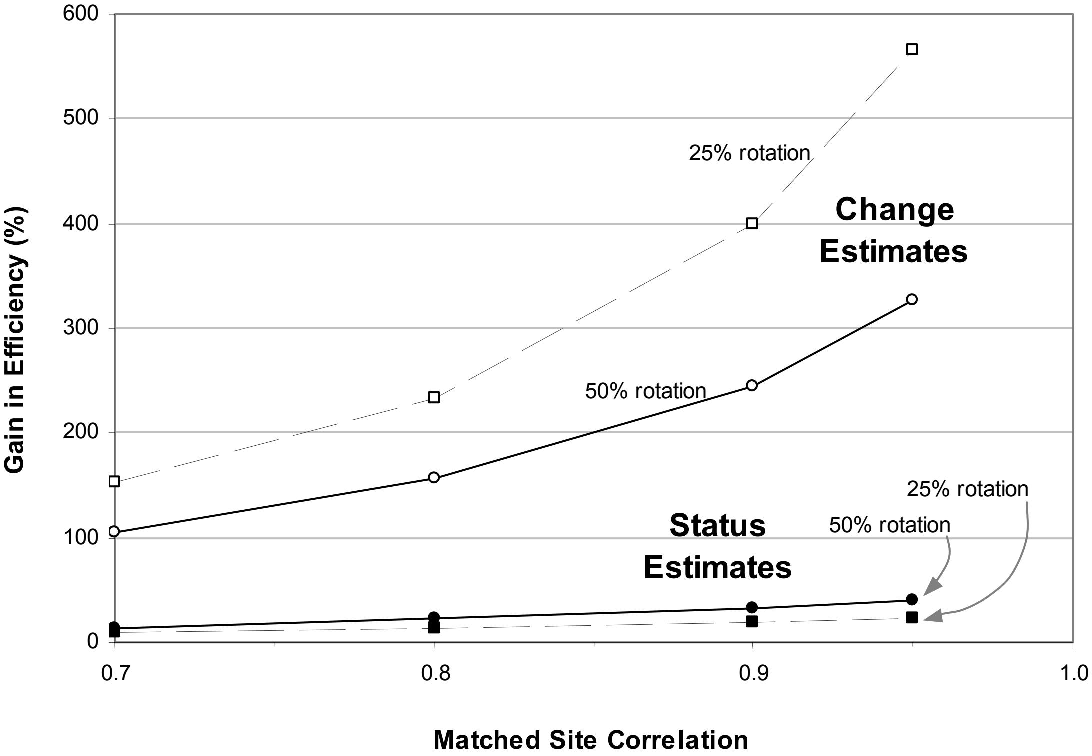


Figure 4-2. Gain in efficiency of both status and change estimates for a range of matched site correlations and two rotation scenarios. Changing from a 50% to 25% rotation (more matching sites) would strongly improve the change estimate, especially at high correlation, but only mildly degrade the status estimate. Values shown are not directly applicable to the SVMP because they are based on the use of fixed weights. Values are from Cochran (1977, Table 12.5, p.353) for the second sampling occasion. Note that the change estimates referred to here are different from the paired-site change estimates calculated for the SVMP.

Note that the change estimate referred to here is different from the paired site change estimate of year-to-year change used in the SVMP. The change estimate discussed here (follows Cochran 1977) is equivalent to calculating year-to-year change using consecutive sound-wide *Z. marina* status estimates.

4.3.2 Theoretical Gain in Precision for SVMP Design

This section develops calculations of gain in precision (reduction in variance) rather than efficiency as depicted in Figure 4-2. These calculations are developed for the situation where the rate of matching is fixed but the weights are not fixed but calculated from the data on each occasion. This more closely resembles the specifics of the SVMP design than the results in the previous section (4.3.1) where fixed weights were used. The calculations presented here are relevant to a single SVMP stratum subject to sampling with partial replacement.

It is important to note that these calculations still diverge from SVMP design in that they are based on the use of simple means rather than the more complex estimators used in the actual SVMP sound-wide extrapolation. Also the change estimate referred to here is not the SVMP paired site change estimate. The change estimate referred to here is analogous to calculating change using annual *Z. marina* status estimates.

Furthermore, the calculations presented here diverge from the SVMP design in that measurement error is not considered. As shown in section 2.5 on p.47, measurement error is a minor component of overall variance. Hence, the absence of this factor in the underlying model, is probably not a major limitation.

The derivations for the main equations used here are shown in Appendix D. Equations Eqn D.1 (p.129) and Eqn D.10 (p.131) can be used to investigate the effects of matching, correlation and number of occasions sampled on variance estimates. Figure 4-3 shows results for the second sampling occasion, i.e. the first opportunity to apply the adjustment calculation.

Higher rates of matching benefit the adjusted change estimate (by reducing variance relative to the initial estimate) at the expense of the adjusted status estimate. This is the same pattern seen in Figure 4-2 although the effects here on status and change estimates are roughly of the same absolute magnitude. The main reason for this difference is that Figure 4-3 simply depicts the reduction in variance relative to the initial estimates, while Figure 4-2 depicts efficiency, a quantity that is computed relative to the adjusted value.

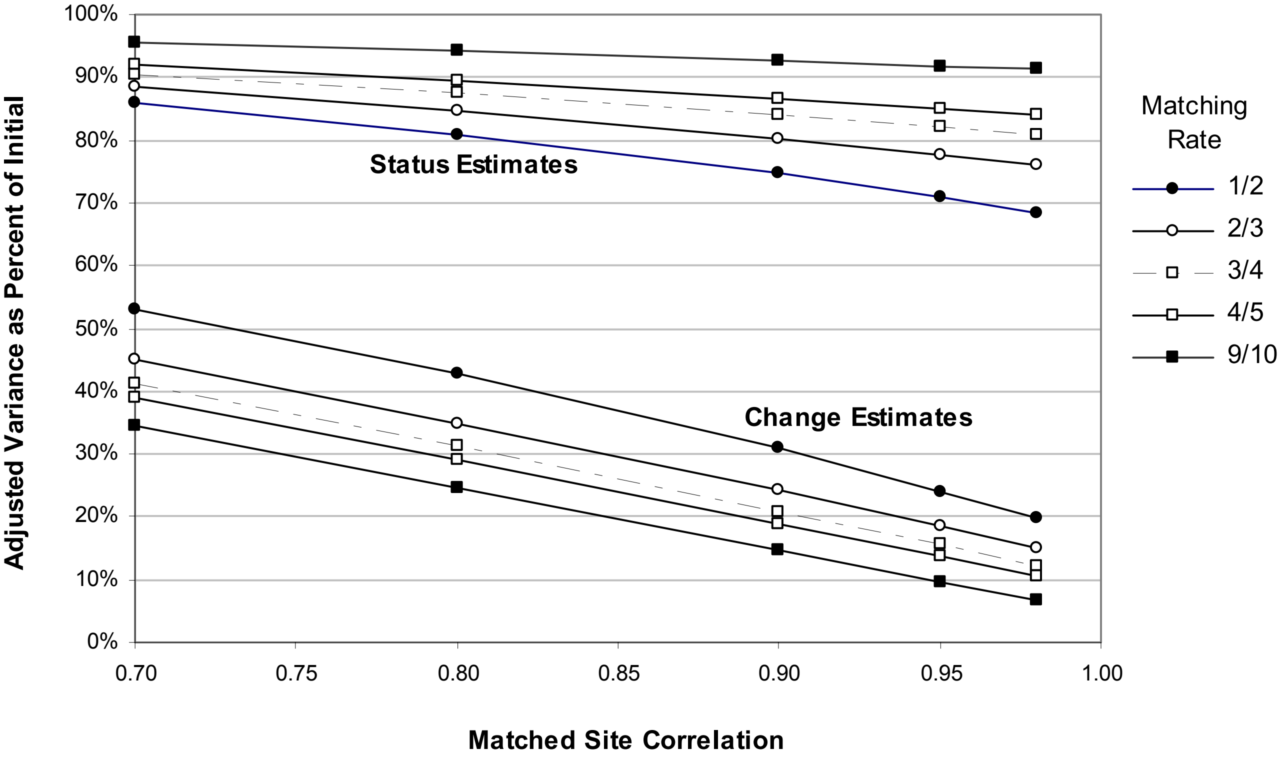


Figure 4-3. The effect of rate of matching and correlation of matched sites on the adjusted variance of both status and change estimates. Note that the change estimates referred to here are different from the paired-site change estimates calculated for the SVMP and are analogous to calculating change from annual *Z. marina* status estimates.

At the rates of matching shown in Figure 4-3, the improvement to the change estimates is much greater than the improvement to the status estimates. As noted earlier, the calculations of improvement in status estimates are more relevant to the SVMP analysis than either the results presented in Cochran (1977) or the change estimates reported here. These status results are useful as a rough guideline for evaluating observed improvements in SVMP status estimates at the stratum level for 2001-2003. As seen in Figure 4-3, for the SVMP rate of matching (4/5) and correlation in matched sites (~98%), these results indicate that the variance of the adjusted status estimate should be roughly 84% of the initial variance. This result is based on the assumption of constant sample variance and neglects covariance in matched site observations in consecutive years (see Eqn D.9 on p.131). Neglecting this covariance results in an underestimate of the improvement in precision of the change estimates.

4.3.3 Recommended Change to SVMP Analysis

In the presentation of Cochran (1977), the estimates from the most recent sampling occasion are adjusted using previously collected data. In this way, the adjusted estimate for the latest occasion is always the current estimate with highest precision. The disadvantage to this approach is that the estimates from the first sampling occasion are never adjusted (Skalski, pers. comm.). To avoid this situation, the SVMP statistical framework uses the most recent initial estimates to adjust the estimates of the previous occasion (Skalski 2003). In this approach the adjusted estimates for the previous occasion are always more precise than those for the latest occasion. At an SVMP team meeting convened in February 2004 to review initial results from 2003, this issue was discussed. The group voiced the need to provide final results for the most recent sampling occasion. This was thought to be important in the effort to build a consistuency that relies on SVMP results. As a compromise, the recommendation here is to follow the design of Skalski (2003) for the maintenance of the long-term SVMP data record and conducting tests for trend. However, the most recent status can also be adjusted using the result from the previous occasion in order to produce new results in a timely fashion. These two approaches must be kept separate, i.e. the adjusted estimate from the most recent occasion would not be used for any further calculations.

In the case where the adjustment is applied to results from the latest occasion, precision should improve over the first several sampling occasions (Cochran 1977, Table 12.4, p.351). When there are more than two occasions, "the only change in procedure from the second occasion is that in the regression adjustment of the estimate from the matched portion we use the improved estimate instead of the [initial estimate]" (Cochran 1977, p.349). The equations derived in Appendix D can be used to predict the improvement in precision over multiple sampling occasions. The results are shown in Figure 4-4. The improvement over time is more pronounced for the status estimates. After approximately seven occasions the incremental improvement becomes negligible.

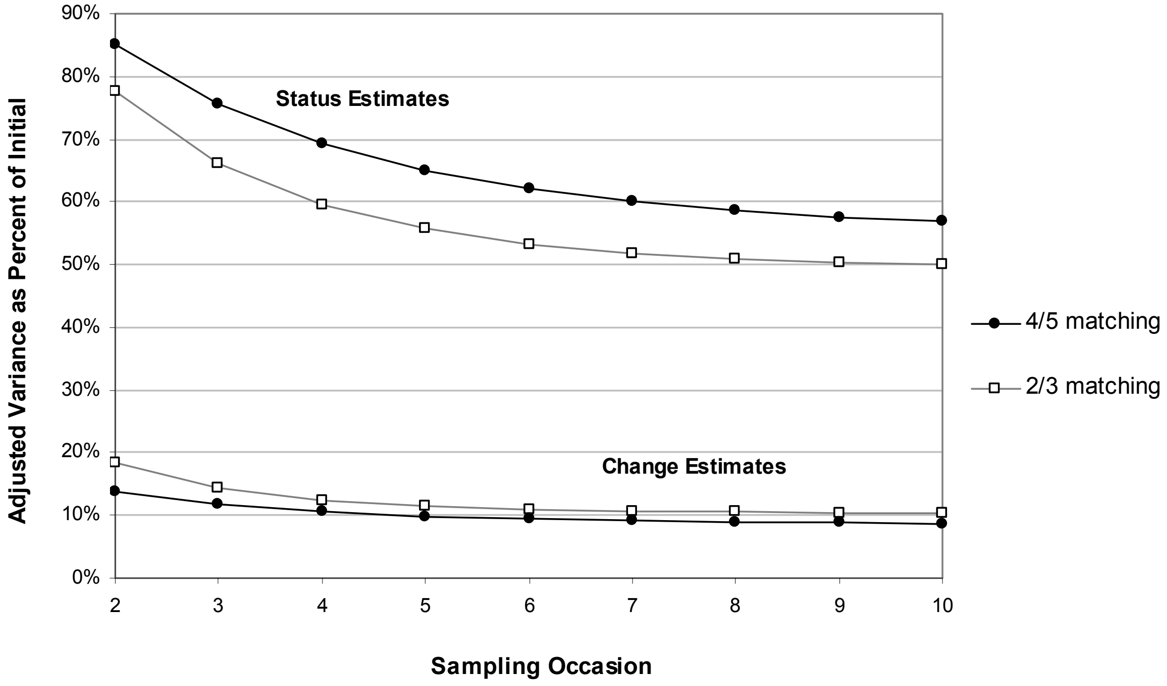


Figure 4-4. Proportional improvements in adjusted variances over multiple sampling occasions. A correlation of 0.95 for matched sites was assumed for all calculations. Two rates of matching are shown.

4.4 Summary

This chapter was initially motivated by a discrepancy between the 80% rate of site matching implemented by the SVMP for 2000-2003 and guidelines found in Cochran (1977, p.347) and Skalski (2003, p.13) that suggest the rate of matching should never exceed 50%. This chapter reviewed approaches to optimization of site rotation primarily based on material in Cochran (1977). Also, a new framework was developed for evaluating hypothetical improvements in status and change estimates that is tailored to the SVMP design.

The following key findings emerged from this chapter:

- The optimization scheme presented in Skalski (2003) and the 50% guideline of Cochran (1977) mentioned above are relevant to the optimization of site rotation for status estimates alone. These do not consider the precision of change estimates and have not been used as the basis for SVMP implementation.
- The correlations in the SVMP matched sites are very high (>0.95%).
- In practical application, Cochran indicates that the rate of rotation need not conform strictly to the calculated optimal rate. He suggests selecting among simple fractions (e.g. 2/3, 3/4 or 4/5 for rates of matching) based on reference to results from optimization calculations.
- The 2000-2003 SVMP target of 80% matching (20% rotation) is at the upper end of Cochran's recommendations where estimates of status and change are both important. This is reasonable considering that matched site correlations are very high (>95%).
- If the differential cost of sampling new and existing sites were considered, this would tend to increase the optimal rate of matching and further support the use of Cochran's upper recommendation for rate of matching.
- The new optimization scheme developed is more relevant to SVMP analysis than the other schemes discussed, although it is still diverges in key details from the SVMP situation. It suggests that the variance of adjusted status estimates may be roughly 85% of the initial variance given the level of matching used and the level of correlation observed.
- In the analyses of 2004 data, supplementary analysis should explore the use of previous adjusted results to adjust the status estimates from the most recent sampling occasion. In this case, results from the most recent sampling would be the best status estimates available. The disadvantage of this approach is that the 2001 results would not be adjusted and precision of these results would be relatively poor.

It is important to note that the optimization of change estimates as discussed in this chapter diverges from details of the SVMP analysis. The change estimates discussed here correspond to using consecutive adjusted status estimates to calculate year-to-year change rather than the paired site approach used by the SVMP. The analysis presented here clearly shows that greater site matching benefits change estimates based on status estimates. Greater matching also benefits change estimates based on the paired site analysis by making the estimate more representative of the population. However, it is not clear what effect greater matching would have on the precision of change estimates derived from the paired site technique. Regardless, the discrepancy in the approaches to calculating change suggests that a rigorous implementation of the optimized rate of matching is not warranted.

Any future SVMP work on optimization of site rotation should analyze the effects of matching on paired site change estimates. Also, the use of multiple occasions for adjustment (Yates 1960; Patterson 1950) and issues related to trend analysis (Scott and Smith 1974) should be addressed.

5 Site Rotation and Effects of Outlier Sites

This chapter was motivated by the observation that a single site, flats11, had a major impact on sound-wide results when it rotated into the sample pool in 2001. As discussed in Chapter 1, flats11 is clearly an outlier with respect to the other sites in the flats stratum. The work described in Chapter 1 included an effort to review the entire flats stratum and identify outliers for placement in a separate stratum. The purpose was to create a more homogeneous stratum and avoid producing spurious results of change that merely reflect the random rotation of outlier sites in and out of the sample pool.

The recommendations in Chapter 1 and the continued reassessment of the stratification following every sample season should minimize the danger of such spurious results. Nevertheless, a simple and somewhat artificial calculation is reported here to explore the effect of outlier sites on the long-term detection of decline in *Z. marina* area.

The following scenario was deliberately constructed and would not realistically ever occur. Two sound-wide status estimates were calculated using two sets of site data artificially constructed from actual site data from 2001 and 2002. The two sets of data were identical except flats11 was included in one but was replaced by flats20 in the other dataset.

These two datasets represent two annual samples. The sample with flats11 was used for the first four years of a ten-year dataset and the other sample was used for the remaining six years (Figure 5-1). This is an unrealistic scenario, not only because of the irregular rotation, but because the flat site with the greatest positive influence on the stratum-level estimates is replaced by the flat site with the greatest negative influence (see section 1.1, e.g. Figure 1-3 on p.9).

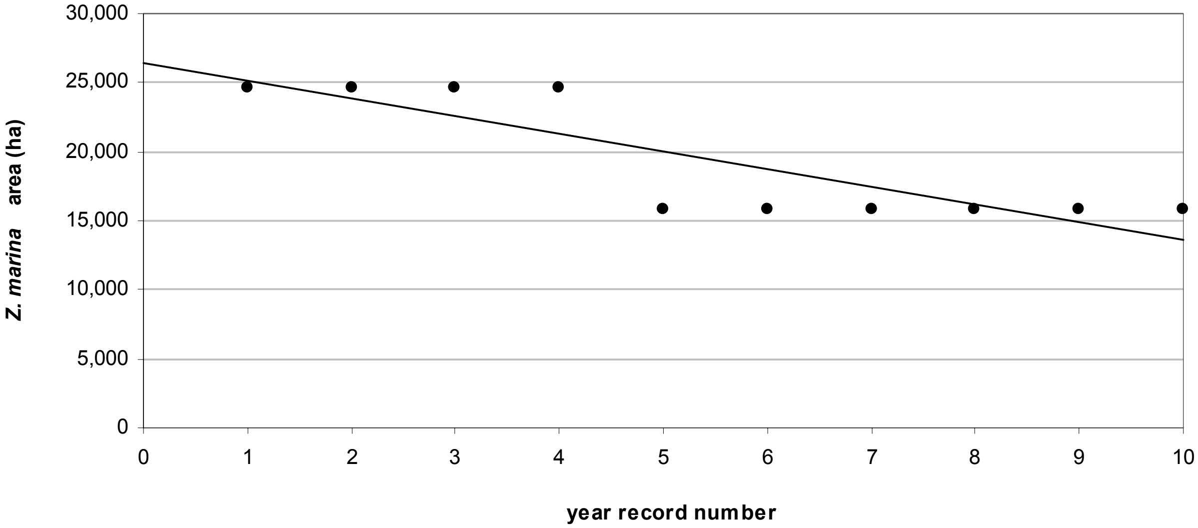


Figure 5-1. Fabricated 10-year dataset of sound-wide *Z. marina* area and linear regression line. All site results were the same for each year except flats11, which appeared in the first 4 years but was replaced by flats20 at year 5.

A least squares regression line was fit to the points giving slope $\hat{\beta}$ and regression mean square error (MSE) of

$$\hat{\beta} = -1275.5 \,{}^{\text{ha}}\!/_{\text{yr}}$$

$$\text{MSE} = 6{,}291{,}439 \text{ ha}^2.$$

The relevant *t*-statistic to test for decline is given by (Skalski 2003, Equation 46)

$$t_{m-2} = \frac{\left|\hat{\beta}\right|}{\sqrt{\dfrac{\text{MSE}}{\sum_{i=1}^{m}\left(t_i - \bar{t}\right)^2}}}$$

where $m=10$ is the sample size and $\sum_{i=1}^{m}\left(t_i - \bar{t}\right)^2$ evaluates to 82.5 yr^2. The resultant *t* statistic is $t=4.62$. A one-tailed *t*-test leads to a rejection of the null hypothesis of no decline with $p<0.001$.

This simple calculation, while artificial and extremely unrealistic, illustrates the hypothetical upper limit of the effect of outlier sites rotating in and out of the sample. In this case a highly significant determination of decline results when there is no actual change in underlying population.

The change in flats stratification (see Chapter 1) in the 2004 season eliminates the possibility of a scenario resembling the one tested here since. This change moves the three most extreme outliers, including flats11 and flats20, to a separate stratum not subject to site rotation.

Nevertheless, the results of this hypothetical scenario suggest that in practical application rotation effects may play a role. When the tests for long-term decline are applied, hypothetical scenarios based on the actual sites in the data record should be tested to assess the potential for rotation effects to bias the test.

6 Effect of Number of Transects on Site Estimates

6.1 Introduction

The overall purpose of this chapter is to investigate the reliability of SVMP site-level results and the effects of sampling intensity on reliability. This work was motivated by observations that features in the pattern of *Z. marina* at a site appear to be over- or under-sampled by the randomly selected video transects in some instances. Such features could be any distinct area with a *Z. marina* density that differs markedly from the site average. This is purely a matter of chance associated with the placement of the random transects.

These observations have led to concern about the reliability of site-level estimates of *Z. marina* area, the confidence intervals about these estimates and the tests for change between years. External reviewers independently raised this concern in more general terms (Sewell et al. in prep.).

In response to this concern, the SVMP team developed a plan to assess site-level estimates as part of the 2003 sampling and analysis effort. Six sites were identified for sampling with additional transects to support this assessment. This chapter summarizes the additional field effort and presents bootstrap analyses using data from these six intensively sampled sites.

The site-level estimates are based on the line intercept, or line intersect, sampling technique (Skalski, 2003). This technique has been well established since its adaptation to vegetation sampling in rangeland systems (Canfield 1941). Estimates of cover based on a single transect have been shown to be unbiased but estimates of the associated variance have been more problematic (Muttlak and Sadooghi-Alvandi 1993; Kaiser 1983; Lucas and Seber 1977). The SVMP design includes the sampling of multiple transects at each site thereby simplifying variance estimation. It is not clear, however, how well the variance estimator performs at relatively small sample sizes for heterogeneous sites. Currently, the general target is to sample 11 transects at each site, although in practice this target varies based on previous variance estimates at each site and scheduling. This target was based on variance estimates in results from the first three years of monitoring.

Imprecision in the estimates of *Z. marina* area is acceptable (though not desirable) as long as we know the level of precision of these estimates with some certainty. Unreliable precision estimates are problematic because this affects the interval estimates and the tests for significant change between sampling occasions.

There were four specific objectives for the work presented in this chapter.

1. Investigate the bias of site-level estimates of *Z. marina* area and variance for different numbers of transects.

2. Determine how precision in *Z. marina* area estimates varies with the number of transects.
3. Assess the reliability of sample estimates of precision and the effect of the number of transects on reliability.
4. Assess the reliability of tests for year-to-year change in site-level *Z. marina* area and the effect of the number of transects.

An additional outcome of this chapter is an assessment of whether the working target of 11 transects per site is reasonable within the overall design of the monitoring program.

Initial plans also called for collecting aerial photographs for a subset of these six intensive sites near the time of underwater videography sampling to support a broader analysis. For logistical reasons, this component was eliminated from the plan.

The chapter does not consider the effect of delineation of the sampling polygon on site-level results. Also, it does not attempt to generate spatially explicit maps of *Z. marina* at the six intensive sites as was originally planned. The issues involved in the choice of random vs. systematic sampling at the site level have been discussed previously by the SVMP team[6] and are not considered here.

6.2 Methods

6.2.1 Field Work

Six sites were selected for more intensive sampling. This number allowed for comparison of contrasting sites and could easily be accommodated within the sampling schedule. The sites were picked to represent flats and fringe sites, both large and small. Core sites were emphasized in the site selection since these sites were originally intended to be more intensively sampled over the course of the project.

The six sites are listed in Table 6-1 with the total number of random transects sampled at each site. The supplementary random transects ranged from a marginal increase over the target 11 transects (12 at core004) to more than double (25 at core002).

In addition to the random transects, other transects were deliberately placed to complete a grid pattern. These transects were intended to be used in the generation of continuous fields of *Z. marina* presence. This latter analysis was not completed as part of this study and the deliberately placed transects were not used in the analysis presented here.

Site maps showing 2003 video transects and the sampling polygons for the six intensive sites are shown in Figure 6-1 through Figure 6-6. The pattern resulting from all transects (random and deliberate) and from random transects alone are shown separately.

---

[6] An email exchange between Jim Norris and John Skalski in March 2002 very nicely highlights the key issues. See also Butler and McDonald (1983).

site code	site name	sampling stratum	region	random transects sampled
core002	Picnic Cove	flat	sjs	25
core004	Lynch Cove	flat	hdc	12
core005	Dumas Bay	wide fringe	cps	20
flats35	Nisqually Delta E	flat	cps	15
nps1363	Village Pt – Lummi Is.	narrow fringe	nps	21
swh1593	Cornell – Camano Is.	narrow fringe	swh	16

Table 6-1. The six sites selected for intensive sampling during the 2003 field season. The number of random transects shown were based on site analysis results and represent the total number of transects available for the bootstrap analysis. These numbers do not include the deliberately placed transects added to complete grid patterns.

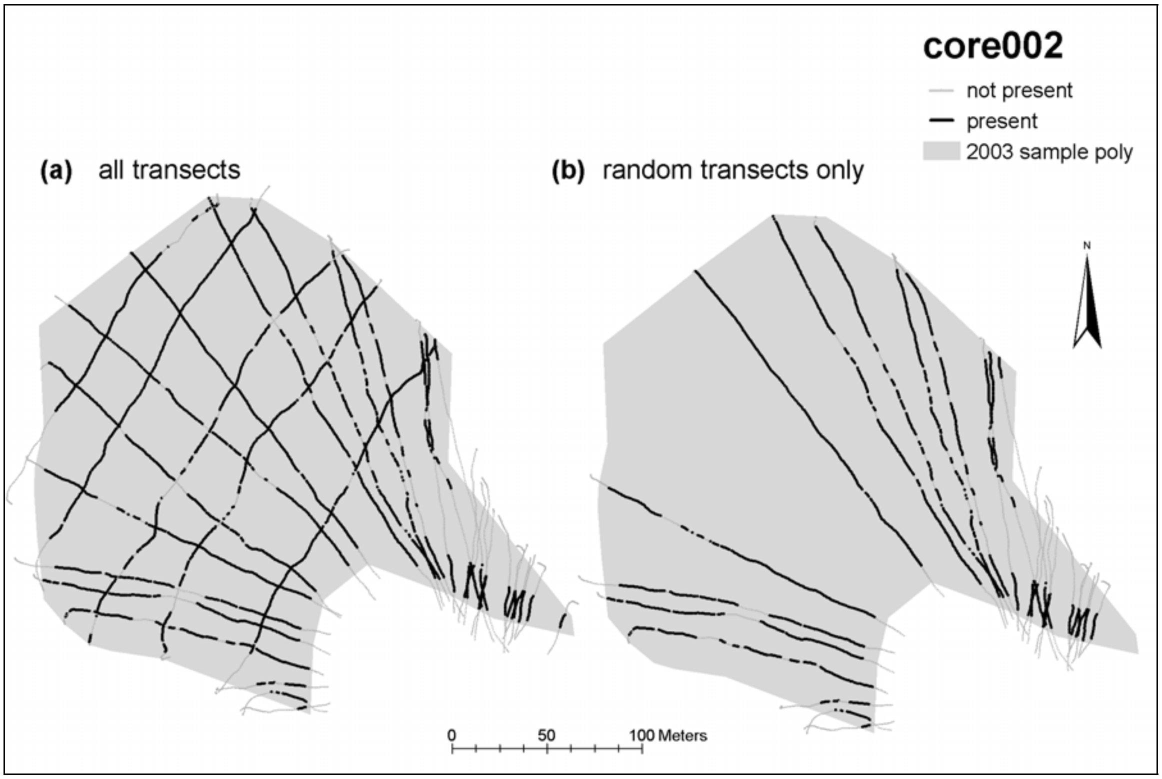


Figure 6-1. 2003 video transects for core002, Picnic Cove, showing *Z. marina* presence and absence.

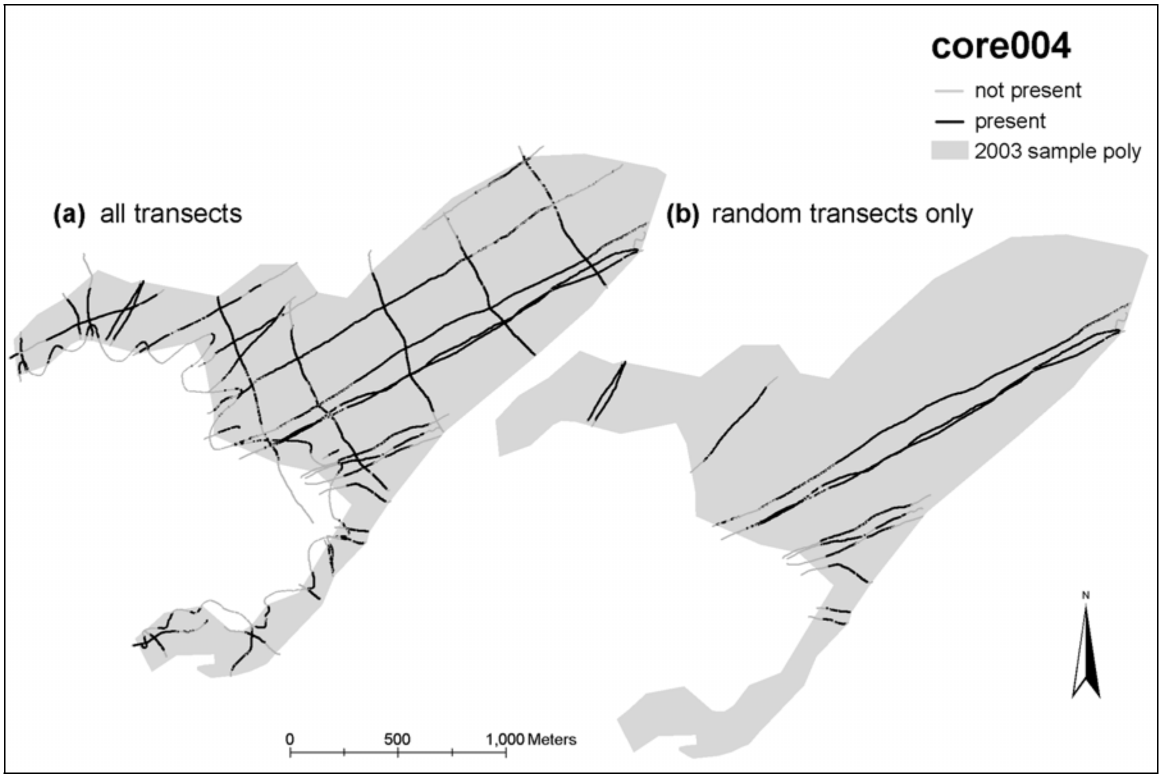


Figure 6-2. 2003 video transects for core004, Lynch Cove, showing *Z. marina* presence and absence.

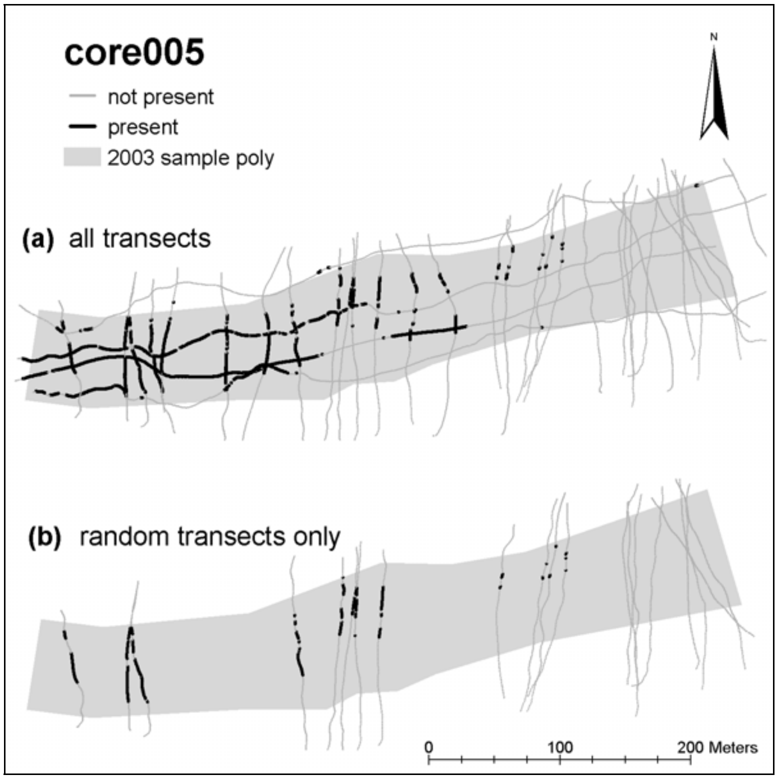


Figure 6-3. 2003 video transects for core005, Dumas Bay, showing *Z. marina* presence and absence.

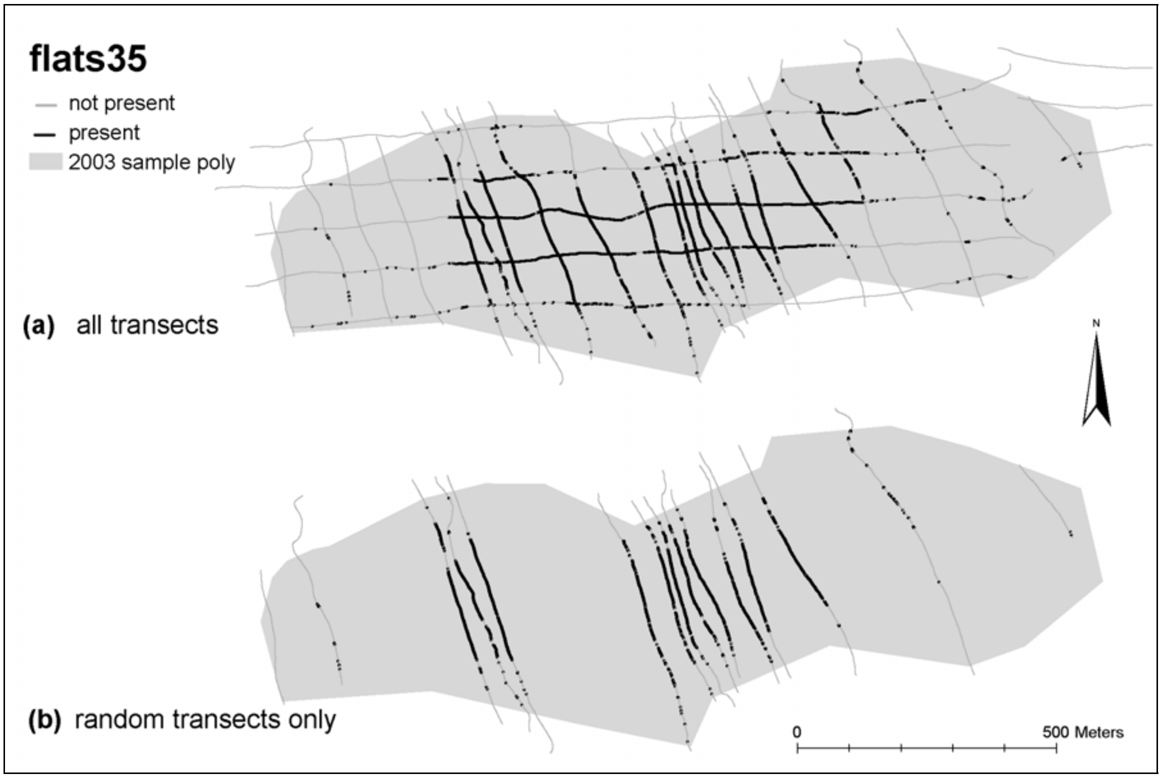


Figure 6-4. 2003 video transects for flats35, Nisqually Delta east, showing *Z. marina* presence and absence.

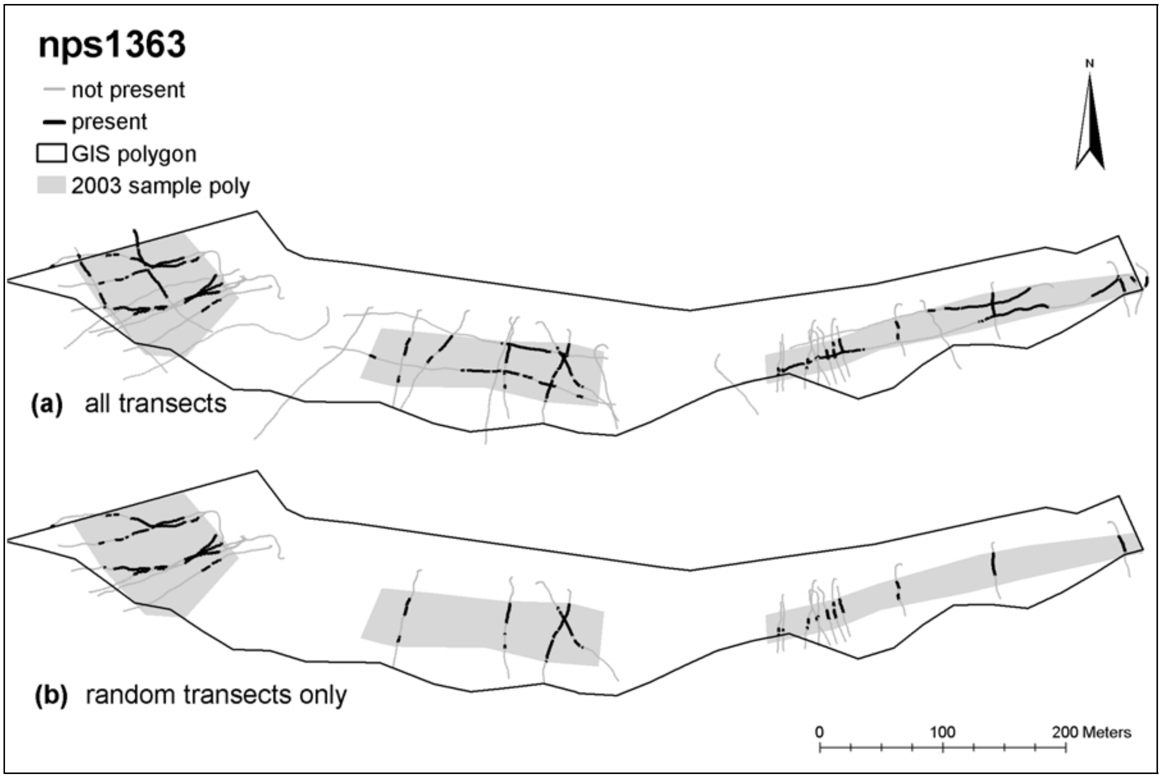


Figure 6-5. 2003 video transects for nps1363, Village Pt., Lummi Island, showing *Z. marina* presence and absence. The GIS polygon delineates the site boundary.

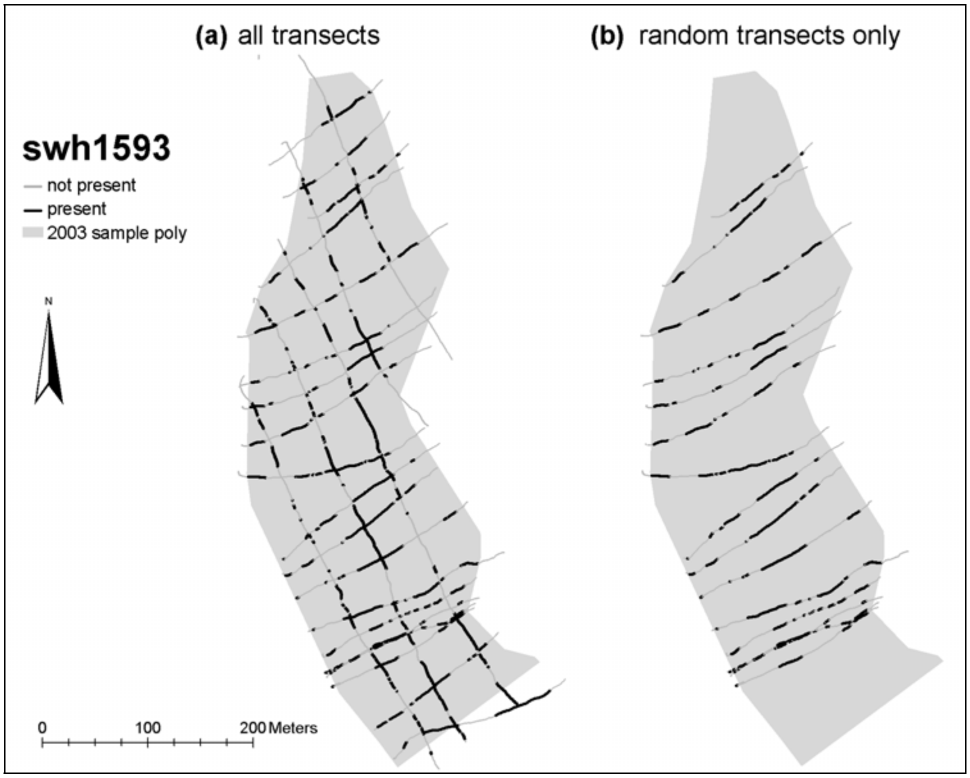


Figure 6-6. 2003 video transects for swh1593, Cornell, Camano Island, showing *Z. marina* presence and absence.

6.2.2 Analysis

The data were analyzed using bootstrap procedures (Efron 1982; Sprent 1989). The bootstrap, as applied here, is a non-parametric technique that has minimal assumptions. The key assumption is that the sample used is a good representation of the unknown population–in this case, the infinite population of possible transects at a site.

The principle underlying the bootstrap is that, in the absence of other information about a population, the values in a random sample are the best guide to the distribution and resampling the sample is the best guide to what can be expected from resampling the population.

A C program was developed for this study that utilizes the random number generator of Matsumoto and Nishimura (1998). For bootstrap estimation, the parameters (*Z. marina* area and variance) are estimated for a number of bootstrap samples taken with replacement from the original data values.

The four objectives of this chapter were addressed in the following four questions.

Question 1: Is there evidence of bias in estimates of *Z. marina* area and variance and how is this affected by sample size?

Analysis: The bootstrap is used to assess bias in the estimation of a site *Z. marina* area at each site for a range of sample sizes as follows. The estimate based on all available random transects is denoted $\hat{X}$. This is the best estimate available for site *Z. marina* area. A series of bootstrap samples of size *n* is drawn with replacement from the original sample. The estimate calculated from the i^{th} bootstrap sample using the SVMP estimator is denoted $\hat{X}_{n,i}$ and the average of all bootstrap samples is denoted $\overline{\hat{X}_{n,i}}$. For this analysis 20,000 bootstrap samples were used. Bias in the estimation of *Z. marina* area when sampling with *n* transects is estimated as

$$\text{bias}_n = \overline{\hat{X}_{n,i}} - \hat{X}.$$

The best estimate of variance in *Z. marina* area estimates when sampling with *n* transects is the bootstrap variance, i.e. the actual variance in the 20,000 bootstrap estimates, where

$$\text{bootstrap variance} = Var_{BS}\left(\hat{X}_{n,i}\right) = \frac{\sum_{i=1}^{20,000}\left(\hat{X}_{n,i} - \overline{\hat{X}_{n,i}}\right)^2}{(20,000-1)}.$$

The mean of the 20,000 variance estimates from each of the bootstrap samples is used to characterize the central tendency of the variance estimator. This is calculated as

$$\overline{\widehat{Var}\left(\hat{X}_{n,i}\right)} = \frac{\sum_{i=1}^{20,000}\widehat{Var}\left(\hat{X}_{n,i}\right)}{(20,000-1)}.$$

Bias in the variance estimate when sampling with *n* transects is estimated as

$$\text{bias}_n = \overline{\widehat{Var}\left(\hat{X}_{n,i}\right)} - Var_{BS}\left(\hat{X}_{n,i}\right).$$

Question 2: How does precision in the estimation of *Z. marina* area vary with sample size?

Analysis: J. Skalski pointed out that this question can be addressed analytically simply by using the role of sample size in the relationship between sample variance and standard error. Nevertheless, the question is answered here using Monte Carlo simulation (i.e. the bootstrap procedure). While this involves considerable additional effort it may have more intuitive appeal since it explicitly simulates the sampling process. It also reveals the actual sampling distribution and allows for the use of distribution free confidence intervals to assess precision, rather than variance or confidence intervals that assume normality.

The sampling distribution for sampling with *n* transects is simply the frequency histogram of the 20,000 estimates of *Z. marina* area. This distribution is determined at each of the six sites for a range of sample sizes. In each case the definition of a confidence interval[7] is invoked to determine the actual 95% confidence interval, i.e. the symmetric interval about

---

[7] A 95% confidence interval for a parameter θ is defined as the interval [L_1,L_2] such that the probability that θ has a value between L_1 and L_2 is 95%, that is $P[L_1 \le \theta \le L_2] = 0.95$ (Milton and Arnold 1990).

the mean *Z. marina* area that contains 95% of the 20,000 bootstrap samples. There are no distributional assumptions involved.

Question 3: How reliable are sample estimates of precision and what is the effect of sample size on reliability?

Analysis: In a sampling context, a confidence interval about a point estimate of a population parameter is itself estimated from the sample. An estimated 95% confidence interval should have the property that under repeated sampling from the population, 95% of the intervals generated cover the population parameter. By chance, 5% of the intervals generated will not cover the population parameter (Milton and Arnold 1990, p.217). In other words, the probability that the interval includes the population parameter is 95% (Zar 1999, p.99).

To answer this question, two aspects of sample confidence interval estimates are evaluated (cf. Skalski et al. 1993). First, the width of the estimated intervals is compared to the width of the actual confidence interval determined from the sampling distributions of *Z. marina* area (see Question 2). Second, the coverage or probability that a sample 95% confidence interval will include the central value of *Z. marina* area is assessed and compared to the nominal 95% value. This analysis is conducted for the target sample size, 11 transects, for each of the six intensive sites.

Question 4: How reliable is the test for change in *Z. marina* area between years?

Analysis: To address this question, the occurrence of false positive results was assessed when testing for change between successive sampling occasions (as represented by two random bootstrap samples). A false positive occurs when the test indicates a significant change when there was no real change since both bootstrap samples are drawn from the same original 2003 sample.

This analysis was done for the six sites at a range of sample sizes. In each case, the rate of false positives in 20,000 tests was calculated. For each test, the relative change between two bootstrap samples was calculated (Skalski 2002, p.25). If the 95% confidence interval about the relative change did not include the origin, the test indicated significant change.

The analysis for nps1363 did not treat the discrete sample polygons separately as described in Appendix B. To simplify the analysis, nps1363 was treated as a single polygon site. Previous analysis has shown that use of the multiple-polygon procedure (Appendix B) only has an effect if the sample polygons within a site are markedly different in terms of their *Z. marina* cover (Blain Reeves, unpub. data). The sample polygons at nps1363 are very similar and therefore the choice of analysis method should have little effect on the results.

6.3 Results

6.3.1 Assessment of Bias

The bias results calculated for both *Z. marina* area and variance estimates are shown in Figure 6-7 (core002, core004 and core005) and Figure 6-8 (flats35, nps1363 and swh1593). For each site, results are shown for sample size ranging from four to the number of transects in the original 2003 sample (Table 6-1).

In the case of *Z. marina* area estimates, bias is depicted for a given sample size as the discrepancy between the best estimate of *Z. marina* area based on the entire 2003 sample, $\hat{X}$, and the mean of the estimates calculated from the 20,000 bootstrap samples with the SVMP estimator. $\overline{\hat{X}_{n,i}}$. For each site, the calculated bias generally increases as sample size decreases. This is a smooth, montonic relationship in the cases of core002, core004 and (for the most part) nps1363 but is more erratic for the other sites, particularly core005.

The magnitude of the calculated bias in *Z. marina* area estimates is small. Even at n=4 transects, the largest bias in relative terms is approximately 4% at core002. At n=11, bias is less than 1% for all sites except core002 where it is approximately 1.3%.

In the case of variance estimates, bias is depicted as the discrepancy between the bootstrap variance and the mean of the estimates calculated from the 20,000 bootstrap samples using the SVMP estimator. Since sample size has a strong effect on variance, the scale of the y-axes shown is expanded to encompass the total range in variance. This tends to mask the level of bias at any given sample size.

On a relative basis, the calculated bias in variance estimates is moderate at n=11 for core002 (-10%), core004 (-11%) and nps1363 (-6%) but much less at core005 (+0.2%), flats35 (+0.6%) and swh1593 (-3%). At n=4, the highest level of bias is -39% (core004) and -38% (core002) with the remaining sites have much lower levels. For each site the relationship between variance estimates and sample size is monotonic and almost completely explained by sample size.

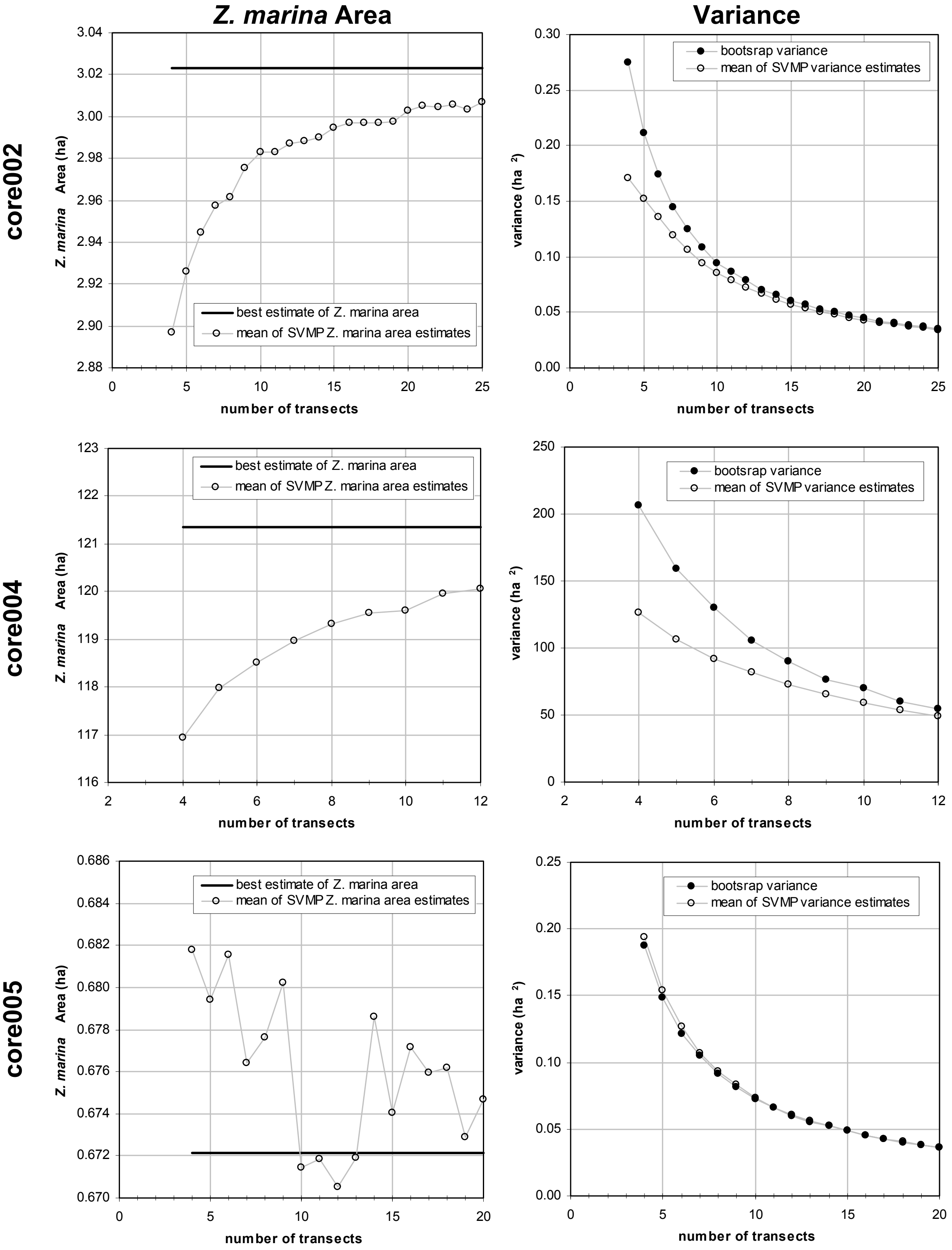


Figure 6-7. An assessment of bias in site level estimates of *Z. marina* area and variance at core002, core004 and core005 for a range of samples sizes (numbers of transects) using all 2003 random transects. The plots to the left compare (1) the estimate of *Z. marina* area at each sample size averaged from 20,000 bootstraps using the SVMP estimator with (2) the single estimate calculated using all available transects (best estimate). To assess bias in the variance estimates, the plots to the right compare the mean estimate from the 20,000 bootstraps using the SVMP estimator to the bootstrap variance.

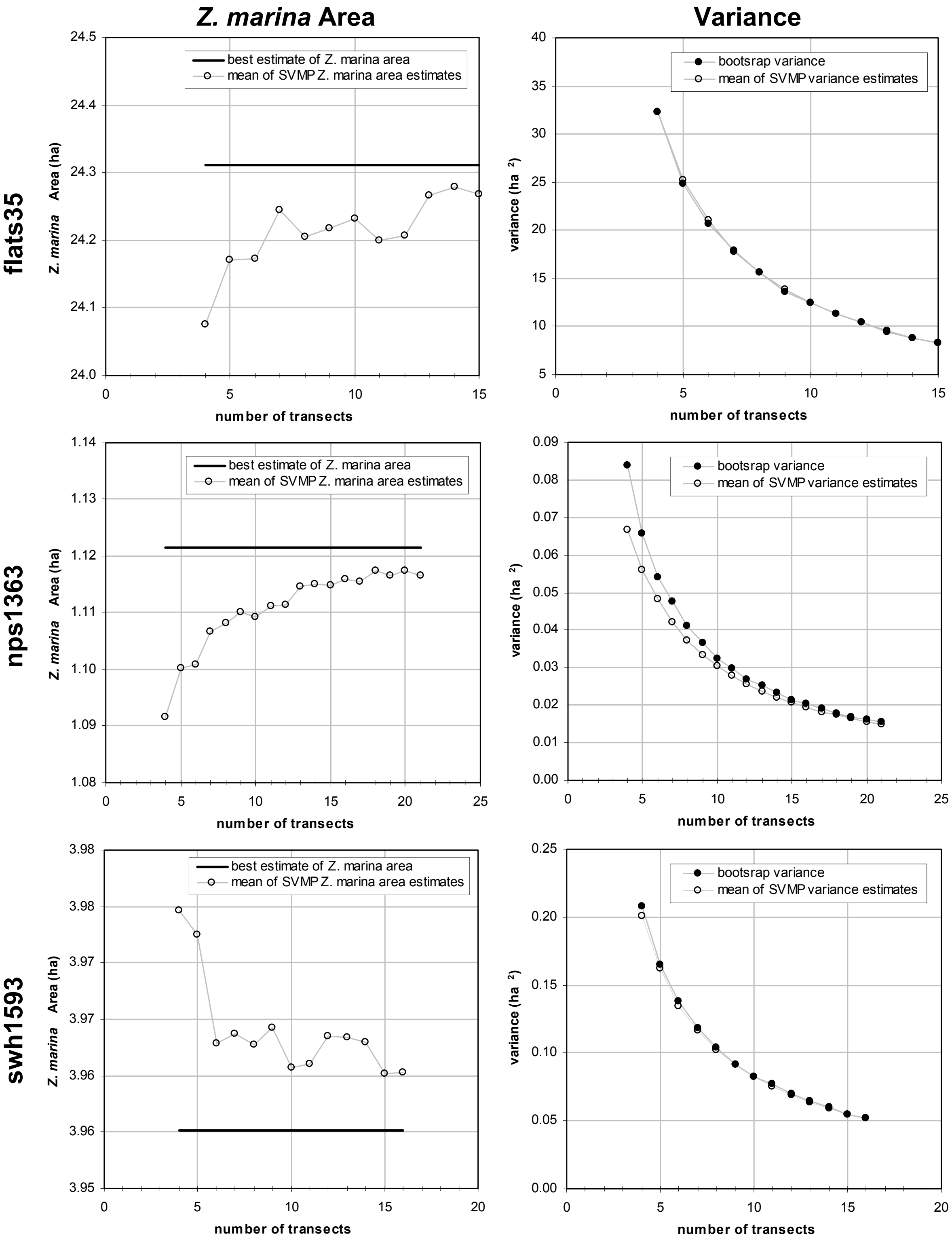


Figure 6-8. An assessment of bias in site level estimates of *Z. marina* area and variance at flats35, nps1363 and swh1593 for a range of samples sizes (numbers of transects) using all 2003 random transects. The plots to the left compare (1) the estimate of *Z. marina* area at each sample size averaged from 20,000 bootstraps using the SVMP estimator with (2) the single estimate calculated using all available transects (best estimate). To assess bias in the variance estimates, the plots to the right compare the mean estimate from the 20,000 bootstraps using the SVMP estimator to the bootstrap variance.

6.3.2 Precision and Sampling Intensity

Figure 6-9 and Figure 6-10 show the distributions of sample *Z. marina* area estimates based on the 20,000 bootstrap samples. For each distribution the 95% confidence intervals are also shown.

Several key points emerge from these results.

1. As sample size increases, the associated increase in precision is clearly reflected through less dispersion in the sampling distributions and narrower 95% confidence intervals.
2. Additional transects are most effective in increasing precision at lower numbers of transects. Precision increases dramatically from 4-5 transects to 10-12 transects, but only moderately beyond this point.
3. The sampling distributions of *Z. marina* area are clearly non-normal at low numbers of transects and some divergence from normality is evident even at higher numbers in some cases—at core002 and core004 in particular.
4. The sampling distributions tend to become more erratic at low sample sizes. This is particularly prominent for core005.

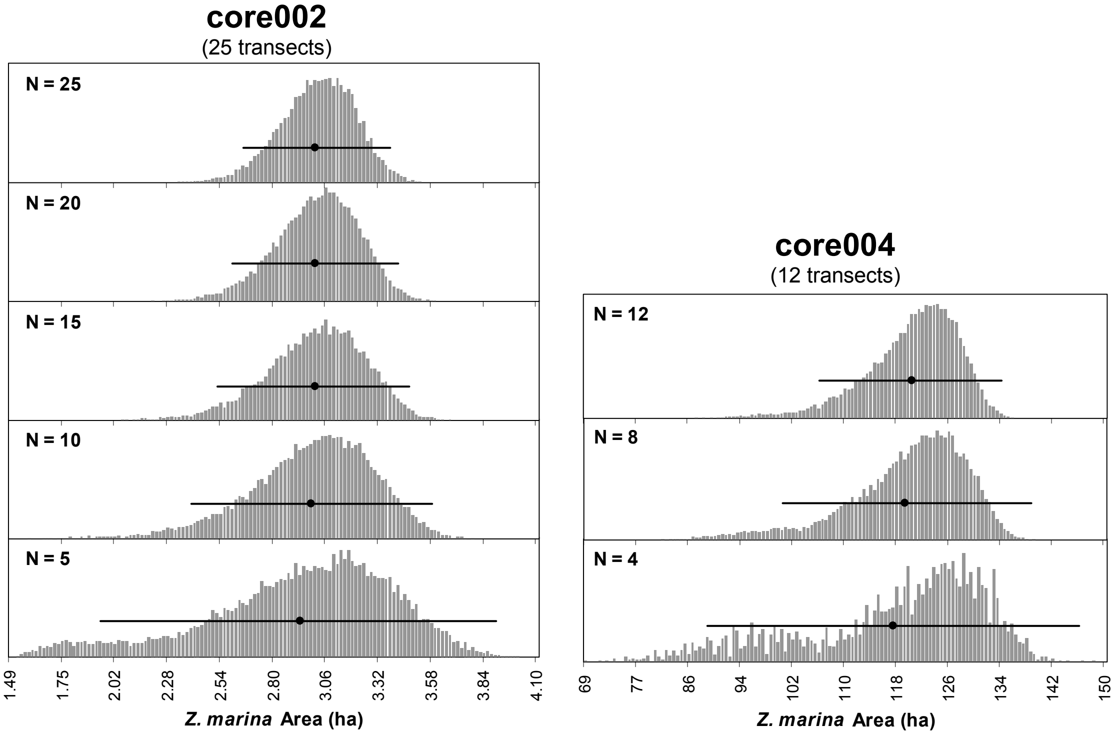


Figure 6-9. Frequency distributions of *Z. marina* area estimates for core002 and core004 for different sample sizes (N). Each distribution summarizes 20,000 estimates based on simulated sampling with replacement from the number of transects actually sampled in 2003. The number of 2003 transects is indicated in parentheses for each site. The mean is shown as a point for each distribution as well as the 95% confidence interval determined directly by coverage of 95% of the samples.

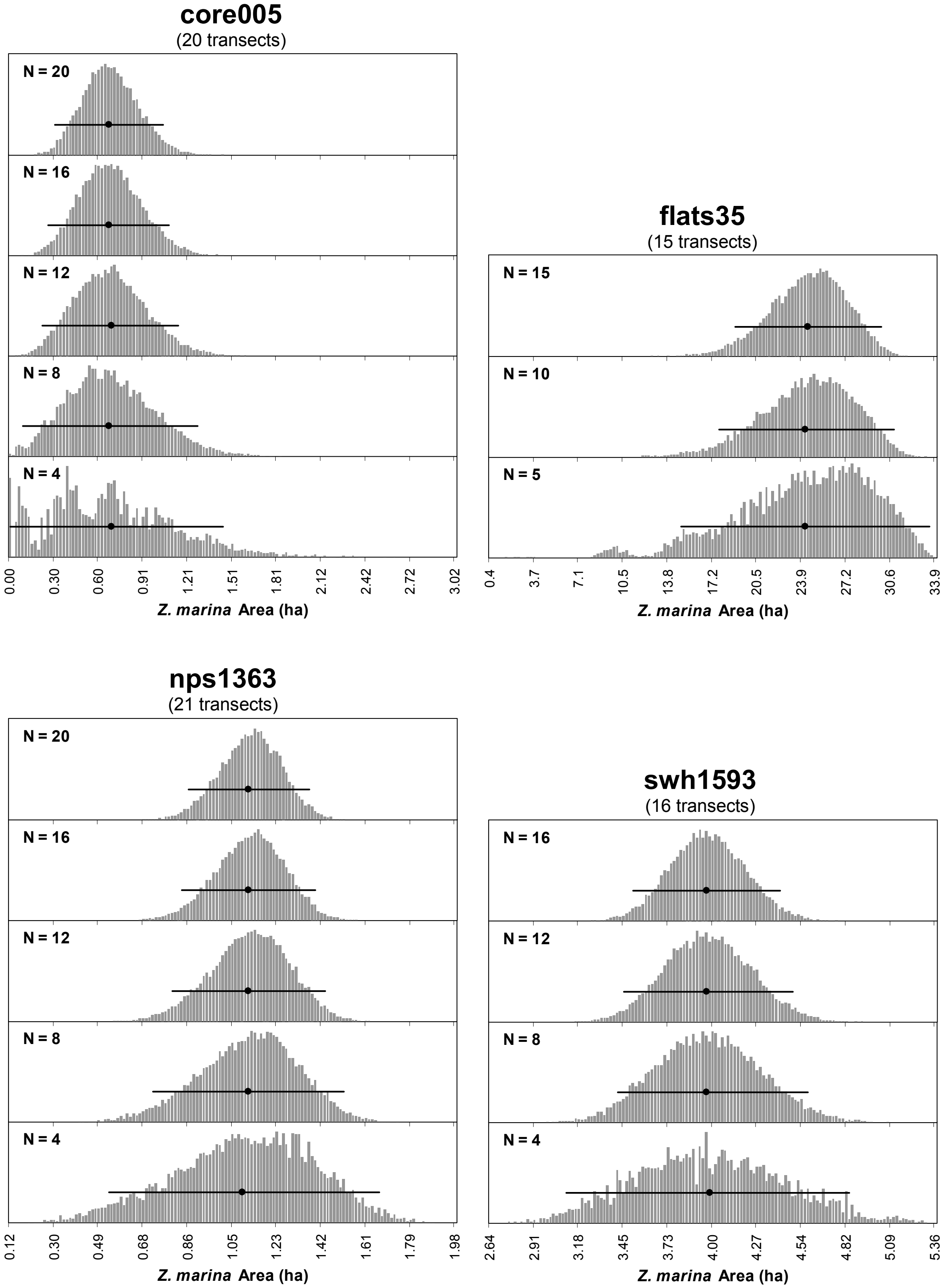


Figure 6-10. Frequency distributions of *Z. marina* area estimates for core005, flats35, nps1363 and swh1593 for different sample sizes (N). Each distribution summarizes 20,000 estimates based on simulated sampling with replacement from the number of transects actually sampled in 2003. The number of 2003 transects is indicated in parentheses for each site. The mean is shown as a point for each distribution as well as the 95% confidence interval determined directly by coverage of 95% of the samples.

6.3.3 Reliability of Precision Estimates

This section presents results on the reliability of precision estimates made with 11-transect samples. Reliability is assessed in terms of the widths and coverage of the estimated 95% confidence intervals. Interval widths are compared to the confidence intervals determined directly from the sampling distributions. The interval coverage is the percentage of estimated confidence intervals that actually cover the best estimate of *Z. marina* area. The coverage of the estimated intervals is compared to the nominal 95% coverage.

Interval Widths

Figure 6-11 shows the frequency distributions of estimated 95% confidence interval widths derived from the 20,000 bootstrap samples for each of the six sites. The *x*-axis of each distribution is the difference between the estimated widths and the actual widths on a percentage basis (percentage of actual widths). The actual widths were determined from the sampling distributions of *Z. marina* area (as shown in Figure 6-9 and Figure 6-10 for other sample sizes).

In addition to the frequency distributions, Figure 6-11 also shows the cumulative distribution function for each site. This is useful in determining the proportion of interval width estimates at different levels of departure from the actual width. This proportion is tabulated in Table 6-2 for departures of ±10% to ±40%.

The frequency distributions (Figure 6-11) give a visual impression that there is substantial dispersion in the estimates of interval widths. The tails of the distributions in some cases still have relatively high probabilities at departures of ±40%. When viewed in cumulative terms however (Table 6-2), the probabilities in the tails are less dramatic. For example, as seen in Figure 6-11 there is substantial probability of obtaining a -40% departure at core004—more than half the probability of the mode (most common value) of the distribution. As seen in Table 6-2 however, only 13% of all sample estimates have widths that diverge more than 40% from the actual width.

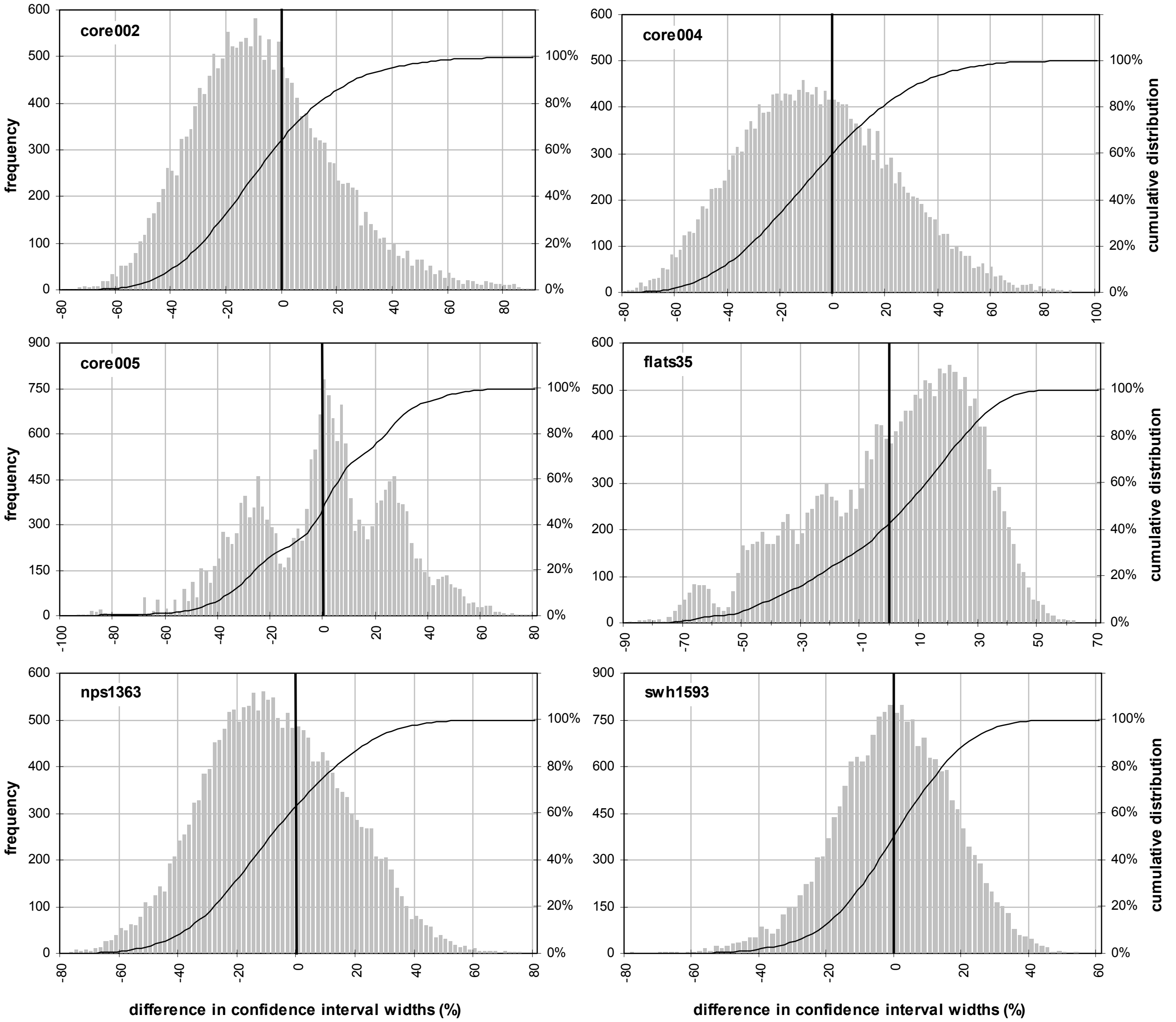


Figure 6-11. Distribution of widths of 95% confidence intervals for *Z. marina* area when sampling with 11 transects. The confidence interval widths are shown in terms of their deviation from the interval width determined from the bootstrap sampling distribution using all 2003 random transects for 20,000 bootstraps. The cumulative distribution is useful in answering questions such as 'what proportion of samples produce intervals with a negative deviation of 30% or more?' (see Table 6-2).

Difference Between Sample Confidence Interval Width and Bootstrap Confidence Interval Width	Relative Frequency					
	core002	core004	core005	flats35	nps1363	swh1593
-40% or more	9%	13%	7%	12%	9%	2%
-30% or more	20%	23%	16%	18%	18%	5%
-20% or more	34%	35%	27%	26%	33%	13%
-10% or more	50%	48%	33%	36%	49%	30%
+10% or more	23%	28%	33%	37%	23%	27%
+20% or more	14%	18%	24%	20%	13%	11%
+30% or more	8%	11%	12%	7%	6%	3%
+40% or more	4%	6%	6%	1%	2%	0%

Table 6-2. Distribution within eight categories of sample confidence interval widths for *Z. marina* area estimates (*n*=20,000 samples) when sampling with 11 transects. These categories represent varying levels of deviation from the confidence interval width derived directly from the sampling distribution produced by the bootstrap procedure.

Coverage of Best Estimate

Table 6-3 shows the coverage of the sample 95% confidence interval estimates for the six sites when sampling with different numbers of transects. Coverage is the probability that the estimated interval includes the central estimate—in this case the best estimate of *Z. marina* area calculated using the original 2003 sample of transects.

The coverage of a sample 95% confidence interval estimate is of course intended to be 95%. If coverage is less than 95% then the intervals tend to be too narrow and if the coverage is greater than 95% then the intervals tend to be too wide.

number of transects	core002	core004	core005	flats35	nps1363	swh1593
5	80.0%	77.2%	84.6%	79.4%	83.9%	85.6%
10	89.9%	88.2%	89.8%	87.4%	89.0%	91.4%
15	91.6%	–	91.5%	90.2%	90.7%	92.3%
20	92.8%	–	92.3%	–	91.6%	–
25	93.2%	–	–	–	–	–

Table 6-3. Coverage of 95% confidence intervals for site *Z. marina* area at different sample sizes (number of transects) determined from bootstrap analysis. For example, a 95% confidence interval generated from an individual sample of 10 transects at core002 actually had an 89.9% probability of covering the best estimate of *Z. marina* area. The best estimate was calculated using all available random transects.

The results in Table 6-3 indicate that interval estimates increasingly approach the nominal 95% coverage as the number of transects increase. For all sites and all sample sizes shown, the coverage is less than nominal indicating that the intervals generally should be wider to achieve 95% coverage. However, the discrepancies are only moderate. For example, when sampling with ten transects the coverage ranges from 87.4% to 91.4%. The coverage when sampling with eleven transects was not calculated, but it is reasonable to infer that in practice the coverage of 95% confidence interval estimates is generally closer to 90%.

6.3.4 Reliability of Test for Year-to-Year Change

Table 6-4 shows the frequency of false positive results in the tests for year-to-year change in *Z. marina* area. Each result is based on 20,000 tests using two different bootstrap samples (40,000 total bootstraps).

Since testing was conducted at α=0.05, the expected rate of false positive results is 5%. The results generally approach this expected rate as the sample size increases, but the rate of false positives is substantially higher when sampling with only five transects—up to 15.6% in the case of core002. Flats35 is an exception in that the rate of false positives is less than 5% for all sample sizes tested and the rate increases slightly as the sample size increases from 10 to 15 transects.

At *n*=10, the results range from 3.7% (flats35) to 10.3% (core005). The general target of *n*=11 was not tested, but results would likely be similar. The worst-case scenario in these results (core005) indicate that tests for change at α =0.05 would actually have a true α value closer to α=0.10 when sampling with *n*=11 transects.

number of transects	core002	core004	core005	flats35	nps1363	swh1593
5	15.6%	14.7%	13.9%	4.0%	11.7%	9.5%
10	7.3%	7.0%	10.3%	3.7%	8.1%	7.0%
15	6.0%	–	8.9%	3.9%	6.8%	6.4%
20	5.7%	–	7.8%	–	6.1%	–
25	5.5%	–	–	–	–	–

Table 6-4. Relative frequency of false positive results indicating change when comparing two site-level samples at α=0.05. All 2003 transects were used to represent 2003 status as well as a hypothetical subsequent year so that the two years in the simulation were identical. The values in the table summarize results from 20,000 tests that calculate the relative change between two bootstrap samples.

6.4 Discussion and Summary

The primary purpose of this chapter was to assess the reliability of site-level SVMP results. The bootstrap results suggest that there are discrepancies between site-level *Z. marina* status estimates, the tests for change and their nominal statistical properties. However, the magnitude of these discrepancies is modest and do not provide a basis for revising the site-level sampling design.

Three lines of evidence support the conclusion that the status estimates of *Z. marina* area are acceptably reliable:

- Analytical results in the literature demonstrate that point estimates of cover from line intercept sampling are unbiased (Lucas and Seber, 1977).
- The empirical results of bias in the point estimates presented in this chapter were on the order of 1% or less when sampling with n=11 transects.
- The empirical confidence intervals presented for *Z. marina* area were reasonably close to the nominal 95% coverage, ranging from 87.4% to 91.4% when sampling with n=10 transects.

Furthermore, the results presented support the conclusion that the test for change in site *Z. marina* area is acceptably reliable. The empirical results presented suggest that the significance level of the test for change can be expected to be reasonably close to the nominal α=0.05. Results from the six intensive sites investigated ranged from α=0.04 to α=0.10 when sampling with n=10 transects.

While the level of reliability may support the continued use of the site-level sampling and analysis procedures, it is critical to consider the discrepancies found when interpreting site-level results. In the absence of additional information, any interpretation must be rejected if it assumes a higher level of reliability than found in the simulation results.

It is considerably more complex to assess whether the general target of sampling eleven transects per site is adequate. The results presented clearly reflect the maxim that precision benefits from greater sample sizes (Figure 6-9 and Figure 6-10). For the site-level results, this is an argument for maximizing the number of transects at each site. This would increase both the precision of the site *Z. marina* area estimates and the reliability of tests for change. From this perspective, it might even be appropriate to reduce the number of sites sampled in order to intensify effort beyond the target 11 transects at the remaining sites.

However, the objective of producing robust site-level estimates competes with the primary monitoring objective of producing robust estimates at the sound-wide scale. As seen in section 2.5 (Table 2-5, p.49), in the sound-wide results the uncertainty associated with selecting only a sample of sites to measure (sampling error) is from one to three orders of magnitude larger than the uncertainty associated with line intercept sampling at a site (measurement error). This is an argument for reducing the number of transects sampled at each site and redirecting the effort into sampling more sites less intensively. This would increase the precision of the sound-wide results at the expense of precision in the site-level results.

The monitoring design clearly must include subjective choices based on the relative importance placed on these two competing objectives. The results presented in this chapter indicate that the site-level results are not optimal but they are better than if the sound-wide results alone were optimized. This is consistent with the objective of producing meaningful results at different scales and suggests that the current target of sampling eleven transects at each site is appropriate.

It is important to keep in mind that the results and interpretation presented in this chapter are based on the fundamental assumption of the bootstrap technique—that the original samples used to study the six sites is representative of the infinite population of transect samples from each site. It is impossible to definitively determine the representativeness of these samples. However, a review of the spatial pattern of *Z. marina* when the deliberately placed transects are included (Figure 6-1 through Figure 6-6) does not suggest that the samples grossly misrepresent the patterns at the sites. In the absence of more information, the results presented in this chapter are the best indicator of the performance of the site-level results.

There are many questions regarding the site-level analysis that remain unanswered and provide opportunities for future investigation. One of the most important issues is the effect of uncertainty in the delineation of the sample polygon on site-level estimates and the effect of changes to this delineation. Also a high priority for the SVMP is the development of multi-parameter tests for change at the site-level. A less immediate issue, but one that needs to be addressed and has also been brought up by external reviewers (Sewell et al., in prep.), is the potential for incorporating airborne remote sensing into site-level monitoring.

It would be interesting from a spatial sampling perspective to study the effects of heterogeneity at different scales on site-level estimates. At scales less then 10^0 m, heterogeneity is ignored by reducing the pattern of *Z. marina* cover to present or absent. Heterogeneity at scales of 10^0-10^1 m is often referred to as patchiness and is summarized in the fractional cover of an individual transect. Heterogeneity at larger scales gives rise to variability between transects and site-level variance estimates. When heterogeneity at this larger scale is organized into coherent zones this could have implications for the appropriate level of sampling effort and stability of the results. Core005 falls into this category where 7 of the 20 transects in the original sample did not intersect any *Z. marina* and these are grouped together in one zone (Figure 6-3).

Ultimately, core005 did not differ in a qualitative sense from the other sites in terms of confidence interval coverage and the reliability of the test for change. However, the bootstrap sub-sampling at the site produced erratic bias results (Figure 6-7), an erratic sampling distribution at low *n* (Figure 6-10) and a unique tri-modal distribution of confidence interval widths (Figure 6-11). The effects of heterogeneity at different scales could be addressed with the use of mathematically generated fields and Monte Carlo sampling. Such generation of spatial data is well established with point data (Diggle 1983) but is considerably more complex with areal data (O'Sullivan and Unwin 2003, p.71). Pielou (1977) presents an approach to generating random areal patterns but it may be easier to work with point patterns of *Z. marina* shoots rather than adopt the Pielou method to generate a pattern of *Z. marina* patches.

Current plans call for dedicating a limited effort to these issues during 2005-2006. The effects of sample polygon delineation and development of multi-parameter tests for change will have the highest priority.

7 Recommendations and Future Priorities

The SVMP is based on a sophisticated statistical framework that was designed without the benefit of pre-existing data. The work described in this report used data from the first four years of the SVMP to assess whether specific aspects of the sampling and analysis design are performing as intended. Several recommendations have emerged from this work that are summarized below.

Many specific issues could not be addressed within the scope of the work reported here. These include previously identified issues and new questions that emerged during the course of the study. Priority issues for future work are also listed below.

7.1 Recommendations

- Four specific recommendations were made for changes to the division of large embayments into discrete flats sampling sites. These changes improve the consistency of the flats sampling frame. These recommendations have already been implemented for the 2004 sampling season.
- A change to the stratification of flat sites was made in order to improve overall precision of estimates for the flat sites. Three very large, anomalous flats sites were removed from the previous flats stratum and moved to a new "persistent flats" stratum that is subject to complete census every year. This recommendation was also implemented prior to the 2004 sampling season.
- A systematic procedure was developed to help create the new flats stratum. This procedure should be used in the future as new flats data is obtained to reassess the flats stratification. In the future, however, if additional sites are moved to the persistent flats stratum, this should only be done after the sites have completed their role as randomly selected sites in the residual rotational flats stratum.
- When data become available to test for long-term decline in *Z. marina* area, starting with the 5-year test, new power curves should be generated that are tailored to the existing data. These should be used to determine the optimal significance level, α, for testing and help interpret the test results.
- The 20% rate of site rotation should be retained. Even if the SVMP shifts emphasis more strongly to trend estimates at the expense of status estimates, site rotation should be maintained (see Skalski 1990).
- The general target of sampling a minimum of 11 transects at each site should be maintained.

7.2 Future Priorities

- A Monte Carlo study based on an expanded model flats dataset should be implemented to definitively answer whether there is significant bias in SVMP estimates associated with low sample size, skewed distributions and the presence of measurement error. Work on this study has already been initiated.
- A power analysis should be performed for the paired site change detection tests, both at the sound-wide and regional scales.
- A simulation study should be conducted to assess the paired site change detection tests for bias at the small sample sizes available for regional estimates. When recommendations for minimal sample sizes are generated, the optimization of site rotation should be revisited.
- A study should be conducted to determine the optimal allocation of sampling effort among the strata. Results in this report suggest that the wide fringe strata would benefit the most from additional sites and it may be appropriate to shift resources from the narrow fringe stratum to the wide fringe stratum.
- The robustness of the sound-wide status estimates should be thoroughly evaluated under the 2004 change to the flats stratification. If these estimates are not sufficiently robust, an alterative approach to assessing long-term decline should be developed that builds on the paired site change detection procedures.
- Multi-parameter tests should be developed for identifying individual sites undergoing change in *Z. marina* area. These tests could involve the site-level *Z. marina* area estimates, maximum depth estimates and patchiness estimates over single or multiple years. The purpose would be to produce a list of candidate sites that might benefit from increased scrutiny and might be of particular interest to local community and governmental entities.
- Procedures for testing for long-term change at the site level should be developed along with estimates of statistical power. These would have ongoing application at core sites and rotating application at sites that have been in the sample pool for 4-5 years.

8 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 State Department of Natural Resources, Olympia, Washington, 60pp. plus appendices.

Berry, H., and R. Ritter, 1997, *Puget Sound Intertidal Habitat Inventory 1995: Vegetation and Shoreline Characteristics Classification Methods*, Washington Department of Natural Resources, Aquatic Resources Division, Olympia, WA.

Butler, S.A. and L.L. McDonald, 1983, Unbiased Systematic Sampling Plans for the Line Intercept Method, *Journal of Range Management*, **36(4)**:463-468.

Canfield, R.H., 1941, Application of the Line Interception Method in Sampling Range Vegetation, *Journal of Forestry*, **39**:388-394.

Cochran, William G., 1977, *Sampling Techniques*, 3rd edition, John Wiley & Sons, 428pp.

Deutsch, C.V. and A.G. Journel, 1998, *GSLIB: Geostatistical Software Library and User's Guide*, 2nd edition, Oxford University Press, New York, 369pp. with CDROM.

Diggle, P.J., 1983, *Statistical Analysis of Spatial Point Patterns*, Academic Press, New York, 148pp.

Dowty, P.R., B.R. Reeves, A.T. Sewell, H. Berry, S. Wyllie-Echeverria and T.F. Mumford, Jr., in prep., *Puget Sound Submerged Vegetation Monitoring Project: 2003-2004 Report*, Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, WA.

Efron, B., 1982, *The Jackknife, the Bootstrap and Other Resampling Plans*, Society for Industrial and applied Mathematics, Philadelphia, 92pp.

Evans-Hamilton, Inc. and D.R. Systems, Inc., 1987, *Puget Sound Environmental Atlas–Volumes I and II*, Prepared for U.S. Environmental Protection Agency, Puget Sound Water Quality Authority and the U.S. Army Corps of Engineers, Seattle, Washington.

Finlayson, D.P., R.A. Haugerud, H. Greenberg and M.G. Logsdon, 2000, *Puget Sound Digital Elevation Model*, University of Washington, (http://students.washington.edu/dfinlays/pugetsound/)

Kaiser, L., 1983, Unbiased Estimation in Line-Intercept Sampling, *Biometrics*, **39**:965-976.

Kulldorff, G., 1963, Some problems of optimum allocation for sampling on two occasions, *Rev. Int. Stat. Inst.*, **31**:24-57.

Lucas H.A. and G.A.F. Seber, 1977, Estimating coverage and particle density using the line intercept method, *Biometrika*, **64(3)**:618-622.

Matsumoto, M. and T. Nishimura, 1998, Mersenne Twister: A 623-Dimensionally Equidistributed Uniform Pseudo-Random Number Generator, *ACM Transactions on Modeling and Computer Simulation*, **8(1)**:3--30.

Milton, J.S. and J.C. Arnold, 1990, *Introduction to Probability and Statistics*, McGraw-Hill.

Muttlak, H.A. and S.M. Sadooghi-Alvandi, 1993, A Note on the Line Intercept Sampling Method, 1993, *Biometrics*, **49**:1209-1215.

Needham, E.K. and E.L. Lanzer, 1993, The Puget Sound Environmental Atlas Update: Context and GIS Database Integration Issues, *Computers, Environment and Urban Systems*, 17:409-424.

Norris, J.G., S. Wyllie-Echeverria, J.R. Skalski and R.C. Zimmerman, 2001, *Washington State Department of Natural Resources Submerged Vegetation Monitoring Project 2000 Final Report*, Revised Draft September 26, 2001, Report from Marine Research Consultants submitted to Washington Dept. of Natural Resources.

O'Sullivan, D. and D.J. Unwin, 2003, *Geographic Information Analysis*, John Wiley & Sons, New Jersey, 436pp.

Patterson, H.D., 1950, Sampling on successive occasions with partial replacement of units. *Joural of the Royal Statistical Society*, **B12**:241-255.

Pielou, E.C., 1977, *Mathematical Ecology*, John Wiley & Sons, New York, 385pp.

Rao, J.N.K., 1968, Some small sample results in ratio and regression estimation, *Journal of the Indian Statistical Association*,**6**:160-168.

Rao, J.N.K. and J.E. Graham, 1964, Rotation designs for sampling on repeated occasions, *Joural of the American Statistical Association*, **59**:492-509.

Ritter, R.A., H.D. Berry, B.E. Bookheim and A.T. Sewell, 1999, *Puget Sound Habitat Intertidal Inventory 1996: Vegetation and Shoreline Characteristics Classification*

Methods, Washington Department of Natural Resources, Aquatic Resources Division, Olympia, WA.

Scott, A.J. and T.M.F. Smith, 1974, Analysis of repeated surveys using time series methods, *Joural of the American Statistical Association*, **69**:674-678.

Sewell, A.T., S. Wyllie-Echeverria, H.D. Berry, T.F. Mumford, B.R. Reeves, P.R. Dowty, in prep., *Submerged Vegetation Monitoring Project: 2003 External Project Review*, Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, WA.

Skalski, J.R., 2003, Statistical Framework for Monitoring *Zostera marina* (Eelgrass) Area in Puget Sound, Appendix L in Berry et al., 2003, *Puget Sound Submerged Vegetation Monitoring Project: 2000-2002 Monitoring Report*, Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, Washington.

Skalski, J.R., 1990, A Design for Long-Term Status and Trends Monitoring, *Journal of Environmental Management*, **30**:139-144.

Skalski, J.R., A. Hoffman, B.H. Ransom and T.W. Steig, 1993, Fixed-Location Hydroacoustic Monitoring Designs for Esimating Fish Passage Using Stratified Random and Systematic Sampling, *Canadian Journal of Fisheries and Aquatic Sciences*, **50**:1208-1221.

Skalski, J.R. and D.S. Robson, 1992, *Techniques for Wildlife Investigations: Design and Analysis of Capture Data*, Academic Press, San Diego, 237pp.

Sprent, P., 1989, *Applied Nonparametric Statistical Methods*, Chapman and Hall, New York, 259pp.

Uhrich, M.A. and T.S. McGrath, 1997, *Conversion of Environmental Data to a Digital-Spatial Database, Puget Sound Area, Washington*, Open-File Report 95-359, U.S. Geological Survey, Prepared in cooperation with the Puget Sound Water Quality Authority, Portland, Oregon, 117pp.

Washington State Department of Ecology, 1979, *Coastal Zone Atlas of Washington*, Volume Four: Island County, Pub. No. DOE 77-21-4, Volume Five: Snohomish County, Pub. No. DOE 77-21-5, Olympia, WA.

Wilkinson, D., 2004, *Normal random variables via the Box-Muller method*, Web page designed to support coursework at Department of Statistics, University of Newcastle upon Tyne. Implementation of Box-Muller in C downloaded in June 2004. http://www.mas.ncl.ac.uk/~ndjw1/teaching/sim/transf/norm.html.

Yates, F., 1960, *Sampling Methods for Censuses and Surveys*, Charles Griffin and Co., London.

Zar, J.H., 1999, *Biostatistical Analysis*, 4th edition, Prentice-Hall.

Appendix A. Sampling and Analysis Review Plan

The table below lists the questions that were finalized on 12/3/03 to serve as the basis for the review of SVMP sampling and analysis issues.

The SVMP team prioritized the questions. Only the first six questions are addressed in this study. The remaining questions may be revisited at a later date.

Rank	Basin[8] or Site-Level	Question and Approach
1	Basin	What is the relationship between number of sites sampled per season and our change detection limit at the study area and regional scales? Would an increase in sites within the constraints of the budget enhancement lead to a substantial improvement in results?
2	Basin	Is the large improvement in CV that resulted from adjusting the 2002 results using 2003 data statistically defensible? If this can be considered in instability in the analysis, how can the methodology be improved (num. of paired/unpaired sites)?
3	Basin	What are the key considerations in dividing the current flats stratum into two strata? Would the rotation effects be reduced given the loss in sample size?
4	Basin	Under our current study design, what criteria are suitable for optimizing the level of site rotation? Given these criteria, what is the optimal level of rotation?
5	Basin	How do the effects of site rotation compare to our stated detection limit? Given these effects, can we meet our detection limit objective?
6	Site-Level	Using data from the 2003 intensive sites, what relationships emerge between number of random transects and variance?
7	Site-Level	Do estimates of eelgrass area from kriging fall within the confidence interval from our standard site-level analysis?
8	Site-Level	What are the effects of uncertainty in sample polygon delineation and year-to-year change in this delineation on site-level results?
9	Basin	Examine reallocation of sampling effort among strata (p.11 of report).
10	Basin	How do the two methods in the statistical framework compare for the analysis on long-term (>5 year) trends? These methods are change detection in paired sites and the comparison of adjusted status estimates.
11	Site-Level	Are there examples in the literature for single systematic sample analysis that provide estimates of variance? What are the tradeoffs of this approach versus random transects?
12	Basin	What are the tradeoffs between our current study design and a mixed design that combines both remote sensing (aerial) and underwater videography transect sampling?
13	Basin	What are the effects of uncertainty in our base GIS layers (particularly the –20ft contour) on regional and study-area extrapolations?

---

[8] 'Basin' denotes questions relevant to analysis at the level of SVMP regions and the entire study area.

Appendix B. Site-Level Estimates from Multiple Sampling Polygons

The following is a slight modification of analysis methods in the Statistical Framework (Skalski 2003) applicable to sites with multiple distinct sample polygons.

B.1 *Z. marina* Area

Total site eelgrass area ($\widehat{X}$) is estimated by a simple sum of the estimates from each distinct sample polygon ($\widehat{X_i}$):

$$\widehat{X} = \sum \widehat{X_i}$$

Eqn B.1 (Area)

The variance of the estimate of total site eelgrass area $\widehat{Var}\left(\widehat{X}\right)$, can be estimated as the sum of variance estimates from each distinct sample polygon:

$$\widehat{Var}\left(\widehat{X}\right) = \sum Var\left(\widehat{X_i}\right)$$

Eqn B.2 (Variance)

This last equation is based on the assumption that the estimates from the individual polygons are uncorrelated. Specifically covariance terms have been ignored. This does not refer to correlation between polygons over time (e.g. in response to the same environmental driver) but correlation under repeated sampling (i.e. the selection of random transects in one has nothing to do with those selected in the other).

The associated estimate of standard deviation of the total site eelgrass area statistic (i.e., the standard error) is

$$\widehat{SE}\left(\widehat{X}\right) = \sqrt{\widehat{Var}\left(\widehat{X}\right)}$$

Eqn B.3 (Standard Error)

and, following the recommendation of the Statistical Framework, the Z-statistic confidence interval on the estimate of total site eelgrass area is

$$\widehat{X} \;\pm\; Z_{1-\frac{\alpha}{2}} \cdot \widehat{SE}\left(\widehat{X}\right)$$

Eqn B.4 (Confidence Interval)

where $Z_{1-\frac{\alpha}{2}}$ is one of the following:

α	$Z_{1-\frac{\alpha}{2}}$
0.05 (95% CI)	1.960
0.10 (90% CI)	1.645
0.20 (80% CI)	1.282

Table-B-1. Z values for α=0.05, 0.10 and 0.20.

B.2 *Z. marina* Fraction

The aggregated site-level eelgrass fraction is most readily estimated by dividing total *Z. marina* area by total sample polygon area, that is

$$\hat{p} = \frac{\hat{X}}{\sum E_i}$$

Eqn B.5 (*Z. marina* Fraction)

where

$\hat{p}$ = estimated mean site-level eelgrass fraction

E_i = area of sample polygon *i*, and the denominator is a sum over all sample polygons in the site.

To derive an equation for variance, we start by expressing the equation for *Z. marina* fraction (Eqn B.5) above in equivalent terms as the weighted sum

$$\hat{p} = \frac{\sum \hat{X}_i}{\sum E_i} = \frac{\sum \left(E_i \cdot \hat{p}_i \right)}{\sum E_i} = \sum \left(w_i \cdot \hat{p}_i \right)$$

where $\hat{p}_i$ is the estimate of *Z. marina* fraction for sample polygon *i*, and the associated weighting is

$$w_i = \frac{E_i}{\sum E_i}$$

The estimate for variance of the *Z. marina* fraction estimate can then be expressed as

$$\begin{aligned} \widehat{Var}\left(\hat{p}\right) &= \widehat{Var}\left[\sum \left(w_i \cdot \hat{p}_i \right)\right] \\ &= \sum \left[\widehat{Var}\left(w_i \cdot \hat{p}_i \right) \right] \end{aligned}$$

This can be further manipulated into a computational formula for the variance estimate

$$\widehat{Var}\left(\hat{p}\right) = \sum \left\{ \left[w_i^2 \right] \left[\widehat{Var}\left(\hat{p}_i \right) \right] \right\}$$

Eqn B.6 (Variance)

As before, since the standard error is simply the standard deviation of the distribution of estimates of our statistic (in this case, $\hat{p}$), it is estimated from variance as

$$\widehat{SE}\left(\hat{p}\right) = \sqrt{\widehat{Var}\left(\hat{p}\right)}$$

Eqn B.7 (Standard Error)

and the confidence interval on $\hat{p}$ is calculated as

$$\hat{p} \pm Z_{1-\frac{\alpha}{2}} \cdot \widehat{SE}\left(\hat{p}\right)$$

Eqn B.8 (Confidence Interval)

using the same Z values as above.

B.3 References

Skalski, J.R., 1993, Statistical Framework for Monitoring *Zostera marina* (Eelgrass) Area in Puget Sound, Appendix L in Berry et al., 2003, Puget Sound Submerged Vegetation Monitoring Project: 2000-2002 Monitoring Report, Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, Washington.

Appendix C. Testing for Difference Between Two Means

This is a brief description of a procedure for the problem of testing for difference between two means for SVMP use.

In the 2000-2002 SVMP Monitoring Report (Berry et al. 2003), two different procedures were used to address this problem at the site level. In each case, the problem is to test whether a parameter estimate from one year is statistically different from the previous year's estimate.

In the case of site-level eelgrass area, the relative change in area was calculated with an associated confidence interval. The test for significant difference consisted of testing whether the confidence interval included the value zero. The calculations are described in the statistical framework (Skalski 2003, pp.25-6) and use a *Z* value for the calculation of confidence interval.

In the case of site-level depth estimates (max, min or mean), a mean and confidence interval (based on *Z* value) for the parameter of interest were calculated for the two years to be compared. The test consisted of determining whether there was overlap in the two confidence intervals. This approach was derived internally.

The approach described below is taken from basic statistical texts and is preferred to the test of overlapping confidence intervals used for depth estimates. In fact this test is "not generally valid" (Zar 1999, p.104). This approach could also be considered as an alternative to the testing for change in site-level eelgrass area.

C.1 Variances – Equal or Unequal?

There are two approaches to comparing two means depending on whether the two associated variances can be considered equal (i.e., not statistically different). These two approaches are referred to here as a *t* test and a *t'* test and are described below.

An F-test can be used to formally test whether the variances can be considered to come from different populations (Milton and Arnold 1990, pp.310-315; Zar 1999, pp.137-139). If the variances can be considered to come from the same population, then they can be used to calculate a pooled variance and a two-sample *t*-test is applied (Milton and Arnold 1990, pp.315-319; Zar 1999, pp.122-127). This case is denoted by Zar simply as a *t* test.

If an assumption of equal variances is not appropriate, then the variances cannot be pooled and a different approach is needed. In this case a modified *t*-test is used that relies on the Smith-Satterthwaite estimate of degrees of freedom (Milton and Arnold 1990, p.320; attributed solely to Smith by Zar 1999, p.128). This case is denoted by Zar as a *t'* test.

Milton and Arnold (1990) note that the *t'* test "performs well when variances are unequal but it yields results that are virtually equivalent to those obtained with the [*t*] test when variances are equal" (pp.321-322). They present this as a matter of opinion, but suggest that the *t'* test is preferred regardless of the equality of the variances thereby avoiding an F-test on the variances (p.322).

Similarly, Zar (1999) states "the routine test of variances is not recommended" (p.129). He notes that if the two sample sizes are different and variances are different then *t'* will provide a better (i.e., more powerful) test than *t*. If the variances are very similar then *t* is the better test. If the sample sizes are equal and variances are equal then the two tests are equivalent.

The recommended method described below is what Zar refers to as the *t'* test.

C.2 Description of Test

This test is for comparing two means from normal populations without assuming equal variances (Milton and Arnold, 1990, pp.320-322; Zar, 1999, pp.128-129). Zar states that numerous studies have shown this test to be robust enough to stand considerable departures from normality.

The test statistic follows a *t* distribution and is given by

$$t' = \frac{\overline{X_1} - \overline{X_2}}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

Eqn C.1

where

$\overline{X_i}$ = mean value for year *i*

s_i = sample standard deviation for year *i*

n_i = sample size for year *i*

The degrees of freedom for the *t* distribution is estimated using the following equation referred to as the Smith-Satterthwaite procedure by Milton and Arnold (1990, p.320 and denoted by γ) but attributed to Smith by Zar (1999, p.129 and denoted by ν').

$$\nu' = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}\right)^2}{\frac{\left(\frac{s_1^2}{n_1}\right)^2}{(n_1 - 1)} + \frac{\left(\frac{s_2^2}{n_2}\right)^2}{(n_2 - 1)}}$$

Eqn C.2

The value of ν' is rounded down to the nearest integer for the estimate of degrees of freedom.

Using the notation of Zar for a two-tailed test, the null hypothesis of no difference in means will be rejected if

$$\left|t'\right| \geq t_{\alpha(2),\nu'}$$

where the critical value $t_{\alpha(2),\nu'}$ is the two-tailed *t* value for significance level α and ν' degrees of freedom.

It is important to note that s_i in Eqn C.1 and Eqn C.2 above refers to **sample standard deviation**, not the **standard error** (i.e., the standard deviation of the distribution of means). If the standard error is available and denoted by $s_{e,i}$, then Eqn C.1 becomes

$$t' = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{s_{e,1}^2 + s_{e,2}^2}}$$

Eqn C.3

and Eqn C.2 becomes

$$\nu' = \frac{\left(s_{e,1}^2 + s_{e,2}^2\right)^2}{\frac{\left(s_{e,1}^2\right)^2}{(n_1 - 1)} + \frac{\left(s_{e,2}^2\right)^2}{(n_2 - 1)}}$$

Eqn C.4

C.3 Power and Sample Size

Zar (1999, pp.132-134) presents an estimate for minimum sample size needed to detect a minimum difference between the two means with specified significance level and power. This is an iterative procedure requiring an initial guess of minimum sample size. It is better to have an initial guess that is too high rather than too low.

The minimum sample size is calculated as

$$n \geq \frac{2s^2}{\delta^2}\left(t_{\alpha(2),\nu} + t_{\beta(1),\nu}\right)^2$$

Eqn C.5

where

δ = smallest difference to be detected
s^2 = sample variance (pooled variance if appropriate)
$t_{\alpha(2),\nu}$ = two-tailed *t* value for significance level α and degrees of freedom ν (a one-tailed test – denoted $t_{\alpha(1),\nu}$ – could be used if the purpose is to test for a particular type of difference)
β = probability of Type II Error (incorrectly accepting false null hypothesis); power of test is (1-β).
$t_{\beta(1),\nu}$ = one-tailed *t* value for significance level β and degrees of freedom ν.

Example (taken from Zar 1999, Example 8.4, p.134)
The procedure is to first make an initial guess of minimum sample size, i.e. sample size for each of the two populations to be compared. As an example, let us guess that *n* = 100

measurements of max. depth is the minimum sample size needed to meet our requirements.

The degrees of freedom to be used in Eqn C.5 is given by
$\nu = (n-1) + (n-1)$, or $\nu = 2(n-1)$ and for this example we have $\nu = 198$.

If we want a significance level of $\alpha = 0.05$ and a power of $1-\beta = 0.8$, then the two t values are

$$t_{\alpha(2),\nu} = t_{0.05(2),198} = 1.9720$$
$$t_{\beta(1),\nu} = t_{0.2(1),198} = 0.8434$$

Let us assume that we can use existing data to estimate variance as 9.6 m^2 (standard deviation = 3.10; standard error = 0.80 with a sample size of 15). If we want to be able to detect a change in maximum depth of 2.0 meter, then first iteration of n gives

$$n \geq \frac{2(9.6)}{(2.0)^2}\left(1.9720 + 0.8434\right)^2 \quad = \quad 38.05$$

For the second iteration we have $\nu = 2(38 - 1) = 74$ and

$$t_{\alpha(2),\nu} = t_{0.05(2),74} = 1.9925$$
$$t_{\beta(1),\nu} = t_{0.2(1),74} = 0.8465$$

and

$$n \geq \frac{2(9.6)}{(2.0)^2}\left(1.9925 + 0.8465\right)^2 \quad = \quad 38.69$$

The final result is that each sample should contain at least 39 measurements of maximum depth to meet our requirements.

C.4 Application to SVMP Depth Sampling

The SVMP analysis of *Z. marina* depth parameters (mean minimum and mean maximum) makes extensive use of testing for difference between means. This is one of the main tests for site level change between years.

Also, calculations of minimum sample size required to detect a target depth difference are used to guide field sampling on a site-by-site basis. A summary of these calculations was presented in Norris et al. (2001, p.24). As part of the effort reported here, a spreadsheet was developed to perform these calculations so that they could be adapted to specific site scenarios.

The spreadsheet calculations were first tested by replicating the above example (section C.3) and then by an attempt to replicate the results of Norris et al. (2000, p.24). There were significant discrepancies with the Norris et al results that have not been resolved.

The results obtained here for minimum sample size in depth sampling are presented in Figure- C-1.

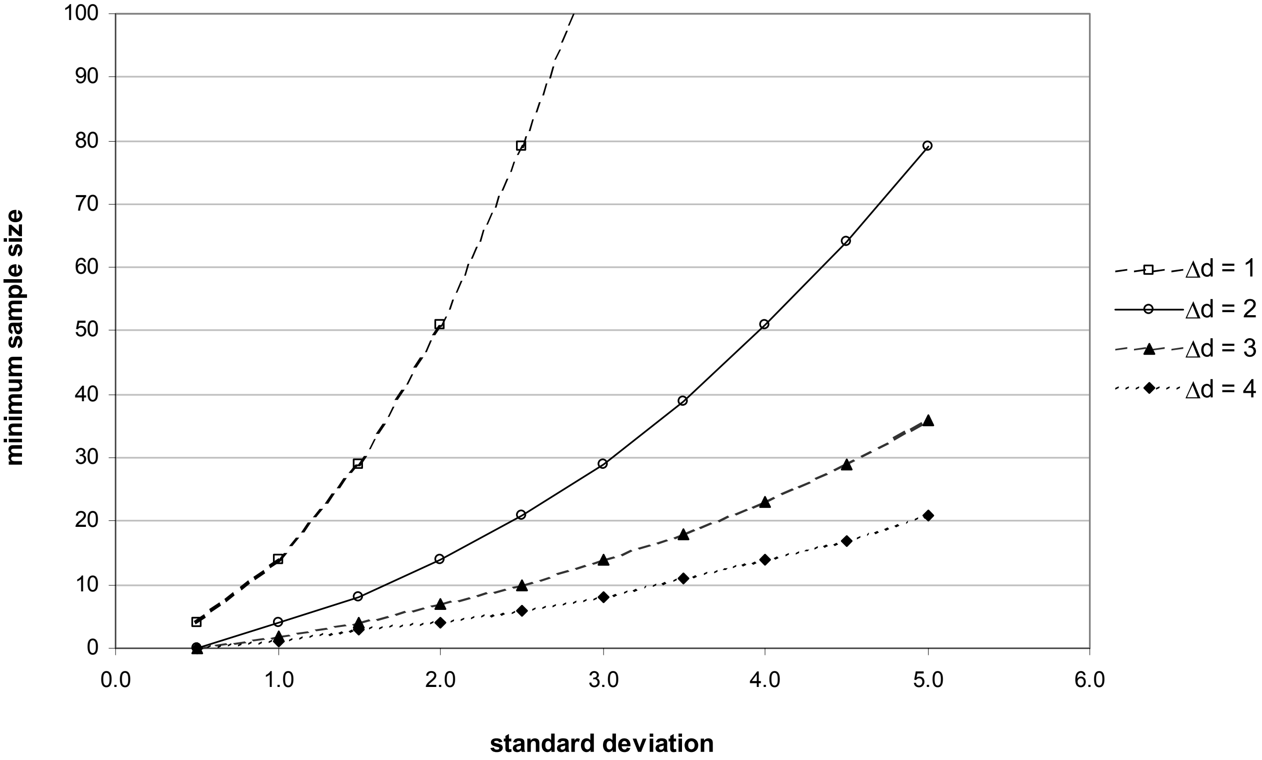


Figure- C-1. Minimum sample size necessary to detect differences of Δd=1, 2, 3 and 4 for a range of standard deviation with significance of α=0.10 and power of 1-β=0.80. Δd and standard deviation are in arbitrary but identical units.

C.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 State Department of Natural Resources, Olympia, Washington, 60pp. plus appendices.

Milton, J.S. and J.C. Arnold, 1990, Introduction to Probability and Statistics, McGraw-Hill.

Norris, J.G., S. Wyllie-Echeverria, J.R. Skalski and R.C. Zimmerman, 2001, Washington State Department of Natural Resources Submerged Vegetation Monitoring Project 2000 Final Report, Revised Draft September 26, 2001, Report from Marine Research Consultants submitted to Washington Dept. of Natural Resources.

Skalski, J.R., 1993, Statistical Framework for Monitoring *Zostera marina* (Eelgrass) Area in Puget Sound, Appendix L in Berry et al., 2003, Puget Sound Submerged Vegetation Monitoring Project: 2000-2002 Monitoring Report, Nearshore Habitat Program, Washington State Department of Natural Resources, Olympia, Washington.

Zar, J.H., 1999, Biostatistical Analysis, 4th edition, Prentice-Hall.

Appendix D. Derivations Used in Section 4.3.2

The objective is to derive equations for the ratio of adjusted to initial variance for both status and change estimates.

In the case of the status estimate this is denoted g_h and is given by

$$g_h = \frac{V\left(\bar{y}_h'\right)}{S^2/n}$$

where the following notation conventions are used:
h = sampling occasion, e.g. 0, 1, 2, 3 ….
n = number of samples (sites).
S^2/n = initial variance of the mean (assumed to apply to all occasions). Equivalent to $V\left(\bar{y}_h\right)$ for all occasions *h*.
$\bar{y}_h$ = initial mean (*Z. marina* area) for occasion *h*.
$\bar{y}_h'$ = adjusted mean for occasion *h*.
$V\left(\bar{y}_h'\right)$ = adjusted variance for occasion *h*.

Starting from Eqn. 12.80 of Cochran (1977, p.350) we have

$$\frac{n}{g_h} = u + \frac{1}{\dfrac{\left(1-\rho\right)^2}{m} + \dfrac{\rho^2 g_{h-1}}{n}}$$

where *n* is the total number of sites, *m* is the number of matched sites and ρ is the correlation in matched sites. The occasion subscripts on *n* and *m* used by Cochran have been dropped since these are fixed in the SVMP design.

Dividing by *n* gives

$$\frac{n}{g_h} = \frac{u}{n} + \frac{1}{\left(\dfrac{n}{m}\right)\left(1-\rho^2\right) + \rho^2 g_{h-1}}$$

Eqn D.1

which can be used to calculate g_h for various scenarios of ρ and rate of matching *m*/*n*.

In the case of the change estimate, the ratio of adjusted to initial variance is denoted g_h' and is given by

$$g_h^{'} = \frac{V\left(\overline{y}_h' - \overline{y}_{h-1}'\right)}{V\left(\overline{y}_h - \overline{y}_{h-1}\right)}$$

Eqn D.2

First we derive a relationship for the weight, ϕ_h, that is used to combine matched ($\overline{y}_{hm}'$) and unmatched ($\overline{y}_{hu}'$) estimates into an overall estimate for occasion h (see equation 12.79, Cochran 1977 and equations 14 and 23 of Skalski 2003 for the equation to calculate the overall estimate). The weight is calculated using (Cochran 1977, p.350; Skalski 2003, equation 15)

$$\phi_h = \frac{W_{hu}}{W_{hu} + W_{hm}}$$

Eqn D.3

where the W_i terms are the reciprocal of matched (m) and unmatched (u) variance on occasion h and are given by (Cochran 1977, table 12.3)

$$\frac{1}{W_{hu}} = \frac{S^2}{u}$$

$$\frac{1}{W_{hm}} = \frac{S^2\left(1-\rho^2\right)}{m} + \rho^2 V\left(\overline{y}_{h-1}'\right)$$

Eqn D.4

Note also that

$$g_{h-1} = \frac{V\left(\overline{y}_{h-1}'\right)}{S^2 / n}$$

Eqn D.5

Eqn D.3, Eqn D.4 and Eqn D.5can be combined resulting in

$$\phi_h = \frac{\dfrac{u}{S^2}}{\dfrac{u}{S^2} + \dfrac{1}{\dfrac{S^2\left(1-\rho^2\right)}{m} + \rho^2\left(\dfrac{S^2}{n}\right)g_{h-1}}}$$

which can be simplified to

$$\phi_h = \frac{\frac{u}{n}}{\frac{u}{n} + \frac{1}{\left(\frac{n}{m}\right)\left(1-\rho^2\right) + \rho^2 g_{h-1}}}$$

Eqn D.6

The variance of the adjusted change estimate is given by (Cochran 1977, equation 12.87, p.352) as

$$V\left(\bar{y}_h' - \bar{y}_{h-1}'\right) = V\left(\bar{y}_h'\right) + V\left(\bar{y}_{h-1}'\right) - 2\operatorname{cov}\left(\bar{y}_h' \bar{y}_{h-1}'\right)$$

Eqn D.7

The variance of the initial change estimate is the same as Eqn D.7 except adjusted estimates are replaced by initial estimates. Cochran (1977, equation 12.88, p.353) shows that Eqn D.7 can be manipulated to obtain

$$V\left(\bar{y}_h' - \bar{y}_{h-1}'\right) = \frac{S^2}{n}\left[g_h + g_{h-1}\left(1 - 2\rho\left(1-\phi\right)\right)\right]$$

Eqn D.8

In the case of the initial change estimate, the analogous version of Eqn D.7 can be manipulated if we neglect the covariance of $\bar{y}_h$ and $\bar{y}_{h-1}$ to obtain

$$V\left(\bar{y}_h - \bar{y}_{h-1}\right) = 2\frac{S^2}{n}$$

Eqn D.9

Finally, Eqn D.2, Eqn D.6, Eqn D.8 and Eqn D.9 can be combined to obtain the ratio of adjusted to initial variance of the change estimate given by

$$g_h' = \frac{1}{2}\left[g_h + g_{h-1}\left(1 - 2\rho\left(1 - \frac{u/n}{u/n + \frac{1}{\left(n/m\right)\left(1-\rho^2\right) + \rho^2 g_{h-1}}}\right)\right)\right]$$

Eqn D.10

Appendix E. Total Variance Equation

The variance of SVMP sound-wide estimates includes both sampling error and measurement error. Sampling error arises from using a random sample of sites within a stratum to calculate the value of the associated sound-wide estimator. Measurement error arises from the line transect sampling at the site level (i.e. within a sampling unit) and represents the error in measuring site-level *Z. marina* area.

The sound-wide *Z. marina* area estimator for a single stratum has a conditional probability distribution. The overall variance of this estimator can be broken down into components using the total variance formula.

For a given stratum, we define the random variable

$\hat{B}$ = estimate of *Z. marina* area in the stratum

and a quantity

x_i = true *Z. marina* area at site *i* in the stratum. This is used loosely in the equation below to represent the true *Z. marina* areas of a sample of sites taken in the stratum, rather than that of a single site.

The overall variance of this estimator is given by the total variance equation (following Skalski and Robson 1992, App. A)

$$Var\left(\hat{B}\right) \;=\; Var_{x_i}\left[E\left(\hat{B}|x_i\right)\right] + E_{x_i}\left[Var\left(\hat{B}|x_i\right)\right].$$

The two terms in this equation can be interpreted in the context of the SVMP as follows, starting from the innermost terms in the brackets and then moving outward:

$Var_{x_i}\left[E\left(\hat{B}|x_i\right)\right]$

$\hat{B}|x_i$ = the estimator of stratum *Z. marina* area given a random selection of sites each with a true *Z. marina* area x_i. There is a population of estimator values associated with repeated measurement of the given selection of sites.

$E\left(\hat{B}|x_i\right)$ = the expected value of the estimator given a random selection of sites each with a true *Z. marina* area x_i. There is a population of these expected values based on multiple selections of samples (sites) from the stratum.

$Var_{x_i}\left[E\left(\hat{B}|x_i\right)\right]$ = the variance of this population of expected values. This variance of expected values represents **sampling error**.

$E_{x_i}\left[Var\left(\hat{B}|x_i\right)\right]$

$\hat{B}|x_i$ = the estimator of stratum *Z. marina* area given a random selection of sites each with a true *Z. marina* area x_i. There is a population of estimator values associated with repeated measurement of the given selection of sites.

$Var\left(\hat{B}|x_i\right)$ = the variance of these estimator values given a random selection of sites each with a true *Z. marina* area x_i. Dispersion in this distribution of estimator values represents the role of measurement error for the given sample of sites. There is a population of these variances based on multiple selections of samples (sites) from the stratum.

$E_{x_i}\left[Var\left(\hat{B}|x_i\right)\right]$ = the expected value of this population of variances. This expected value of variance represents **measurement error**.