EVALUATION OF ADAPTIVE CLUSTER SAMPLING FOR REMOTELY-OPERATED UNDERWATER VEHICLE SURVEYS OF INSHORE ROCKFISH (*Sebastes* spp.)

by

Jessica Harris
B.Sc., University of Toronto, 2003

RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT
OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF RESOURCE MANAGEMENT

In the
School of Resource and Environmental Management

Project No. 456

SIMON FRASER UNIVERSITY

Fall 2008

APPROVAL

Name:	**Jessica Harris**
Degree:	**Master of Resource Management**
Title of Research Project:	**Evaluation of Adaptive Cluster Sampling for Remotely-operated Underwater Vehicle Surveys of Inshore Rockfish (*Sebastes* spp.)**
Project No.:	**456**

Examining Committee:

Chair: **Pier van Dischoeck**
Master Student, School of Resource and Environmental Management, Simon Fraser University

____

Dr. Sean P. Cox
Senior Supervisor
Assistant Professor, School of Resource and Environmental Management, Simon Fraser University

____

Dr. Carl J. Schwarz
Supervisor
Professor, Department of Statistics and Actuarial Science, Simon Fraser University

____

Dr. Andrew B. Cooper
Supervisor
Associate Professor, School of Resource and Environmental Management, Simon Fraser University

Date Defended/Approved: October 9, 2008

ABSTRACT

Remotely-operated vehicle surveys of inshore rockfish (*Sebastes* spp.) are low-cost and non-lethal, but frequently result in density estimates with high variance. Adaptive cluster sampling may result in more precise abundance estimates when the distribution of organisms is rare and clustered. Using computer simulations, I investigate the ability of adaptive cluster sampling to reduce sampling variance for remotely-operated vehicle surveys of rockfish. In general, adaptive cluster sampling was less precise than simple random sampling in surveys of equivalent sampling effort. Adaptive cluster sampling was only more precise (for equivalent sample size) for highly clustered rockfish populations, large initial sample sizes, and small numbers of adaptively sampled units, conditions which are not likely met for ROV surveys of rockfish, based on my analyses. Adaptive cluster sampling is not recommended for remotely operated vehicle surveys of rockfish and other strategies should be investigated to increase the precision of abundance estimates.

Keywords: adaptive cluster sampling; simple random sampling; rockfish; remotely operated vehicle; spatial distribution; underwater visual surveys

ACKNOWLEDGEMENTS

I thank my senior supervisor, Sean Cox for his expertise, guidance, and enthusiasm for tackling questions. I would also like to thank my committee members, Carl Schwarz and Andy Cooper, for their thoughtful input and reviews of this work. I am grateful to Lynne Yamanaka and Lisa Lacko at Fisheries and Oceans Canada, Robert Pacunski and Wayne Palsson at the Washington Department of Fish and Wildlife, and Cleo Brylinsky and Mike Byerly at the Alaska Department of Fish and Game for generously providing underwater visual survey data as well as helpful comments on this research project. I would like to acknowledge my colleagues in the Fisheries Science and Management Research Group at Simon Fraser University who provided valuable suggestions and support throughout the course of this research as well as the rest of the Resource and Environmental Management community who continue to create a stimulating academic environment and are inspiring in their dedication to addressing resource management problems. Financial support was provided by Fisheries and Oceans Canada via a grant to Sean Cox and Randall Peterman.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

.

1 INTRODUCTION

Detecting biologically important changes in population abundance is crucial to fisheries management. Therefore, monitoring surveys must be designed with low enough sampling variance to detect changes in population size and respond with appropriate management action (Peterman 1990). An important determinant of the variance of a population estimate is the underlying spatial distribution of organisms within the sampling area. For example, if spatial distribution is patchy, sampling variance will be large and the power of a survey to detect changes in population numbers will be low if distribution is not taken into account at the survey design stage. For these patchy distributions, an adaptive cluster sampling approach may be one means of reducing the sampling variance. Adaptive sampling can be more efficient than conventional sampling for rare and clustered distributions because sampling effort is increased when organisms are encountered. In populations with low densities and aggregated distributions, many sampling units will have no organisms and conventional sampling results in a high proportion of zero observations, which leads to population estimates with large variances (Seber and Thompson 1994).

The application of adaptive cluster sampling to remotely operated vehicle surveys of inshore rockfish (*Sebastes* spp) in the northeast Pacific could result in more precise density estimates for these rockfish populations and, therefore, better information for fisheries managers. Although abundance levels of rockfish are currently uncertain, what data are available suggest that rockfish populations may be threatened. For example, copper (*S. caurinus*), quillback (*S. maliger*), and brown (*S. auriculatus*) rockfish populations are at very low levels in Puget Sound, Washington and stock assessments completed in 1999 to 2001 indicate the biomass of at least seven of the major commercial rockfish species in California, Oregon, and Washington are possibly only 25% of 1970 levels (Love et al. 2002). Rockfish are often taken as incidental catch in fisheries directed at other species, such as those for Pacific halibut (*Hippoglossus stenolepis*) and sablefish (*Anoplopoma fimbria*) (Love et al. 2002). In British Columbia, vessels must possess quota to account for their directed and incidental rockfish catch, and vessels landing rockfish in excess of their holdings will be restricted from further fishing until additional quota is acquired (Fisheries and Oceans Canada 2008). Thus, detecting changes in rockfish density is necessary to avoid conservation losses from overharvesting rockfish or economic losses from underutilizing other fisheries.

Unfortunately, most conventional methods of estimating fish biomass are inadequate for inshore rockfish (O'Connell and Carlile 1993) because they lack accuracy or incur unacceptable rates of mortality, or are not feasible on rocky reef habitats. Underwater visual surveys allow the collection of population data without incurring incidental mortality; however, the precision of alternative sampling designs for underwater visual surveys has not been evaluated for rockfish.

Small ROVs are considered a relatively low-cost option for conducting underwater visual surveys and are currently employed by a number of agencies and research institutions to survey fish populations (Pacunski et al. 2008), including rockfish (e.g. Pacunski et al. 2008; Johnson et al. 2003; Fox et al. 2004; Karpov et al. 2006). However, despite increasing use of ROVs, population estimates from these surveys typically have high variance and thus low power to detect changes in rockfish density. For example, rockfish density estimates from an ROV survey at Perpetua Reef, an unfished nearshore rocky reef in Oregon, had relative standard errors ranging from 23 to 200 % (Fox et al. 2004). Although there are many potential causes for high sampling variance, the underlying distribution of individual fish may be a contributing factor. Qualitative observations of individual fish suggests that many species of inshore or nearshore rockfish are clustered at scales of less than 100 metres (Richards 1986, 1987; Love et al. 2002; Fox et al. 2004; Krieger 1993; Murie et al. 1994). In addition, quantitative data from trawl surveys show that the deep water rockfish, pacific ocean perch (*S. alutus*), are aggregated and patchy across the larger scale of several kilometres (Lunsford et al. 2001; Hanselman et al. 2001).

If the high sample variance of density estimates obtained from ROV surveys is caused by the clustered spatial distribution of rockfish, then using an adaptive cluster sampling, rather than a random or stratified random approach, could be a cost-effective means of increasing the precision of density estimate. Adaptive sampling has been found to be a more efficient sampling method than random sampling in other aggregated populations. For example, compared to random sampling, adaptive sampling increased the efficiency of trawl surveys for some species of Gulf of Alaska rockfish (Hanselman et al. 2001, 2003). Adaptive sampling was also found to be more efficient than random sampling for highly aggregated populations simulated using the Poisson cluster process (Su and Quinn 2003). Efficiency is often defined in relative terms; for example, the efficiency of one test relative to another is the ratio of sample sizes for the two tests when they have equivalent variances (Devore 2000).

Adaptive cluster sampling (ACS) begins by selecting an initial sample of survey units, which may be quadrats or transects. When the observed value of the variable of interest (e.g. number of fish) within a unit satisfies some condition for resampling *C*, its neighbours are added to the sample and a second round of sampling is performed on the neighbours. The neighbourhood of a unit can include the units on two, four, or eight sides of the initial unit as long as the neighbourhood is symmetric; if unit A is in the neighbourhood of unit B, then unit B must be in the neighbourhood of unit A (Thompson 2004). If the neighbours satisfy *C*, their neighbours are also surveyed, making a third round. Each initial unit and its subsequent rounds of neighbours are called a cluster and all units within a cluster that satisfy *C* are called a network. Units that do not satisfy *C* and are adjacent to a network are known as edge units (Thompson 1990). The adaptive sampling process continues until no new units satisfy *C* or the total number of sampling units or number of sampling rounds reaches a predefined restriction (Thompson and Seber 1996) (Figure 1).

Since Thompson first provided the theory and method for ACS, many papers have addressed its performance relative to simple random sampling (SRS) (Turk and Borkowski 2005). A primary area of interest has been on the population characteristics under which ACS outperforms conventional sampling. The efficiency of ACS relative to SRS depends on population characteristics and sample design (Thompson and Seber 1996). ACS tends to be more efficient than SRS for the same final sample size when (i) the population is aggregated as measured by the ratio of the variance of values in networks to the variance of values over the entire population of survey units (*rSSQ*); (ii) the population is rare or there are a small number of units with organisms relative to the total number of units; and (iii) the expected final sample size is not much larger than the initial sample size (Thompson and Seber 1996). Su and Quinn (2003) found that although in some cases ACS was more efficient than SRS for populations that were more aggregated (using the *rSSQ* measure), the relationship was not straightforward and depended on other parameter values as well. For example, allowing an increased number of initial samples reduced efficiency at low levels of aggregation and increased efficiency at high levels of aggregation (Su and Quinn 2003). Thus, descriptions of spatial structure that are more detailed than the *rSSQ* may be necessary to determine the spatial distribution of the population. Clear guidelines on when the population is sufficiently aggregated (and how to measure aggregation) to warrant ACS do not exist (Turk and Borkowski 2005).

Even if a population exhibits sufficient aggregation to make adaptive sampling more precise than conventional sampling, this efficiency may come at a cost of increased logistical difficulties because the final sample size is unknown *a priori* when using the standard ACS approach. If C is set too low, and each sampling unit meets the condition for resampling, the final sample size could be very large (Su and Quinn 2003). Even if C is set at the appropriate level, limits on fuel capacity, time to travel to the study area, personnel availability, or budget may make the uncertainty in time and cost required to complete an adaptive sample unacceptable to those planning surveys (Brown and Manly 1998). A number of studies have proposed ways to reduce the uncertainty in final sample size, but all involve compromises. Thompson (1994) suggests that if sample size becomes too large during the course of sampling, adaptive sampling could be halted and only the primary units would then be sampled in the remaining study area. The data could then be analyzed as if there were two strata: one adaptively sampled and one conventionally sampled. This strategy may result in different sampling intensities across the study area which may be undesirable. Restricted adaptive cluster sampling, in which sampling continues only as long as the total number of distinct units sampled is less than a target sample size is an alternate method of restricting sample size (Brown and Manly 1998). Initial units are sampled only after adaptive sampling is completed for the preceding network. Because sampling cannot be terminated in the middle of a network for estimators to be valid, final sample sizes may still be higher than the target, particularly if clusters are large (Brown and Manly 1998). In addition, density estimates from restricted ACS typically have a positive bias because of the error in estimating the expected value of the initial sample size (Brown and Manly 1998). Lo et al. (1997) used a stopping rule to terminate adaptive sampling after *S* fixed rounds of adaptive sampling. Although the use of a stopping rule also produces biased estimates, Su and Quinn (2003) recommend its use over restricted ACS because it is more effective at limiting final sample size. Using a stopping rule also allows the researcher to effectively set a limit on the proportion of adaptively sampled units relative to initial randomly sampled units. For example, when the study area is divided into transects and $S = 3$, each initial, randomly sampled unit will be followed by a maximum of 6 adjacent transects in the adaptive phase. Thus, in the final sample, at least 14% of the units will be randomly sampled. In comparison, restricted ACS only limits the final sample size. Sampling the first randomly selected unit could result in so many rounds of adaptive sampling that the maximum number of units would be reached before the second randomly selected unit was even sampled. Therefore, coverage of the study area by transects may be more uneven when restricted sampling is used than when a stopping rule is used.

My research goal is to determine whether estimates of population density from ROV surveys for rockfish using ACS, with or without a stopping rule, have lower sampling variance than surveys of equivalent sampling effort using SRS. The feasibility of ACS for ROV surveys of rockfish is unknown primarily because the aggregation and other population characteristics under which ACS is more precise than SRS are unknown. Further, the small-scale aggregation properties of inshore rockfish are unknown. My analyses therefore target these knowledge gaps by quantitatively assessing the small-scale aggregation of individual rockfish locations collected from ROV and manned submersible surveys across the northeast Pacific. I fit two alternative point process models to the data. The first is a random model of distribution and the second is a clustered model of distribution. I use the models to simulate ROV surveys of study sites with mixed-species rockfish aggregations and simulate adaptive cluster sampling and random sampling on these study sites.

2 METHODS

I present my analysis and simulation methods for evaluating adaptive cluster sampling in four main sections. First, I provide a brief summary of the remotely operated underwater vehicle surveys that I used to develop an empirically based simulation model of rockfish spatial distribution. A more complete description of empirical data is presented in Appendix A. The second section describes the simulation approach I used to evaluate combinations of sampling designs and density estimator. This evaluation is done with respect to two possible models – clustered and random – of small-scale rockfish spatial distribution. Model parameterization, based on empirical data from underwater visual surveys, is given in Appendix B. Sections 2.3 and 2.4 describe the performance measures and sensitivity analyses, respectively, that I use to complete the evaluation of ROV surveys with adaptive cluster sampling against those employing simple random sampling protocols.

2.1 Survey data and spatial model

Visual survey data were from two line transect manned submersible surveys and two strip transect ROV surveys of inshore and nearshore rockfish conducted at seven sites in the northeast Pacific (Figure 2). Data were from inshore rockfish surveys in the San Juan Channel (SJC) in 2004 using a Deep Ocean Engineering (DOE) Phantom HD2+2 ROV (R.E. Pacunski, Washington Department of Fish and Wildlife, Mill Creek, WA, pers. comm.), inshore rockfish stock assessment surveys using the *Aquarius* submersible, in Juan Perez Sound in the Queen Charlotte Islands (QCI), British Columbia in 2005 (L. Yamanaka, Fisheries and Oceans Canada, Nanaimo, BC, pers. comm.), demersal shelf rockfish stock assessment survey using the *Delta* submersible in the Southeast Outside (SEO) subdistrict in the eastern Gulf of Alaska (C. Brylinsky, Alaska Department of Fish and Game, Sitka, AK, pers. comm.), and lingcod and rockfish surveys on Chiswell Ridge (CHR) in the northern Gulf of Alaska along the Kenai Peninsula in 2005 using the DOE Phantom HD 2+2 ROV (M. Byerly, Alaska Department of Fish and Game, Homer, AK, pers. comm.; Table 1). All surveys involved some form of stratification to select for rocky habitat (Appendix A).

2.2 Simulated sampling and estimation of rockfish density

A simulation approach to evaluating alternative sampling designs requires a stochastic model for simulating “true” spatial distributions of fish. In this paper, I used empirical survey observations of individual rockfish locations to parameterize two distinct types of spatial processes (Appendix B; Table 2). In the first - the Thomas process or "clustered model" - invisible cluster centres are randomly distributed over the survey area following an isotropic bivariate normal distribution with variance ρ^2 (m^2) (where ρ is approximately equal to the cluster radius), mean number of fish per cluster μ, and cluster intensity λ (m^{-2}). For the second - the Poisson spatial process or "random model" - an intensity of γ (m^{-2}) fish are uniformly distributed at random throughout the survey area.

I simulated rockfish spatial distributions, ROV surveys, and density estimation in 500 survey sites where each covered a 1 km^2 area. I divided each site into 500 non-overlapping, parallel ROV transects that were 1 km in length and 2 m in width where the latter is approximately equal to the average field of view in ROV surveys (Appendix A). Total rockfish abundance in each site was estimated by applying one of five sampling-estimator protocols, which are defined by the following combinations of transect sampling method and density estimator, respectively: (1) simple random sampling with classical density estimator (SRS); (2) adaptive cluster sampling with Horvitz-Thompson density estimator (HT); (3) adaptive cluster sampling with Hansen-Hurwitz density estimator (HH); (4) ACS with a stopping rule and Horvitz-Thompson density estimator (HT-S); and (5) ACS with a stopping rule and Hansen-Hurwitz density estimator (HH-S).

In describing the various density estimators, I follow typical notation as used in the sampling literature (e.g., Thompson 1990 and 2002). Because all transects are identical in dimensions, I use the terms population numbers and population density interchangeably. The observed sample mean number of rockfish per transect $\bar{y}$ is the average of simulated rockfish counts y in each ROV transect. An estimate of the total site abundance can be obtained by multiplying the sample mean by the total number of potential transects N = 500 in each site.

Simple random sampling

In the simple random sampling (SRS) protocol, transects were randomly selected without replacement and the site mean density was estimated using

$$\hat{\mu}_{SRS} = \frac{1}{m}\sum_{i=1}^{m} y_i \; , \qquad (1)$$

where y_i is the observed density in the i^{th} transect and m is the total number of transects sampled (Thompson 2002). The number of transects sampled for SRS was set equal to the number of unique transects sampled in the ACS protocol (see below) on the same study site. The total site abundance was estimated by multiplying the number per transect by the total number of possible transects in the site N, i.e.,

$$\hat{\tau}_{SRS} = N\hat{\mu}_{SRS} \qquad (2)$$

which has the variance

$$\hat{V}(\hat{\tau}) = N(N-m)\frac{s^2}{m} \; , \qquad (3)$$

where

$$s^2 = \frac{1}{m-1}\sum_{i=1}^{m}(y_i - \hat{\mu}_{SRS})^2 \; .$$

Adaptive cluster sampling with Horvitz-Thompson estimator

For adaptive cluster sampling, n_1 initial transects were selected randomly without replacement. If none of these initial transects contained fish, randomly sampled transects were added until at least one fish was observed. If at least a threshold of C fish were observed in the initial transect, one transect on each side of the initial transect was also sampled. If the same condition for resampling was satisfied in either of these transects, their neighbours were also sampled (if they had not already been). This procedure continued until no additional transects satisfied the condition for resampling (Figure 1). Each initial transect and its neighbours that met the condition for resampling are defined as a network. Transects that did not satisfy the condition for resampling were not considered part of a network unless they were initial transects, in which case they were a network of one transect with zero observations. Thus, the adaptive cluster sample with Horvitz-Thompson estimator (*HT*) computes the sample mean density as

$$\hat{\mu}_{HT} = \frac{1}{N} \sum_{k=1}^{K} \frac{y_k^*}{\alpha_k}, \tag{4}$$

where y_k^* is the total number of fish in the k^{th} network, K is the number of distinct networks in the sample, and α_k is the probability that network k is included in the sample, which is given by:

$$\alpha_k = 1 - \binom{N - N_k}{n_1} \bigg/ \binom{N}{n_1} \tag{5}$$

where N_k is the number of transects in the k^{th} network. The site total abundance was estimated by $N\hat{\mu}_{HT}$. An unbiased estimator of the site variance is

$$\hat{V}(\hat{\tau}_{HT}) = \sum_{j=1`}^{K} \sum_{k=1}^{K} \frac{y_j^* y_k^*}{\alpha_{jk}} \left(\frac{\alpha_{jk}}{\alpha_j \alpha_k} - 1 \right), \tag{6}$$

where α_{jk} is the probability that both networks *j* and *k* are included in the sample, which is given by:

$$\alpha_{jk} = 1 - \left[\binom{N - N_j}{n_1} + \binom{N - N_k}{n_1} - \binom{N - N_j - N_k}{n_1} \right] \bigg/ \binom{N}{n_1} \quad (7)$$

Adaptive cluster sampling with Hansen-Hurwitz estimator

Adaptive cluster sampling with Hansen-Hurwitz estimator (HH), proceeded as described above, but the unbiased estimator of the population mean is defined by (Thompson 1990)

$$\hat{\mu}_{HH} = \frac{1}{n_1} \sum_{i=1}^{n_1} w_i \,, \quad (8)$$

where n_1 is the number of initial transects (and also the number of networks) and w_i is the mean number of fish per transect in the network that contains transect *i*, i.e.,

$$w_i = \frac{y_i^*}{x_i} \,, \quad (9)$$

where y_i^* is the total number of fish observed in x_i transects in the network that contains transect *i*. The site total abundance was estimated by $N\hat{\mu}_{HH}$ and an unbiased estimator of the site variance is

$$\hat{V}(\hat{\tau}_{HH}) = N^2 \frac{N - n_1}{Nn_1(n_1 - 1)} \sum_{i=1}^{n_1} (w_i - \hat{\mu}_{HH})^2 \,. \quad (10)$$

For adaptive cluster sampling with a stopping rule, sampling proceeded as in the other ACS protocols; however, adaptive sampling concluded after *S* rounds. The Horvitz-Thompson (HT-S) or Hansen-Hurwitz (HH-S) estimators were used to estimate site mean and variance as described above.

In the baseline simulation scenario, the number of initial transects was n_1 = 10 and the condition for sampling neighbouring transects in the protocols using ACS was observing at least C = 20 fish. When a stopping rule was used, sampling concluded after S = 3 rounds. Thus, a minimum of 14% of samples were initial randomly sampled transects versus adaptive transects. Analyses of sensitivity to alternative values of these design parameters are described in detail in Section 2.4.

2.3 Performance measures

The performance of alternative estimators can be characterised by measures of accuracy such as relative bias and relative standard error. Relative cost is also important for ROV surveys because of shiptime requirements and the need for highly trained vessel and ROV operators. Relative percentage bias is the amount by which the sampling protocol over or under-estimates the number of fish in the site as a percent of the true population size over repeated estimates. The HH, HT, and $\hat{\tau}_{SRS}$ estimators are unbiased (Thompson 1990); however, I calculate the bias of these methods to allow comparison with the bias of ACS using the stopping condition (ACS-S). For example, a bias of ACS of 5% and a bias of ACS-S of 7% would suggest that using a stopping condition does not introduce substantial bias, a finding that would not be evident if the bias of ACS was not calculated. I use the term relative percentage difference to refer to the difference between the estimated value and true value from a single estimate as a percentage of the true value, i.e.,

$$R\left(\hat{\tau}_r\right) = 100\frac{\hat{\tau}_r - \tau_r}{\tau_r}, \tag{11}$$

where $\hat{\tau}_r$ is the estimated number of fish in a site and τ_r is the true number of fish in the site for the r^{th} simulation. Relative standard error is a useful measure of precision when comparing estimates from sampling protocols with different population sizes:

$$RSE\left(\hat{\tau}_r\right) = 100\frac{\sqrt{V\left(\hat{\tau}_r\right)}}{\tau_r}. \tag{12}$$

Total cost was estimated for each sampling event by adding the number of initial transects n_1 to the number of adaptively sampled transects n_2 multiplied by the cost ratio c of adaptively sampled transects to initial transects, i.e.,

$$Q_r = n_{1r} + cn_{2r}, \tag{13}$$

where I initially assume that $c = 0.5$.

In addition to bias and precision, I also examined the relative variance of ACS to SRS for each simulation of the clustered rockfish model because the effect of spatial distribution on the performance of ACS relative to SRS could be masked by the considerable variation in rockfish distribution between simulations. These analyses were also used in sensitivity analyses investigating distributional properties that affect the variance of ACS. Relative variance (*RV*) was calculated as

$$RV = \frac{V(\hat{\tau}_{HH})}{V(\hat{\tau}_{SRS})}. \quad (14)$$

Relative variance values $RV > 1$ indicate that SRS has lower variance for equal numbers of sampled transects and is thus more "efficient" than ACS. I use the term relative variance rather than relative efficiency because the latter is based on variances with equal final numbers of transects, which excludes edge units (Su and Quinn 2003). In contrast, I compared ACS and SRS surveys that sample the same total number of units. Therefore, the sample size used in SRS estimates of mean density will almost always be larger than the sample size in ACS estimators because edge units are networks of size one with zero observations and are thus excluded from the calculation of the HH and HT density estimates. Therefore, RV does not coincide exactly with relative efficiency.

2.4 Sensitivity analyses

I examined the sensitivity of performance measures for adaptive sampling strategies to input parameters for adaptive cluster sampling as well as input parameters for the simulated spatial processes. For adaptive sampling protocols, I tested the following ranges of parameters (bold values indicate the baseline level): condition for resampling C = (1, 10, **20**, 50), initial transect number n_1 = (5, **10**, 20, 40), stopping condition S = (1, 2, **3**, 4, 5), cost ratio c = (0.25, **0.5**, 0.75, 1).

Spatial population aggregation, small network sizes or small m-n_1, and rarity are all factors that affect efficiency (Thompson and Seber 1996; Brown

2003). Therefore, I also conducted sensitivity trials for spatial process parameters that determined characteristics of the simulated populations. For the Thomas process, I varied (i) the cluster radius to examine a range of clustering, (ii) cluster intensity to represent rarity, and (iii) fish per cluster, which indirectly affects network size when combined with certain values of the resampling condition. I varied each of the three parameters between two levels and simulated 500 sampling-estimator protocols for each of the eight combinations (Figure 3). I used relative variance as the sensitivity indicator in these trials.

3 RESULTS

Simple random sampling resulted in population estimates that were more precise and closer to the true value than adaptive cluster sampling did in the majority of simulations regardless of the particular model used to simulate small-scale spatial distribution (Figure 4, Figure 5); however, the difference may not be practically significant because, for example, relative percentage differences for the two designs were typically within one standard error of each other (Table 3). When simulated ROV surveys were performed on clustered rockfish populations, SRS had the smallest relative percentage difference and also had the smallest standard error of relative difference (Table 3). Similarly, the mean and standard deviation of relative standard error was smallest for SRS (Table 3). When ROV surveys were performed on randomly distributed rockfish populations, relative percentage difference and relative standard error were smaller for all protocols compared to those under the clustered model (Figure 5). This difference is likely due, at least in part, to the larger sample sizes obtained in simulations using the random model of small-scale spatial distribution. For example, the average number of unique transects sampled for the SRS, HT, and HH protocols was 139 under the random model and 75 under the clustered model. In the random model, fish were spread more evenly over the study area resulting in the majority of transects meeting the ACS condition for resampling. Thus, the sample sizes tended to be large because many adaptively sampled units were added to the initial sample. The effect of these large sample sizes can also be seen in the large costs associated with ACS. Like the clustered model, SRS had a smaller mean relative difference and relative standard error, but greater costs than ACS under the random model.

Differences in cost between simple random sampling and ACS for surveys involving equal sample sizes depended on the cost ratio (*c*). The ACS surveys with stopping rule tended to cost less because total sample sizes were smaller in these surveys. The mean cost and standard error of cost were greatest for SRS, moderate for ACS (HT and HH), and least for ACS-S (HH-S and HT-S) (Table 3). Using a stopping condition with ACS reduced survey cost while not significantly decreasing the accuracy of population estimates compared to conventional ACS.

Differences in statistical performance among the four adaptive cluster sampling protocols were slight; however, cost of the four protocols did differ. Under both models, adaptive cluster sampling with a stopping condition for the HH and HT estimators (ACS-S) had smaller costs without significantly greater relative standard errors than ACS, or SRS. Relative percentage difference of ACS-S was only slightly larger than relative percentage difference of ACS and only for the HT-S estimator, which tended to have a positive relative percentage difference (the estimate was greater than the true value) relative to HT under both models. The bias effect was larger under the random model because (i) network sizes tend to be larger and (ii) repeat networks are excluded in the HT estimator, yet it was not always possible to identify a repeat network when a stopping rule was used. Sampling beyond the stopping condition is required to reveal that two networks are actually the same one. Such a failure to identify repeat units is not problematic for the HH-S estimator because the population total is obtained by averaging the number of units in each network. Both estimators resulted in roughly equivalent differences between estimated values and true values. Because of these similarities, I only present sensitivity analyses using the HH estimator for both ACS and ACS-S.

Adaptive sampling protocols were more precise than SRS only when the proportion of transects with zero observations was large and the difference between final and initial sample sizes was small (Figure 6b, c). No relationship was visible between relative variance and the ratio of within network sum of squares to total sum of squares (*rSSQ*) (Figure 6a). The strongest correlation appeared between relative variance and the difference between initial and final sample size; as change in sample size increased, relative variance increased. All populations where HH was more precise than SRS occurred at changes in sample sizes of less than approximately 50.

The Thomas process parameters, cluster radius (ρ) and cluster intensity (λ) also appeared to influence the precision of ACS relative to SRS (Figure 6d, e). Relative variance was less than one when cluster radius was 10 m or less and cluster intensity was 0.0001 clusters/m^2 or less. The mean values of cluster intensity and radius were also less in populations where $RV < 1$ than in all populations (Table 4). Small values of cluster radius resulted in small clusters and small network sizes and a small ratio of initial to final sample size. Small values of cluster intensity indicate that clusters will be rare and less likely to overlap and resulting in large network sizes. There was no apparent relationship between mean number of fish per cluster and *RV*, which suggests that total population size is less important than the size and density of clusters (Figure 6f). None of the factors examined wholly explained relative variance, however, and many populations had a large relative variance even when cluster radius and intensity were small.

In the sensitivity trials for the spatial process parameters, population estimates from simple random sampling had a lower variance than those from ACS for each of the eight combinations of the parameter values (Figure 3); however, cluster radius and intensity affected the degree to which SRS was more precise than ACS (Figure 7). Population estimates from surveys simulated on sites with small values of cluster radius had greater relative variance than those estimates from surveys on sites was a large value for cluster radius (Figure 7a,b,d). This finding is counter to the results above, where smaller cluster radii were associated with smaller relative variances, and likely occurs because sites simulated with the large value for cluster radius were densely packed with fish (Figure 3; left side). In these spatial distributions, both adaptive and random sampling surveys resulted in sampling most of the transects, and, therefore, relative variances close to one. Looking at populations where cluster radii were small, relative variances were smaller in population estimates from surveys simulated on populations with the small cluster intensity than the large cluster intensity (Figure 7). Fish per cluster had a minimal effect on relative variance as above (Figure 7).

Decreases in the condition for resampling (*C*) led to modest decreases in relative standard error but also large increases in cost and either no effect or an increase in relative percentage difference (Figure 8). As *C* decreased, more adaptively sampled units were added to ACS and the total sample size increased. This increased the sample size used for SRS as well. This increased sample size decreased the relative standard error and increased the cost but the effect on relative percentage difference was not consistent. Starting from $C = 1$, increases to $C = 10$ resulted in decreased relative percentage difference for SRS and ACS. This decrease is likely just the result of the mix of samples observed in each population and the resulting population estimates. Changing from $C = 10$ to $C = 20$ again led to a decrease in relative percentage difference for SRS but an increase for ACS-S and ACS. Moving from $C = 20$ to $C = 50$, relative percentage difference was largely unchanged. All performance measures for ACS-S and ACS were most similar at $C = 50$ because the condition for resampling is so large that it limits additional rounds of adaptive sampling rather than the stopping condition. At $C = 50$, the cost of SRS and the ACS methods is closer than at $C = 1$ because there are fewer adaptively sampled units, which are less costly than initial units. The relative standard errors are also similar because the proportion of units sampled that are identical in each method are greater.

Increasing the number of initial transects sampled (n_1) led to decreases in relative percentage difference and relative standard error for all protocols with the exception of a slight increase in relative percentage difference when the number of initial transects sampled increased from 10 to 20 (Figure 9). However, for ACS the more consistent increases in performance observed with increases in the number of initial transects sampled suggest that adding new randomly sampled units increases the accuracy of the population estimate more than decreasing the condition for resampling. This result was not due to a difference in total sample size because fewer units were actually sampled in total when $n_1 = 40$ (a mean of 136 unique samples) compared to when $C = 1$ (182 unique samples). Instead, increasing the initial number of transects lead to decreases in relative standard error because of the way samples were added. For example, when the number of initial transects sampled is increased, more new clusters are potentially added, whereas when the condition for resampling is decreased, more units within a cluster are added. Thus, relative percentage difference decreases more when the number of initial transects sampled is increased because a more accurate estimate of how many clusters are in a population has greater value than determining the number of individuals in a cluster. In addition, when clusters are large and rare, a small value of the condition for resampling and number of initial transects sampled will result in over-estimates of the population size because the one large cluster will dominate the calculations. Increasing the number of initial transects sampled counteracts this tendency. At high numbers of initial transects sampled, the relative advantage of SRS over ACS was decreased.

Changing the stopping rule had little impact on sampling-estimator performance measurements. Relative percentage difference and relative standard error decreased slightly and cost increased slightly as the number of rounds of adaptive sampling (S) increased (Figure 10). As the results of the sensitivity analysis on the condition for resampling showed, increasing the number of adaptive samples that are added does not result in significant improvements in relative standard error or relative difference performance measures. When these results were compared to SRS with sample size from ACS-S rather than ACS, SRS was still more accurate, and only marginally more costly.

Total survey costs increased linearly for both adaptive cluster sampling procedures as the ratio of cost-per-unit-sampled between simple random sampling and adaptive sampling increased. As expected, costs did not increase as much overall for ACS with stopping rule.

A useful way of thinking about the trade-offs between cost and variance is to consider which method gives the lowest relative standard error for a given cost. If n_1 = 20 and C = 20, c = 0.5, and S = 3, an ACS-S survey results in a mean relative standard error of 40.3%, a relative percentage difference of approximately 3% and a cost of 33.7 units. Spending 33.7 units on an SRS or selecting 34 random units still resulted in a relative standard error of 31.6% and a relative percentage difference of 1.9%. Therefore, although the results shown here indicated SRS was more costly than ACS when the sample sizes were equal, when surveys of equivalent cost were compared, SRS surveys almost always had lower variance than ACS surveys, despite the fact that adaptively sampled units were assumed to cost less. Even when the cost ratio is reduced to 0.25, an SRS survey of equivalent cost would allow 28 random samples and would have a relative standard error of 33.8%. The exception to this finding is when large numbers of initial transects were sampled. For instance, when n_1 = 40, an ACS-S survey has a relative standard error of 28.3% and a cost of 64; whereas, an SRS survey with 64 units has a relative standard error of 33.8%. The relative percentage difference of the SRS survey is still lower at 0.9% than the ACS-S survey at 3%.

The apparent superiority of SRS to ACS was robust to alternative values for the maximum separation distances used in the Thomas process parameter estimation procedure (Table 3). When h_f = 25, the difference in relative standard error and relative percentage difference between ACS and SRS was smaller. This likely occurred because the estimate of cluster radius was smaller at this value of h_f and as found above, there was a correlation between relative variance and cluster radius. The value of cluster intensity was also larger at h_f = 25, which would be expected to have the opposite effect on relative variance; however, the effect on cluster radius may have been more important. The opposite pattern was seen at h_f = 100; there was a greater difference between SRS and ACS, which may have been the result of large estimates of cluster radius.

4 DISCUSSION

Adaptive cluster sampling did not result in more precise population estimates than simple random sampling for most of the simulation scenarios I examined. This contrasts other studies that found ACS more efficient under a variety of situations including cases where data followed a clustered point process (Brown 2003; Su and Quinn 2003). Such inconsistency could arise from the underlying spatial models I used or the method I used to compare ACS and SRS.

4.1 The relative variance of ACS to SRS

Population spatial characteristics such as the amount of clustering and rarity affect the relative efficiency of alternative estimators and/or sampling protocols (Thompson and Seber 1996). For example, in other studies based on the Thomas process, ACS is typically more efficient than SRS in populations that are more clustered as measured by the ratio of within network variance to total variance (Su and Quinn 2003; Brown 2003). This suggests that SRS outperformed ACS in my study because the simulated rockfish populations were not sufficiently clustered. Although many of the Thomas process parameter values I used were outside the ranges used in other studies, a subset of my simulations had similar parameters to Brown (2003), who found that ACS outperformed SRS. However, even for these most clustered populations, ACS was more precise than SRS in only 50% of my simulations, which suggests that insufficient clustering is not the sole cause of the discrepancy between my results and others. ACS could have been less precise than SRS in those 50% because cluster radii were smaller than those used by Brown. The probability that an individual cluster is contained within a single sampling unit increases as cluster radius decreases, which means that ACS would frequently lead to unnecessary sampling of edge units containing no fish. Therefore, sampling with ACS only increases the sample size of the SRS estimator, without providing any additional information to the ACS estimator when cluster sizes are too small (Turk and Borkowski 2005).

The performance of ACS relative to SRS was only compared over the range of parameter values observed in the rockfish data I had available. ACS may perform better than SRS for different parameter values, such as when rockfish are much rarer than I observed. This could be the case in other less productive environments than those in CHR, SEO, QCI, or SJC.

Differences in the way SRS and ACS are compared in my study versus other studies may also play a role in the different results obtained for the relative performance of the sampling methods.

The majority of ACS studies compare the efficiency of ACS and SRS rather than the precision; that is, they use the ratio of variances of ACS to SRS when the SRS sample size is equal to the “final” sample size of the ACS survey, which excludes edge units (those units surveyed which did not contain fish; Thompson 1990). In contrast, I used an SRS sample size equal to the number of transects sampled in the ACS survey including edge units. Edge units are frequently excluded from the calculation of sample size for SRS because edge units are not included in calculating the HH and HT estimators and comparing the statistical efficiency of the estimators requires that the sample size in both estimators is equivalent (Su and Quinn 2003). I included edge units in my calculation of sample size because I am interested in evaluating the relative performance of ACS and SRS with equal effort and I do not have evidence that surveying empty units would require significantly less effort than surveying units with fish. Edge units could be more easily sampled than network units if enumerating organisms requires a substantial amount of time such as, when catch from a trawl survey tow must be processed or when empty units could be identified prior to the survey by using hydroacoustic equipment (Hanselman et al. 2003). If empty units can be sampled more quickly than units with organisms, then it would be reasonable to assume that empty units in SRS have negligible cost as well; however, none of the studies make this assumption. The practice of excluding edge units and not empty SRS units could be justified if empty units were identified through the process of surveying adjacent units, such as by using SONAR equipment on an ROV to scan neighbouring units while conducting a visual survey of an adjacent unit. Excluding edge units from the sample size calculation does not seem justified based on current ROV survey practices.

Including edge units in the sample size calculation means that a larger sample size can typically be obtained for an SRS survey than an ACS survey for equal sample effort. Variance of SRS surveys will therefore tend to be smaller than the variance of ACS surveys simply because sample size is larger. ACS surveys, therefore, have an inherent disadvantage compared to SRS surveys, when equal total sample sizes are used in comparing the two methods. This disadvantage likely explains at least part of the discrepancy in my results and those of Brown (2003) and Su and Quinn (2003). Other studies in which ACS was not more precise than SRS similarly included edge units in the sample size calculation and used equal total sample sizes (Smith et al. 2003; Lo et al. 1997).

An alternative to either comparing equal total sample sizes (including edge units) or equal "final" samples sizes (excluding edge units) is to explicitly account for the time used to complete an ACS survey and spend the same amount of time on an SRS survey. Hanselman et al. (2003) compared SRS and ACS studies of equivalent length of time as well as surveys in which the distance travelled was equivalent for both methods and found ACS to be more precise in both cases. These results suggest that ACS can be more efficient when adaptively sampled units require significantly less time to survey. If adaptive units can be sampled at a fraction of the cost of initial units, then more units can be surveyed in an ACS survey than in an SRS survey for equivalent cost, and the ACS survey may have lower variance. In their study, sampling adaptive units required a third of the time of initial units. I still found SRS to be more cost-effective even when I reduced the cost ratio to 0.25; however, if I conducted the sensitivity analysis on the cost ratio using only the most clustered distributions, I might have found that ACS was more cost-effective than SRS when the cost of adaptively sampled units was assumed to be small relative to randomly sampled units.

One final difference between my study and others was that I simulated transects, where most other studies simulated quadrats. When using transects, only the clustering on the *x*-axis is important. Some of the distributions examined in my analysis may not have been particularly clustered when the clustering along the y-axis is ignored (Figure 3 b,d,f,g). However, other distributions would still be quite clustered (Figure 3 a,c,e,h), considering that transects are 2 m in width, and would fall within the range of clustering observed in other studies. Further research would be needed to determine if study designs with transects, rather than quadrats, are less suited for ACS.

4.2 Factors affecting the precision of ACS

The sampling protocols in this study and other studies were similar and likely not responsible for the discrepancy in the relative efficiency of estimators under ACS; nevertheless, the sampling protocol affects the precision of estimators under ACS and should be considered in evaluating survey design. Sampling protocol includes the selection of sampling parameters such as the initial number of samples (n_1), the criteria for resampling (C), the use of a stopping condition (S), transect size, and the estimator used (HH or HT).

Increasing the number of initial samples increased the efficiency of ACS relative to SRS. Even when the population was not consistently clustered, ACS was more cost-effective than SRS at the highest number of initial samples. Using a large number of initial samples, however, means a large number of units will be sampled, which can be quite costly and is likely not feasible for many studies.

Changing the condition for resampling did not have a large effect on the precision of the ACS survey; however, the condition for resampling may affect variance in some circumstances by affecting network size. Adaptive cluster sampling is also only more efficient than SRS when the difference between the initial and final sample size is small, which occurs when networks are small. Network size depends on population characteristics, such as clustering, but also on sampling protocol. A large condition for resampling or a stopping rule can prevent networks from becoming too large. My results did not suggest that the bias introduced by a stopping condition was much larger than the average relative percentage difference of ACS surveys; however, significant differences in bias between ACS and ACS-S have been observed in studies of more clustered populations (Su and Quinn 2003). Limiting network size by using a large condition for resampling may be preferable to potentially increasing bias through the use of a stopping condition; however selecting a condition for resampling which allows the optimal amount of adaptive sampling to occur may be difficult without extensive prior information about the population. Using a stopping rule could be a reasonable option in this situation.

Transect size was not examined here; however, the size of sampling units affects network sizes, and therefore precision (Philippi 2005). Larger units will contain more individuals and have smaller network sizes compared to smaller units. On average, large units will also capture more of the population compared to an equal number of small units, thus leading to more precise population estimates, in general. However, if unit sizes are too large, entire clusters will be contained within a unit making ACS less efficient than SRS (Philippi 2005). In ROV surveys, unit width is dependent on the cameras used and is not an easily changed survey parameter. The effect of unit length could be investigated; however, this too may be constrained by physical characteristics of the study site, such as walls. Using shorter transects than necessary may also be inefficient because it would require increased manoeuvring of the ROV in between transects.

The Horvitz-Thompson estimator is typically considered to be more efficient than Hansen-Hurwitz estimator (Turk and Borkowski 2005); however, I did not observe a significant difference between the estimators. Christman (1997) investigated the difference in performance between HH and HT and found that the differences between HH and HT increased as the number of initial samples increased (n_1) and the largest differences occurred when the criteria for resampling (C) was small and the number of clusters was small. I may not have observed a difference between HH and HT because, I did not covary the number of initial samples and the condition for resampling, and the number of clusters in the region ranged widely. The percentage of initial samples out of total samples (n_1/N) was also much larger in Christman (1997). Given that, in most field surveys, n_1 is likely to be small relative to N, these estimators may have little practical difference.

4.3 Using adaptive cluster sampling

I do not recommend using adaptive cluster sampling for ROV surveys of rockfish. First, given the variability of the rockfish populations examined, the distribution of rockfish may not be clustered enough in many locations. Second, the time required to complete an ACS survey would likely be similar to the time required to complete a SRS survey. When equal total sample sizes are used for SRS and ACS, SRS surveys typically have lower variance. Third, using a large number of initial samples would likely lead to prohibitively costly ROV surveys. Increased precision of ROV surveys is needed to make ROVs a more widely-used and reliable survey tool, however, my simulation suggests that using adaptive cluster sampling is not the answer. Resources might be better used by developing more accurate habitat maps to allow more precise survey stratification, which is recommended as an effective way to reduce sampling variance (Cochran 1963).

Using adaptive cluster sampling can be an effective way to reduce sampling variance for other types of organisms and survey methods; however, its applicability may be limited. Because the cost-effectiveness of ACS depends on the aggregation of individuals in a population, using ACS for surveys when little information is available about the population is not recommended (Hanselman et al. 2003). One should know *a priori* that the population is clustered and rare. Perhaps most importantly, however, the cost of adaptively sampling units should be 1/3 or less than the cost of randomly sampling units to allow more units to be sampled with ACS than an equivalently priced SRS survey.

4.4 Study limitations

The conclusions reached here about the performance of random and adaptive sampling designs for surveys of rockfish could be questioned based on several study limitations. These limitations fall into three main categories: (i) the quality and quantity of data, (ii) the method used for characterizing rockfish distributions, and (iii) the assumptions made about ROV sampling of rockfish.

The data I used could be questioned based on the quality of spatial information they contain and the number of study sites. The accuracy of the spatial information is subject to fish behavioural biases, observational error and equipment error. Behavioural biases may occur if the vehicle attracts or repels fish. Anecdotal evidence suggests rockfish do not exhibit strong attraction or avoidance behaviour; however, this hypothesis has not been explicitly tested (Pacunski et al. 2008). Observational error may occur if the location or number of fish is recorded inaccurately, or when observers miss fish entirely. Richards (1986) suggests that this type of bias is greatest in underwater visual surveys for small, cryptic, or schooling species. In addition to observers, the tracking equipment used in these surveys may also have introduced errors in location data. Tracking was not constant; vehicle fixes occurred every 1-5 seconds, necessitating interpolation for some surveys. In addition, tracking locations are imprecise; for example, the ORE Offshore ® Trackpoint II Ultrashort baseline system acoustic transponder, used in the SJC (Pacunski et al. 2008) and QCI surveys (L. Yamanaka, pers. comm.), is ±5 m, when functioning correctly (Karpov et al.2006). However, numerous difficulties can prevent even this level of accuracy, including inaccurate depth or heading information and interference with the acoustic signal (Pacunski et al. 2008). To some extent, computer algorithms can be used to smooth the data but filling in large gaps in the data requires making assumptions based on the video or support vessel tracking. For the purposes of this study, I assumed location errors did not mask the true fish distributions; however I have no way of testing this assumption.

Even if the data from the sites was of sufficient quality, there is a small number of sites. Transects sampled may not have been representative of rockfish distributions within a site because sampling was not strictly random in any of these surveys. The sites sampled may not be representative of rockfish over a larger area because only seven were used in this study; five sites were used to parameterize the model and the fit of this model was tested against an additional two sites. Other rockfish sites may have entirely different distributions. In addition, different seasons or years may have exhibited different distributions within these sites. Data from multiple years are available for some the sites used in this analysis; however, analyzing these data was beyond the scope of this analysis.

The second main category of limitation in this study is the method I used to characterize the spatial data, which was fitting a Poisson cluster model to the data. I could have selected a different model or I could have taken a non-parametric approach, by bootstrapping the transect data to create new transects. Bootstrapping would have captured more of the characteristics in the real data; would not have required making any assumptions about the distribution of the population; and would have been less complicated to apply, and therefore, less prone to errors. Using a Poisson cluster model assumes that the population is characterized by clusters that are the same size and contain the same number of fish and that the intensity of these clusters is constant across the study area. In reality, rockfish distributions likely exhibit trends in intensity, with habitat, for example, and populations contain clusters of multiple sizes with different numbers of individual fish. However, the use of the Poisson cluster process also allowed me to easily compare my results to other ACS studies that simulated data using the Poisson cluster process. Further, using the Poisson cluster process allowed me to examine the range of parameter distributions obtained in further depth to determine which aspects of spatial distribution were important to the performance of ACS.

The method of fitting the Poisson cluster process required several assumptions. I had to assume that the probability of detection in the original distance sampling surveys followed a half-normal distribution. Cowling (1998) found that the estimation method was robust to other detection functions, provided the shoulder of the detection function was at least as wide as the shoulder in the half-normal detection function. I also assume that the probability of detection on the line, $g(0) = 1$. However, in highly complex habitat, some fish may have been missed on the line. I also assume that the value of the strip half-width σ in the detection function is the same for each study.

The final category of assumption was in the way I simulated rockfish distributions and sampling on the distributions. Because each simulated distribution is sampled only once, sampling variability and variability in rockfish distributions are mixed. I set-up the simulations in this way because I was interested in the effect of, both, sampling and distribution variability and including multiple samplings of individual populations would have resulted in an intractable number of simulations. However, if I had known what the distribution of rockfish was, I would have repeated sampling on that one distribution multiple times in order to more fully capture the average sampling performance.

Another potentially relevant detail my simulated sampling protocol was that I did not allow a population estimate of zero to occur; I increased the number of initial transects until at least one fish was observed. This likely resulted in a slight overestimate of the bias of all estimators because the population would have otherwise been underestimated in those simulations. This effect is likely quite small because only marginal increases in sample size were required to ensure that at least fish was observed.

I also assumed that all fish were observed in a transect, when in real ROV surveys, a number of fish will be missed. I assumed that a survey area could be divided into equally sized transects and that all transect surveys could be completed. In reality, variable and challenging habitat make pre-assigning transects carrying out all transects as planned almost impossible. Likewise, completing surveys of adjacent transects that seamlessly abut one another would also be quite difficult. I simulated transects that were 1000 m long, while actual ROV transects were approximately 400 m in length in SJC and CHR. There are likely many more assumptions of this nature. Modelling more realistic ROV simulations would be possible, and may even be desirable before an ACS survey of rockfish was carried out with an ROV. However, it is questionable that making the survey simulations more realistic would change the overall finding that SRS resulted in more precise density estimates than ACS.

I suggest that the conclusion that ACS ROV surveys of rockfish are not more efficient than SRS surveys would not be affected by the limitations discussed above. Although the rockfish data used may have contained errors and the model used may not represent all rockfish populations, changes to the model did not appear to affect the relative variance of ACS and SRS. Moreover, the conditions required to make ACS more precise than SRS could be quite difficult to fulfil for ROV surveys of rockfish. The site being surveyed would have to have clusters as small as those observed at QCI with an intensity of clusters closer to SEO. In addition, the cost of adaptive units would have to be significantly less than the cost of initial units or identifying edge units prior to sampling would have to be possible. Alternatively, adaptive sampling could be pursued in absence of some of these criteria if there was a budget that allowed for sampling a large number of initial transects and therefore, a large proportion of the total study area.

5 CONCLUSIONS

Simulation modelling provides a useful tool for evaluating alternative ROV sampling protocols. Using data from visual surveys of rockfish to parameterize a model of rockfish distribution increases the chances that simulation results will be transferable to field situations. I did not find adaptive cluster sampling (ACS) for ROV surveys of rockfish to be any more precise simple random sampling (SRS), for equivalent survey cost. However, because the differences between ACS and SRS were not substantial in many cases, the use of ACS may be justified in some circumstances, such as when organisms are rare and distributed in small, tight clusters, a large number of initial samples is planned, or the cost of adaptively sampled units is expected to be much less than the cost of randomly sampled units due to travel costs. In these situations, the use of a stopping rule is recommended because it was found to reduce costs with only minor increases in bias and precision. Those planning studies may also wish to use adaptive cluster sampling if observing as many organisms as possible is important for collecting auxiliary data. For many ROV surveys of rockfish, however, adaptive cluster sampling would not offer any gains in the precision of density estimates and resources might be best utilized exploring other methods of reducing sampling variability.

6 REFERENCE LIST

Aldrin, M., M. Holden, and T. Schweder. 2003. Comment on Cowling's spatial methods for line transect surveys. Biometrics **59**: 186-188.

Baddeley, A. and R. Turner. 2005. Spatstat: an R package for analyzing spatial point patterns. Journal of Statistical Software **12**: 1-42. ISSN: 1548-7660. URL: www.jstatsoft.org

Batista, J.L.F. and D.A. Maguire. 1998. Modeling the spatial structure of tropical forests. Forest Ecology and Management **110**: 293-314.

Brown, B.M. and A. Cowling 1998. Clustering and abundance estimation for Neyman-Scott models and line transect surveys. Biometrika **85**: 427-438.

Brown, J. A. 2003. Designing an efficient adaptive cluster sample. Environmental and Ecological Statistics **10**: 95-105.

Brown, J.A. and B.J.F. Manly. 1998. Restricted adaptive cluster sampling. Environmental and Ecological Statistics **5**: 49-63.

Buckland, S.T., D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers, and L. Thomas. 2001. Introduction to Distance Sampling. Oxford University Press, Oxford.

Byerly, M. 2007. Project 6: Lingcod population parameters in the Northern Gulf of Alaska. *In* Nearshore marine research in Alaska (V): Final comprehensive progress report. NOAA Cooperative Agreement NA03NMF4370202. *Edited by* D. Carlile.pp. 42-76. Alaska Department of Fish and Game, Division of Commercial Fisheries, Juneau.

Christman, M.C. 1997. Efficiency of some sampling designs for spatially clustered populations. Environmetrics **8**: 145-166.

Cowling, A. 1998. Spatial methods for line transect surveys. Biometrics **54**: 828-839.

Cochran, W.G. 1963. Sampling Techniques. John Wiley & Sons, New York.

Cressie, N.A.C. 1993. Statistics for Spatial Data. John Wiley & Sons, Inc., Toronto.

Devore, J,L. 2000. Probability and Statistics for Engineering and the Sciences. Duxbury, California.

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

DiCosimo, J. 1998. Groundfish of the Gulf of Alaska: a species profile. North Pacific Fisheries Management Council, Anchorage.

Fisheries and Oceans Canada. 2008. Pacific Region Amended Integrated Fisheries Management Plan, Groundfish, March 8, 2008 to February 20, 2009.

Fox, D., A. Merems, M. Amend, H. Weeks, C. Rosmos, and M. Appy. 2004. Comparative characterization of two nearshore rocky reef areas: a high-use recreational fishing reef vs. an unfished reef.66 pp. Oregon Department of Fish and Wildlife Marine Program, Newport.

Hagen, G. And T. Schweder. 1995. Point clustering of mike whales in the northeastern Atlantic. *In* Whales, Seals, Fish and Man. *Edited by* A.S. Blix, L. Walløe, and Ø. Uiltang. pp. 27-33. Elsevier Science B.V., Amsterdam.

Hanselman, D.H., T.J. Quinn, J. Heifetz, D. Clausen, and C. Lunsford. 2001. Spatial inferences from adaptive cluster sampling of Gulf of Alaska rockfish. *In* Proceedings of the 17th Lowell-Wakefield symposium: Spatial Processes and Management of Marine Populations. *Edited by* G.S. Kruse, N. Bez, A. Booth, M.W. Dorn, S. Hills, R.N. Lipcus, D. Pelletier, C. Roy, S.J. Smith, D. Witherell. pp. 303-325. University of Alaska Sea Grant Program, Fairbanks.

Hanselman, D.H., T.J. Quinn, C. Lunsford, J. Heifetz, and D. Clausen. 2003. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. Fishery Bulletin **101**: 501-513.

Johnson, S.W., M.L. Murphy, D.J. Csepp. 2003. Distribution, habitat, and behaviour of rockfishes, *Sebastes* spp., in nearshore waters of southeastern Alaska: observations from a remotely operated vehicle. Environmental Biology of Fishes **66**: 259-270.

Karpov, K.A., A. Lauermann, M. Bergen, and M. Prall. 2006. Accuracy and precision of measurements of transect length and width made with a remotely operated vehicle. Marine Technology Society Journal **40**: 79-84.

Krieger, K.J. 1993. Distribution and abundance of rockfish determined from a submersible and by bottom trawling. Fishery Bulletin **91**: 87-96.

Lo, N.C.H., D. Griffith, and J.R. Hunter. 1997. Using a restricted adaptive cluster sampling to estimate pacific hake larval abundance. CalCOFI Report **38**: 103-113.

Love, M.S., M. Yoklavich, L. Thorsteinson. 2002. The Rockfishes of the Northeast Pacific. University of California Press, Berkeley.

Lunsford, C. L. Haldorson, J.T. Fujikoka, and T.J. Quinn. 2001. Distribution patterns and survey design considerations of Pacific Ocean Perch (*Sebastes alutus*) in the Gulf of Alaska. *In* Proceedings of the 17th Lowell-Wakefield symposium: Spatial Processes and Management of Marine Populations. *Edited by* G.S. Kruse, N. Bez, A. Booth, M.W. Dorn, S. Hills, R.N. Lipcus, D. Pelletier, C. Roy, S.J. Smith, D. Witherell. pp. 281-302. University of Alaska Sea Grant Program, Fairbanks.

Murie, D.J., D.C. Parkyn, B.G. Clapp, and G.G. Krause. 1994. Observations on the distribution and activities of rockfish, *Sebastes* spp., in Saanich Inlet, British Columbia, from the *Pisces IV* submersible. Fishery Bulletin **92**: 313-323.

O'Connell, V.M. and D.W. Carlile. 1993. Habitat-specific density of adult yelloweye rockfish *Sebastes ruberrimus* in the eastern Gulf of Alaska. Fishery Bulletin **91**: 304-309.

Pacunski, R.E., W.A. Palsson, H.G. Greene, D. Gunderson. 2008. Conducting visual surveys with a small ROV in shallow water. *In* Marine Habitat Mapping Technology for Alaska. *Edited by* J.R. Reynolds and H.G. Greene. pp. 109-128. University of Alaska Sea Grant Program, Fairbanks.

Peterman, R.M. 1990. Statistical power analysis can improve fisheries research and management. Canadian Journal of Fisheries and Aquatic Sciences **47**: 2-15.

Philippi, T. 2005. Adaptive cluster sampling for estimation of abundances within local populations of low-abundance plants. Ecology **86**: 1091-1100.

R Development Core Team 2008. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Richards, L.J. 1986. Depth and habitat distributions of three species of rockfish (*Sebastes*) in British Columbia: observations from the submersible PISCES IV. Environmental Biology of Fishes **17**: 13-21.

Richards, L.J. 1987. Copper rockfish (*Sebastes caurinus*) and quillback rockfish (*Sebastes maliger*) habitat in the Strait of Georgia, British Columbia. Canadian Journal of Zoology **65**: 3188-3191.

Seber, G.A.F. and S.K. Thompson. 1994. Environmental adaptive sampling. *In* Handbook of Statistics. Volume 12: Environmental Statistics. *Edited by* G.P. Patil, and C.R. Rao. pp. 201–220. Elsevier Science, the Netherlands.

Smith, D.R., R.F. Villella, and D.P. Lemarié. 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environmental and Ecological Statistics **10**: 7-15.

Su, Z. And T.J. Quinn. 2003. Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environmental and Ecological Statistics **10**: 17-41.

Thompson, S.K. 1990. Adaptive cluster sampling. Journal of the American Statistical Association **85**: 1050-1059.

Thompson, S.K. 1994. Factors influencing the efficiency of adaptive cluster sampling. Technical Report Number 94-0301, Technical Reports and Reprints Series, Pennsylvania State University, Department of Statistics.

Thompson, S.K. 2002. Sampling. John Wiley & Sons, Inc., New York.

Thompson, S.K. and G.A.F. Seber 1996. Adaptive Cluster Sampling. John Wiley & Sons, Inc., Canada.

Thompson, W.L. 2004 Sampling Rare or Elusive Species: Concepts, Designs, and Techniques for Estimating Population Parameters. Island Press, Washington.

Turk, P. and J.J. Borkowski. 2005. A review of adaptive cluster sampling: 1990 – 2003. Environmental and Ecological Statistics **12**: 55-94.

Washington State. 2003. San Juan Islands Marine Preserve Area, Washington Administrative Code 220-16-440.

Table 1 **Survey vehicle, transect type, depth, number and species of fish observed, and number and length of transects for each of the four survey locations. The survey tools used were remotely-operated vehicles (ROV) and manned submersibles (Sub.), which surveyed fixed-width transects or line transects, respectively. The most common rockfish species observed included quillback (*Sebastes maliger*; QB), redstripe (*S. proriger*; RS), black (*S. melanops*; BK), yelloweye (*S. ruberrimus*; YE), tiger (*S. nigrocinctus*; TI), rosethorn (*S. helvomaculatus*; RT), sharpchin (*S. zacentrus*; SC), pygmy (*S. wilsoni*; PY), Puget Sound (*S. emphaeus*; PU), and copper (*S. Caurinus*; CO).**

Location	Survey tool	Transect type	Mean survey depth (m)	Max. survey depth (m)	Most common species	No. of transects in survey	No. of transects in analysis	Mean transect width (m)	Length of transects in analysis (m)	No. rockfish in analysis
Chiswell Ridge, AK	ROV	Strip	66	136	QB, RS, BK	62	62	1.94	25 000	1051
Southeast Outside, AK	Sub.	Line	115	226	YE, QB, TI, RT	264	257	-	90 000	4064
Queen Charlotte Islands, BC	Sub.	Line	111	170	SC, RS, QB, PY	17	13	-	8000	3692
San Juan Channel, WA	ROV	Strip	128	160	PU, CO, QB	58	18	1.58	5000	2378

Table 2 **Parameter estimates including number of fish per cluster (μ), intensity of cluster centres (λ), and cluster radius (ρ) for the Thomas process, the Poisson process intensity (γ), and strip half-width of the detection function (σ) for Queen Charlotte Islands, BC (QCI) and four sub-areas within Alaska's Southeast Outside Subdistrict (SEO).**

Site	No. of fish locations	Parameter estimates				
		μ	λ (m^{-2})	ρ (m)	γ (m^{-2})	σ (m)
QCI	3692	246.9	2.5×10^{-4}	0.8	0.060	1.9
SEO-NSEO	169	399.5	2.2×10^{-5}	51.2	0.009	4.0
SEO-CSEO	1892	269.7	3.4×10^{-5}	36.7	0.009	4.1
SEO-EYKT	1113	188.2	5.5×10^{-5}	23.1	0.010	3.7
SEO-SSEO	1430	239.5	2.4×10^{-5}	32.6	0.006	3.5
Mean	-	268.8	7.7×10^{-5}	28.9	0.020	3.4
SD	-	78.9	9.8×10^{-5}	18.7	0.023	0.9

Table 3 **Mean (standard error) of performance indicators from five estimators for 500 repetitions of sampling on study sites simulated using different Thomas process parameters. Parameters were estimated from the least squares minimization procedure (Appendix B) using the maximum separation distance h_f = 50 (baseline), as well as h_f = 25 and h_f = 100. ACS-S indicates adaptive cluster sampling with a stopping condition. All sampling parameters are baseline values. For SRS, the cost is equal to the sample size.**

h_f	Method	Estimator	Relative Bias (%)	Relative Standard Error (%)	Cost	Proportion of Fish Sampled
50	SRS	$\bar{y}$	-1.5 (73.6)	39.3 (70.3)	75.4 (123.5)	0.15 (0.25)
	ACS	HT	5.1 (86.7)	52.9 (81.5)	43.2 (60.8)	0.20 (0.26)
		HH	4.4 (86.0)	52.9 (81.9)		
	ACS-S	HT	8.5 (88.7)	54.4 (82.5)	19.2 (10.9)	0.08 (0.07)
		HH	5.7 (87.5)	54.5 (82.6)		
25	SRS	$\bar{y}$	-0.9 (32.4)	25.3 (26.2)	82.1 (129.0)	0.16 (0.26)
	ACS	HT	1.3 (44.1)	35.6 (31.4)	45.8 (63.6)	0.18 (0.27)
		HH	0.7 (44.2)	34.6 (32.1)		
	ACS-S	HT	3.6 (45.0)	35.7 (31.5)	19.5 (9.2)	0.07 (0.04)
		HH	1.4 (44.5)	35.7 (31.6)		
100	SRS	$\bar{y}$	-4.8 (75.1)	31.3 (57.8)	95.5 (133.9)	0.19 (0.27)
	ACS	HT	5.8 (74.5)	49.2 (69.0)	53.5 (66.1)	0.29 (0.30)
		HH	5.9 (77.9)	50.3 (70.5)		
	ACS-S	HT	11.3 (82.0)	52.5 (70.8)	20.8 (14.0)	0.10 (0.09)
		HH	7.7 (79.5)	52.7 (71.0)		

Table 4 Mean Thomas process parameter values from simulated data where relative variance of ACS is greater than 1 compared to the mean parameter values over all simulated data. Relative variance (*RV*) is the ratio of variance from SRS sampling to ACS sampling.

	Parameter		
Mean value	μ	λ **(m^{-2})**	ρ **(m)**
RV > 1	260.1	3.4 x 10^{-5}	2.2
All	256.4	4.6 x 10^{-5}	13.7

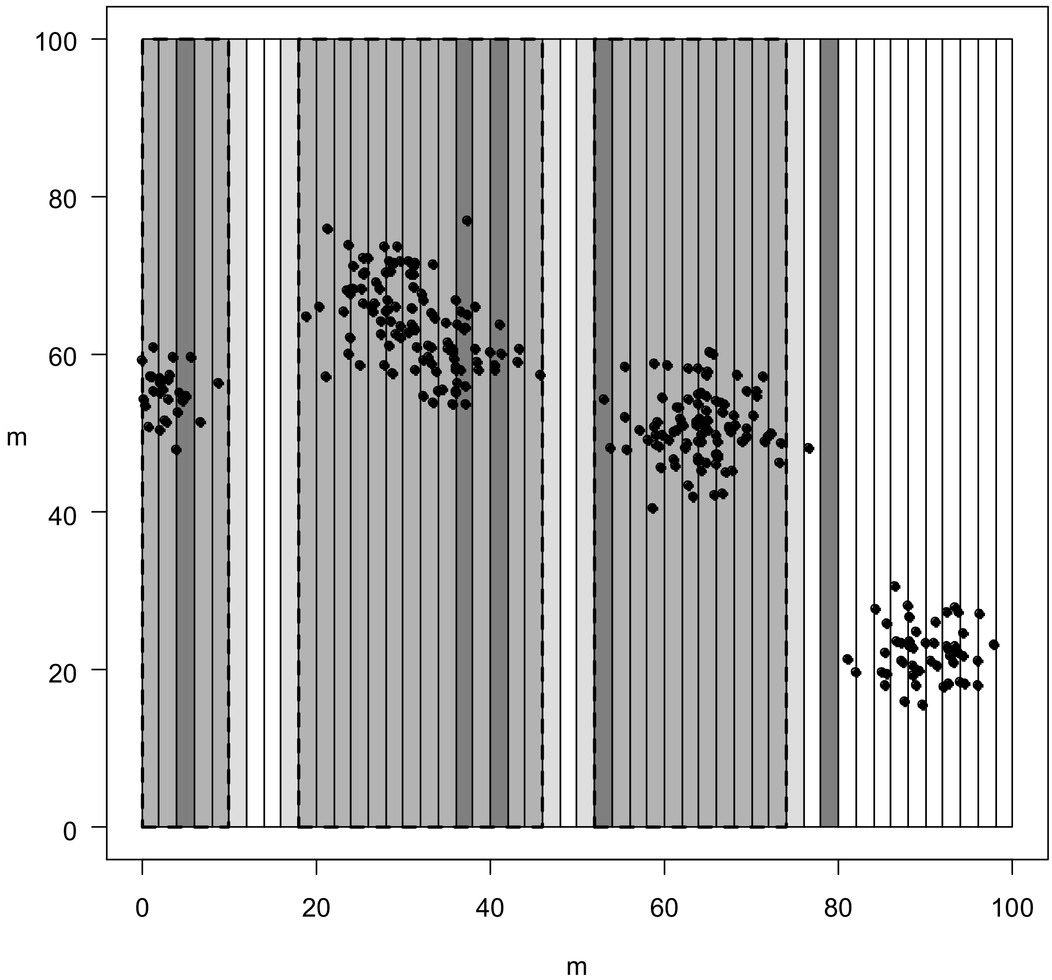

Figure 1 **Example of unrestricted adaptive cluster sampling (ACS) with area 100 m^2 and 50, 2 m-wide transects. The darkest five transects are the initially sampled units. Among these transects, four meet the condition *C* = 1 for resampling. The medium grey transects are the neighbours, sampled in additional rounds of sampling, which meet *C*. Each initial unit and its neighbours that meet *C* form a network, enclosed by a dashed line. The lightest grey bars are sampled as neighbours, but because they do not meet the condition *C*, they are excluded from the network and the population estimate.**

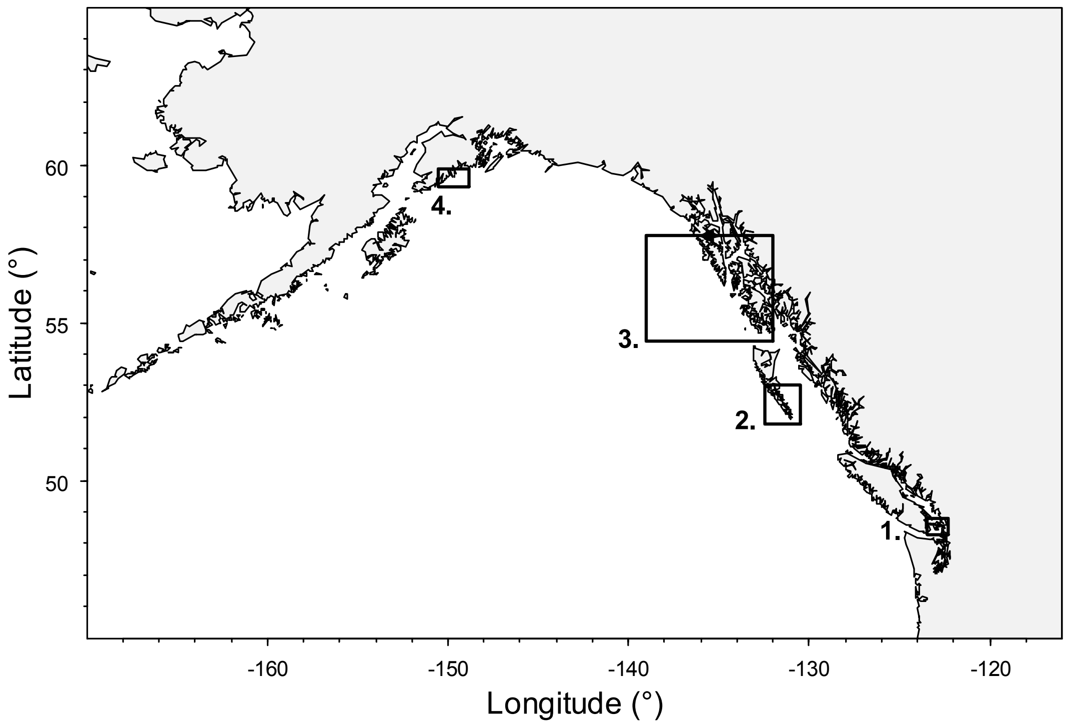

Figure 2 Underwater visual survey sites: 1) ROV survey in San Juan Channel, Washington, 2) submersible survey in Juan Perez Sound, Queen Charlotte Islands (QCI), British Columbia, 3) submersible survey in Southeast Outside Subdistrict (SEO), Alaska, and 4). ROV survey on Chiswell Ridge, Alaska. The SEO included 4 sites including East Yakutat (EYKT), Northern Southeast Outside (NSEO), Central Southeast Outside (CSEO), and Southern Southeast Outside (SSEO).

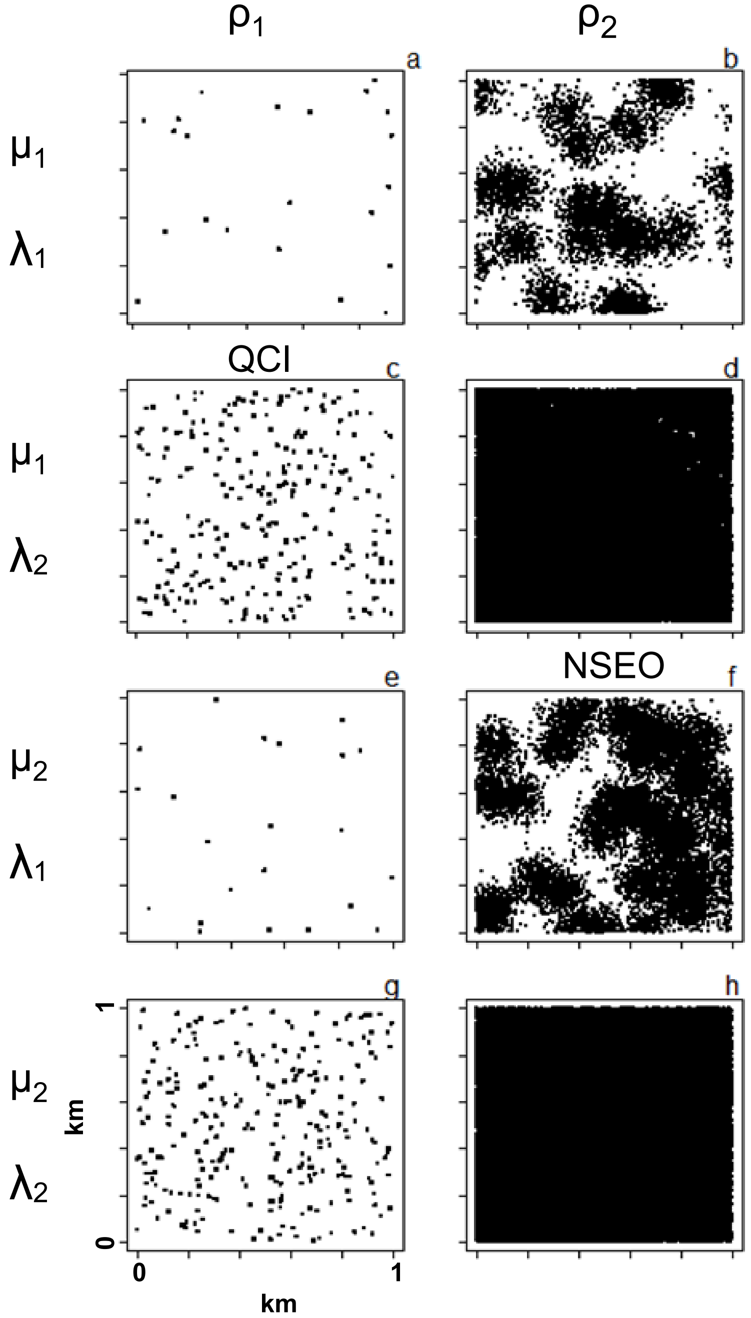

Figure 3 **Simulated spatial distributions of rockfish within 1-km^2 survey areas based on the estimated Thomas process parameters, where μ_1 = 247, μ_2 = 399, ρ_1 =0.8, ρ_2 =51, λ_1 = 2.2 x 10^{-5}, λ_2 = 2.4 x 10^{-5}). Simulations represent the extreme values for all sites. Patterns typical of the Queen Charlotte Islands, BC (μ_1, ρ_1, λ_2) and NSEO within the Southeast Outside District, AK (μ_2, ρ_2, λ_1) are labeled above the representative plot. Points represent individual fish.**

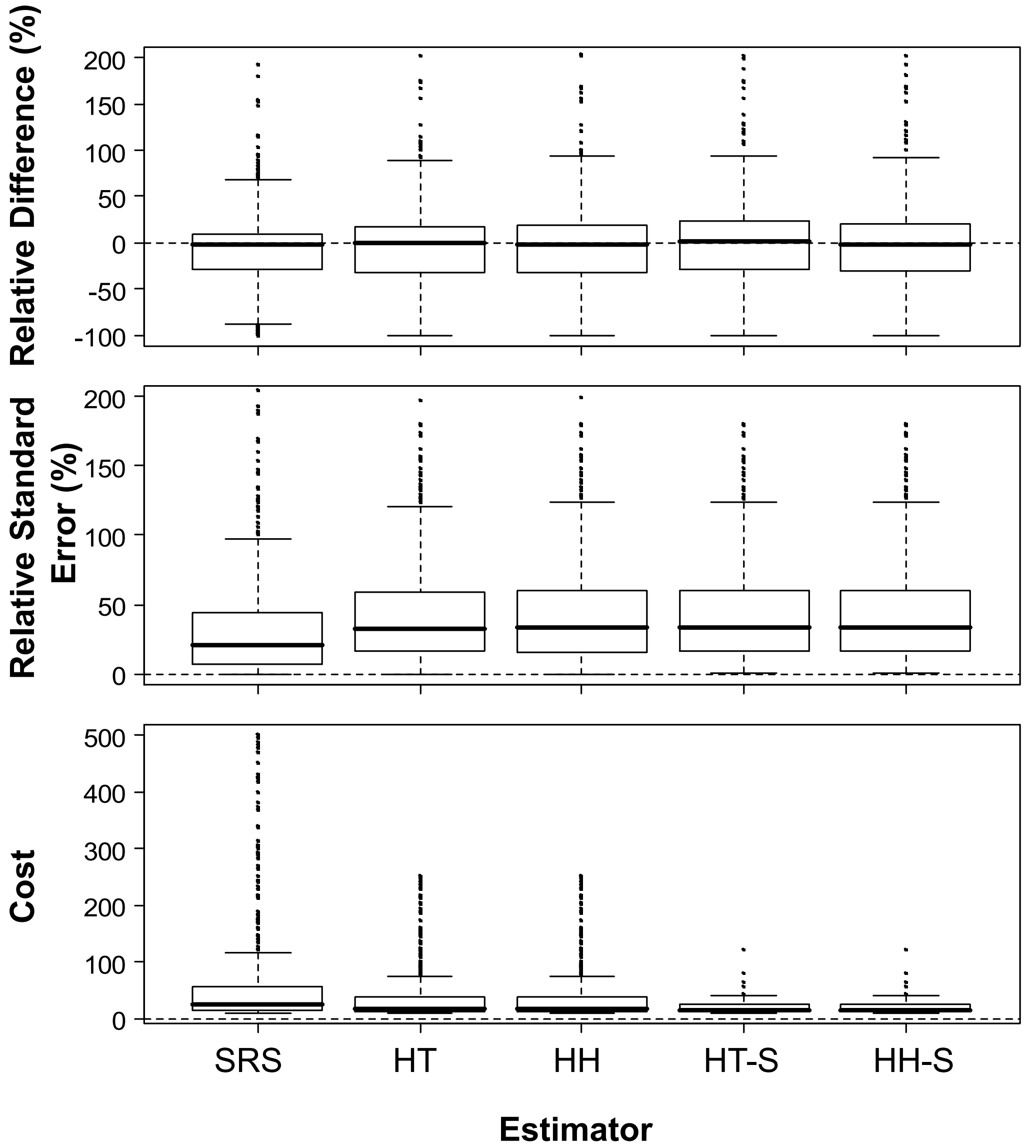

Figure 4 **The performance of the five sampling protocols under the clustered rockfish model. Performance measures include: relative difference (%), relative standard error (%), and cost. The Sampling protocols used were: simple random sampling (SRS), adaptive cluster sampling with Horvitz-Thompson estimator (HT), adaptive cluster sampling with Hansen-Hurwitz estimator (HH), adaptive cluster sampling with stopping rule and Horvitz-Thompson estimator (HT-s), and adaptive cluster sampling with stopping rule and Hansen-Hurwitz estimator (HH-s). Outliers are more than 1.5 times the distance of the inter-quartile range.**

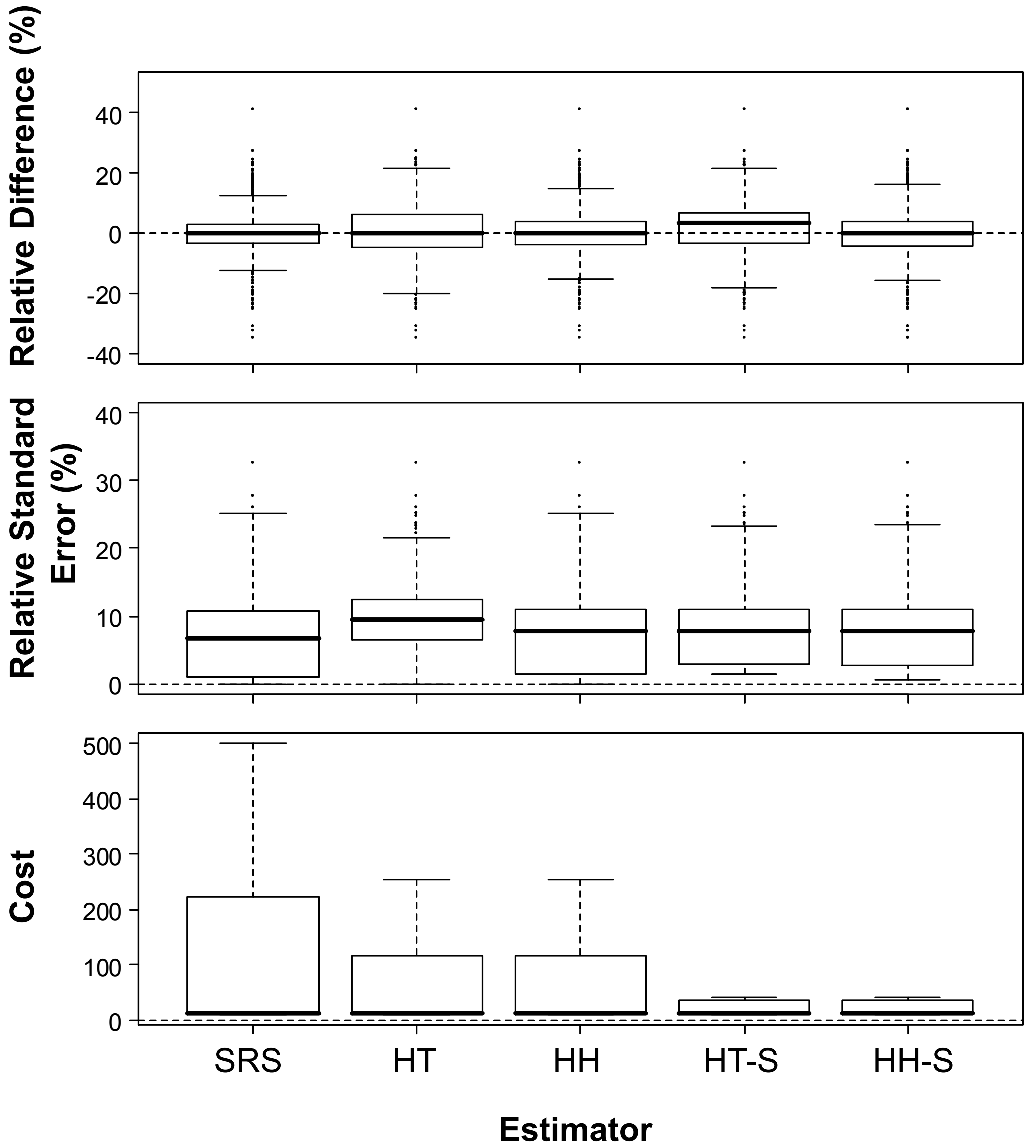

Figure 5 **Simulated performance of five sampling protocols under the random rockfish model. Sampling protocols are simple random sampling (SRS), adaptive cluster sampling with Horvitz-Thompson estimator (HT), adaptive cluster sampling with Hansen-Hurwitz estimator (HH), adaptive cluster sampling with stopping rule and Horvitz-Thompson estimator (HT-s), and adaptive cluster sampling with stopping rule and Hansen-Hurwitz estimator (HH-s). Outliers are more than 1.5 times the distance of the inter-quartile range.**

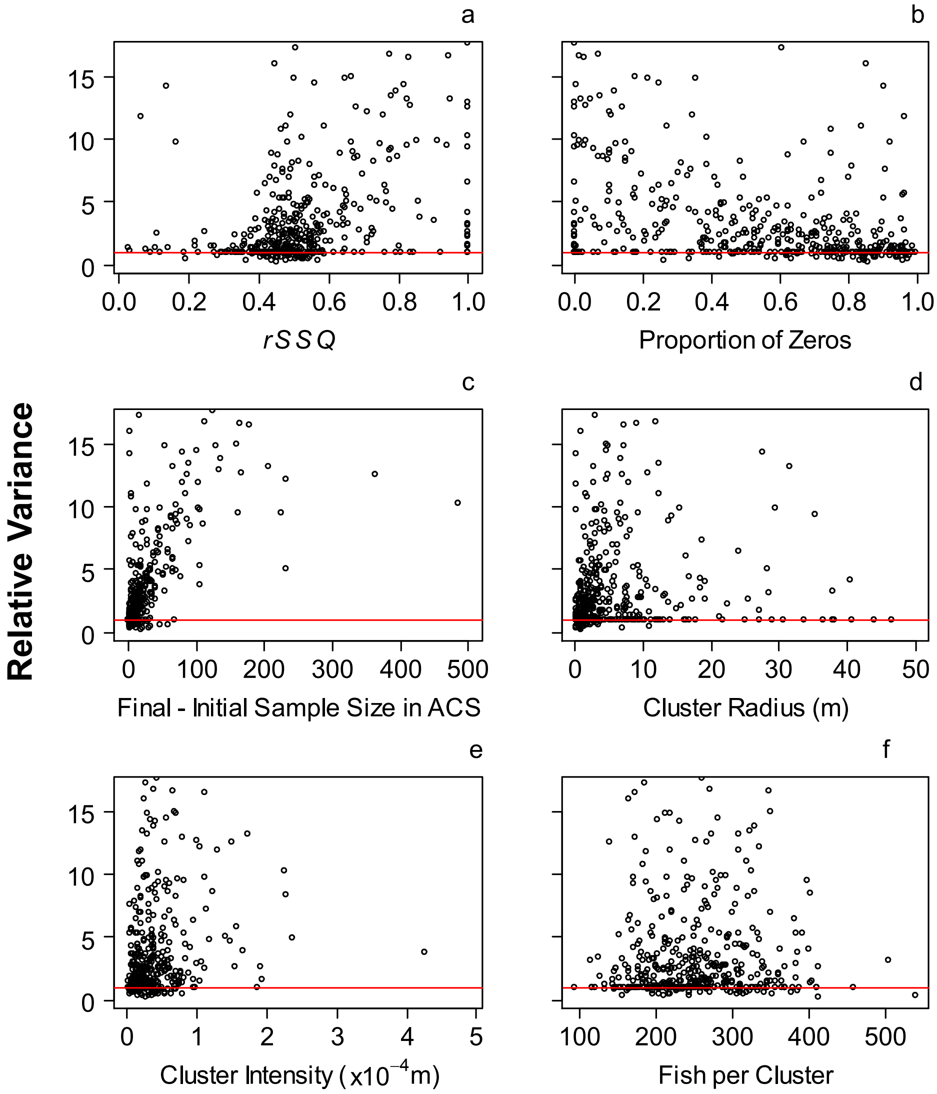

Figure 6 **The relative variance, that is, the ratio of the sample variance for ACS to SRS using the HH estimator are shown versus the factors identified as affecting the efficiency of adaptive cluster sampling. Relative variance < 1 indicates that ACS is more precise than SRS. Factors include: a. the ratio of within network sum of squares to total sum of squares (*rSSQ*), b. the proportion of zeros (proportion of transects with no fish), c. the difference between the initial and final sample size in ACS, and the Thomas process parameters, d. cluster radius (ρ), e. cluster intensity (λ), and f. fish per cluster (μ). Each point represents one simulation of sampling using the baseline conditions: clustered population, condition for resampling (*C*) = 20, number of intial samples (n_1) = 10.**

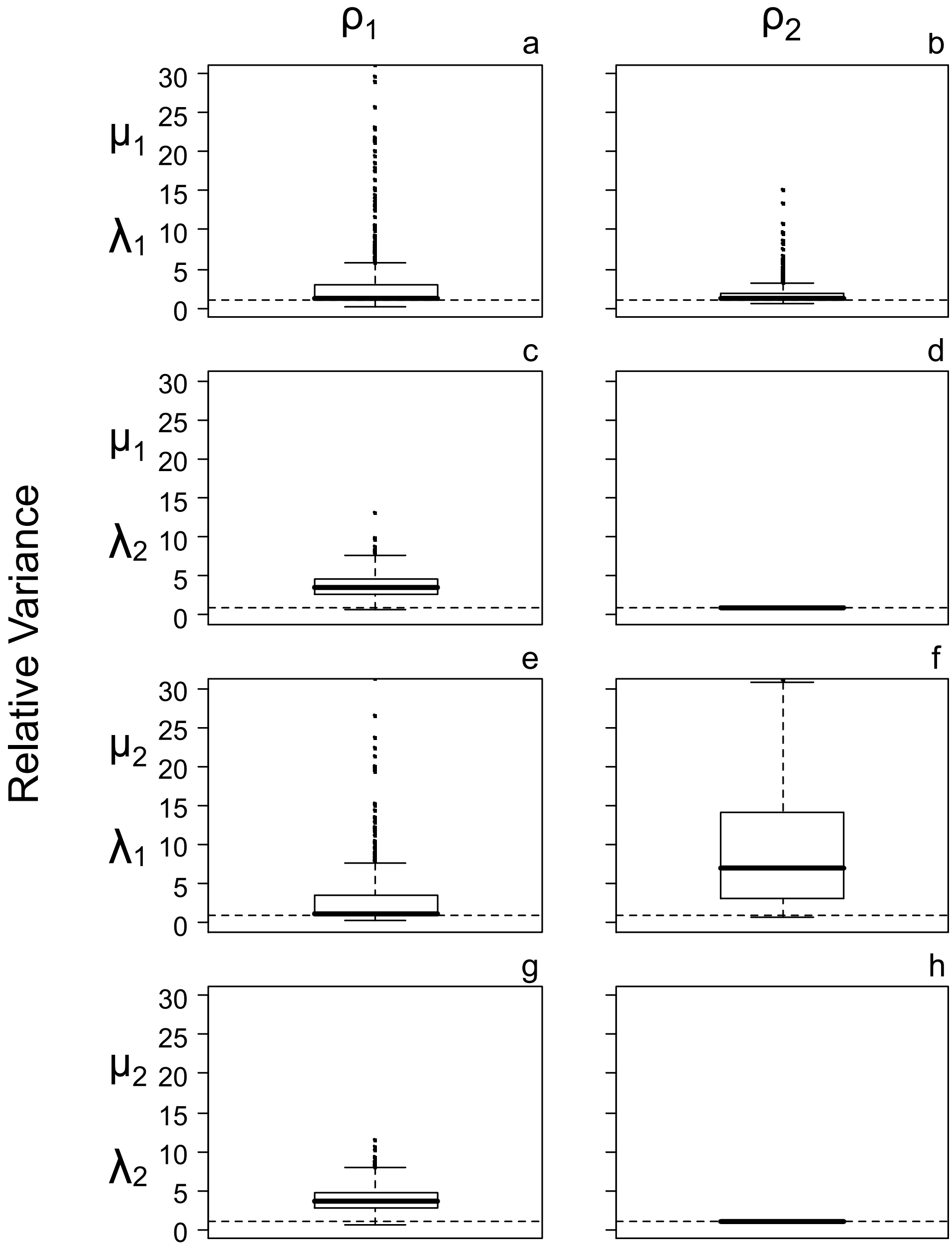

Figure 7 **The relative variance of ACS to SRS for populations simulated with constant parameter values, where fish per cluster μ_1 = 247, μ_2 = 399, cluster radius ρ_1 =0.8, ρ_2 =51, and cluster intensity λ_1 = 2.2 x 10^{-5}, λ_2 = 2.4 x 10^{-5}. Relative variance< 1 indicates ACS has a lower variance than SRS.**

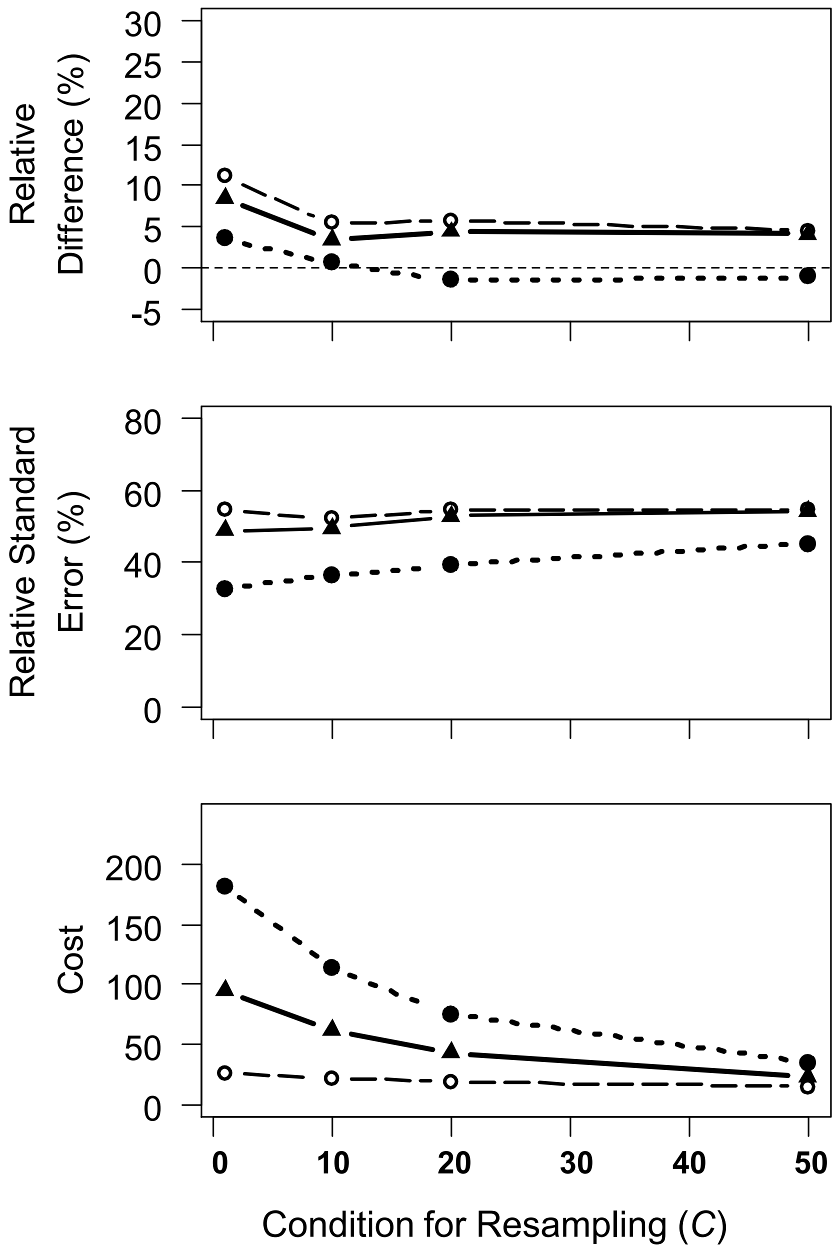

Figure 8 **Change in mean relative percentage difference, relative standard error, and cost with changes in resampling condition *C* for SRS (dotted line with solid circles), ACS with HH estimator (solid line with triangles), and ACS with stopping condition and HH estimator (thin dashed line with open circles). Based on 500 simulations with all other parameters set to baseline values.**

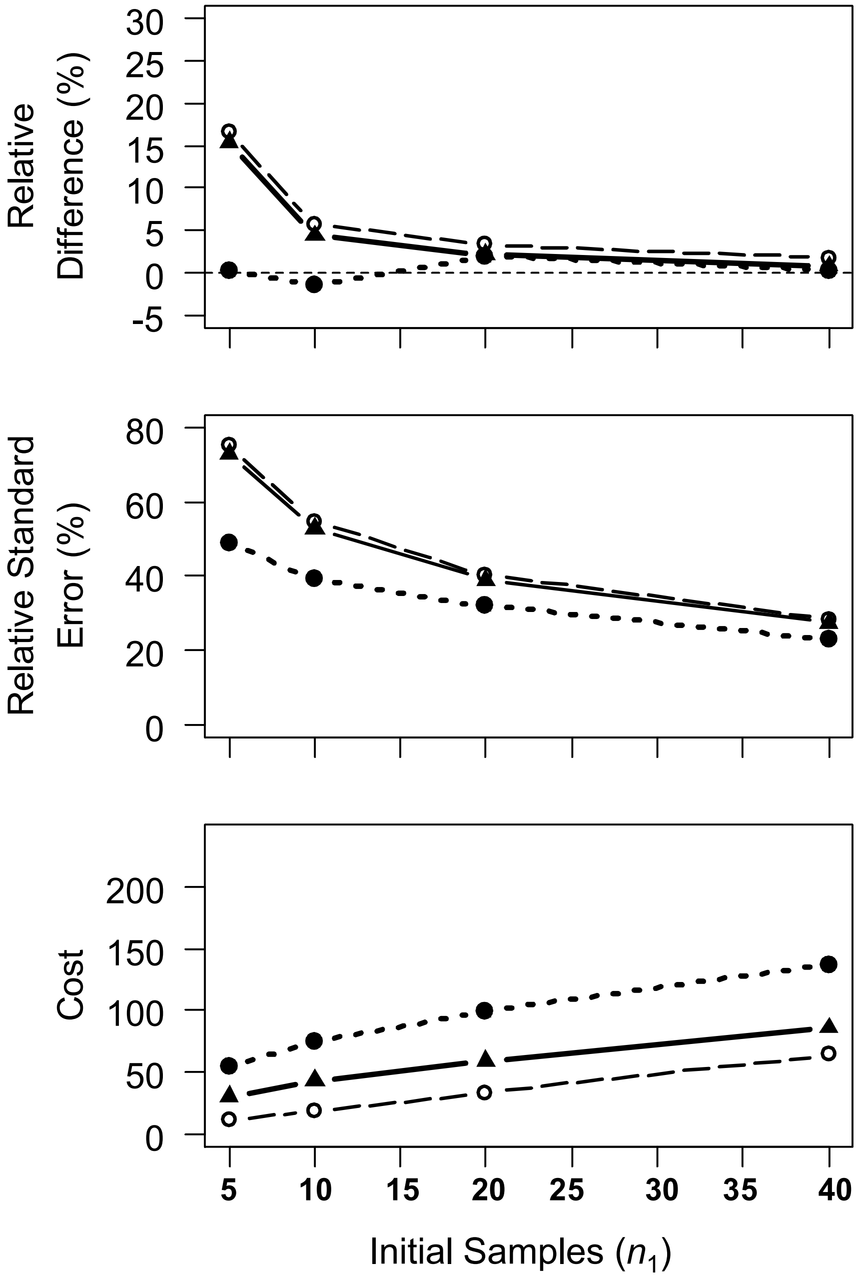

Figure 9 **Change in mean relative difference, relative standard error, and cost with changes in the number of initial samples n_1 for SRS (dotted line with solid circles), ACS with HH estimator (solid line with triangles), and ACS with stopping condition and HH estimator (thin dashed line with open circles). Based on 500 simulations with all other parameters set to baseline values.**

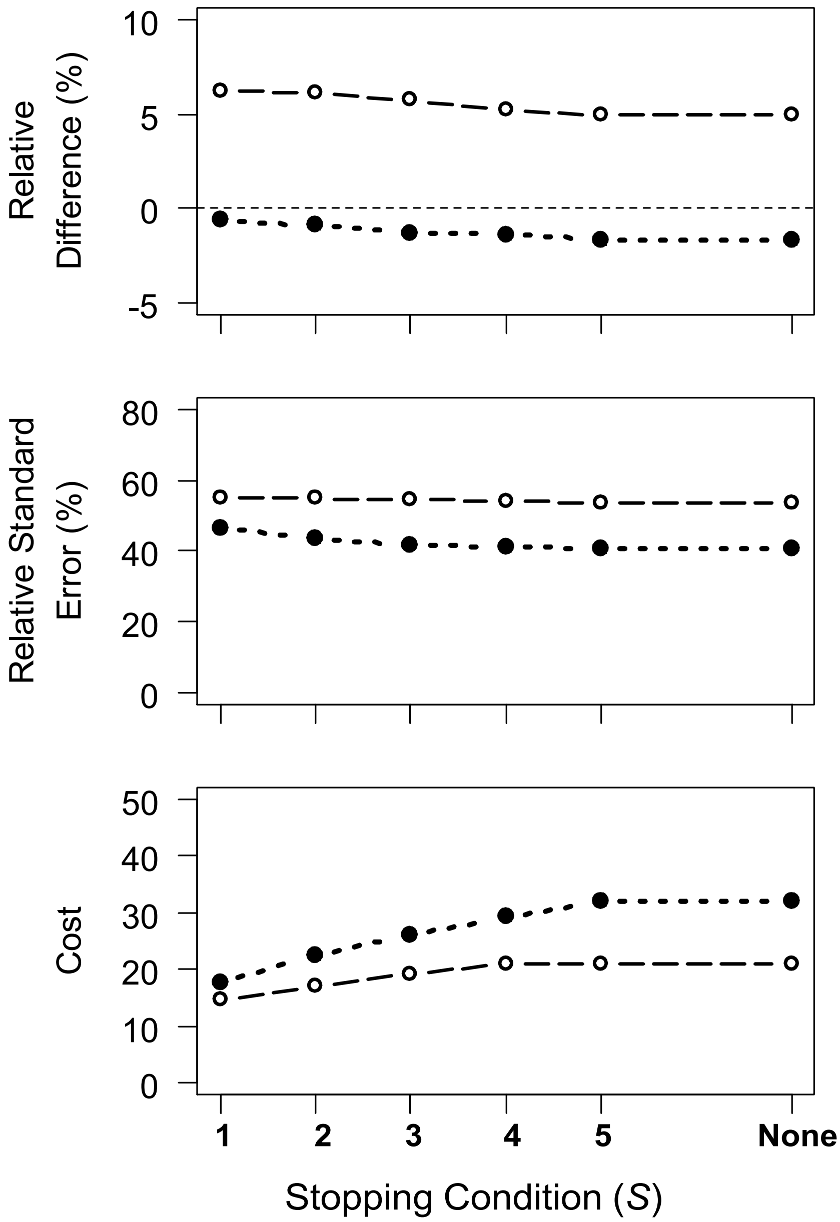

Figure 10 **Change in mean relative difference, relative standard error, and cost with changes in stopping condition *S* for ACS with stopping condition and HH estimator (thin dashed line with open circles) and SRS with sample size equal to that of ACS with stopping condition (dotted line with solid circles). When *S* = "None", the results for ACS with HH estimator and no stopping condition are shown, as well as SRS with equivalent sample size. Based on 500 simulations with all other parameters set to baseline values.**

APPENDIX A: UNDERWATER VISUAL SURVEY DATA

This appendix describes the data sources and survey method for collecting inshore rockfish visual survey data from manned and unmanned submersibles.

The Alaska Department of Fish and Game (ADFG) has conducted surveys using the *Delta* submersible in four management sections of the Southeast Outside (SEO) subdistrict in the eastern Gulf of Alaska. Although the ADFG has conducted these surveys on a rotational basis since 1989 (C. Brylinsky, pers. comm.), I only used one year of data for each of the four sites to avoid potential pseudoreplication: East Yakutat (EYKT) in 1997, Northern Southeast Outside (NSEO) in 2001, Central Southeast Outside (CSEO) in 1997, and Southern Southeast Outside (SSEO) in 1999.

Mean and maximum survey depth were least in CHR and greatest in SJC and SEO, respectively (Table 1). All sites were exposed to recreational and/or commercial fishing (including targeted rockfish fishing) in the year the data were collected (DiCosmo 1998, Byerly 2007, L. Yamanaka, pers. comm.), with the exception of the SJC. At this site, three of the seventeen transects were located in the Shaw Island marine preserve, which has been closed to recreational and commercial bottom fishing since 1990 (Washington State 2003).

All datasets included longitude and latitude of the submersible or ROV and associated fish observations at one (QCI, CHR), two (SJC), or five (SEO) second intervals. All fish were identified to species. Distances from the transect line to each fish and strip width were included for line and strip transect surveys, respectively. I interpolated fish locations based on the time of observation for the QCI and SEO datasets; the latitude and longitude of individual fish were available for CHR and SJC. I removed transects where positional data were unavailable or no fish were observed. The greatest number of transects had to be removed from the SJC data and these removals were mostly because of ROV positional problems; only 18 of the 58 transects surveyed were used in the analysis. In contrast, all of the 62 transects surveyed in CHR were included in the analyses. The number of removals for the other surveys fell somewhere in between CHR and SJC (Table 1).

All surveys employed variations of random-stratified sampling. Details on survey design for Chiswell Ridge and the San Juan Channel appear in Byerly (2007) and Pacunski et al. (2008), respectively. In the Southeast Outside subdistrict, starting points were randomly placed across each of the four study areas and only those points lying within what is believed to be rocky habitat based on fishermen logbook data or remote sensing were surveyed (C. Brylinsky, pers comm.). In the Queen Charlotte Islands, Juan Perez Sound was divided into 2 km grid blocks. Blocks were assessed as low, medium or high rocky ridge habitat, based on a bathymetric position index analysis, and a random selection of medium and high habitats were surveyed (L. Yamanaka, pers. comm.).

Although species composition varied among sites (Table 1), all surveys observed inshore or nearshore rockfish (as opposed to deep water or pelagic species) and some species, including quillback rockfish (*S. maliger*), yelloweye rockfish (*S. ruberrimus*), tiger rockfish (*S. nigrocinctus*), and lingcod (*Ophiodon elongatus*) were observed in all locations. The greatest number of rockfish were observed in SEO (Table 1); however, this location also had the longest survey length; the mean number of rockfish observed per metre of transect is approximately 0.045 rockfish/m. In CHR, the density of rockfish was similar at 0.042, while the densities in QCI and SJC were an order of magnitude larger at 0.46 and 0.47 rockfish/m respectively.

APPENDIX B: PARAMATERIZING THOMAS AND POISSON MODELS OF INSHORE ROCKFISH SPATIAL DISTRIBUTION

This appendix describes my approach to parameter estimation for models of inshore rockfish small-scale spatial distribution. The models are based on Thomas and Poisson spatial process, which differ in their degree of clustering or patchiness. The Thomas process generates highly clustered rockfish locations, while the Poisson process generates locations that are uniformly random within the survey area. By parameterising these spatial models for 5 sites surveyed by manned submersibles, I show that both processes appear to represent actual rockfish location observations in the northeast Pacific. All analyses were completed using **R** statistical software (R Development Core Team 2008).

Methods

Because the Poisson is a special case of the Thomas point process, I only describe parameter estimation details for the latter. In the Poisson process, the number of fish in a region follows a Poisson distribution with mean γA, where *A* is the area of the region and γ is the intensity (individuals per unit area). The intensity parameter γ is estimated as part of the procedure for estimation parameters of the Thomas process.

In the Thomas process, invisible parent events are distributed through a 2-dimensional sampling area according to a Poisson distribution with λ parent events per unit area. Each parent event independently produces a random number of offspring from a Poisson distribution with mean μ. The realised offspring events are then spatially grouped around their invisible parents according to a bivariate normal distribution with variance ρ^2 in all directions (i.e., isotropic; Cressie 1993). These properties of the Thomas process make it a suitable candidate for modelling small-scale spatial distributions of rockfish. For example, Su and Quinn (2003) used this model to simulate rockfish spatial distribution data in their evaluation of alternative sampling designs; however, they did not use actual data to parameterize the model. Hagen and Schweder (1995) and Cowling (1998) used survey data to estimate Thomas process model parameters for minke whales in the northeastern Atlantic; however, they did not test survey designs using data simulated from this process.

I estimated the Thomas process parameters using a procedure based on a summary statistic of the point process, rather than a direct parametric analysis of clustering (Diggle 1983). Although such parametric approaches are sometimes used for estimating the Thomas process parameters from line transect data (e.g., Brown and Cowling 1998), I selected the summary statistic method based on Ripley's *K*-function for its relative analytical simplicity. Ripley's *K*-function counts the number of pairs of fish separated by less than a given distance *h.* By calculating the number of pairs over a range of distances $h = \{0, h_1, h_2, ...h_f\}$, one obtains a function that is a measure of aggregation at different spatial scales. The method relies on least squares estimation in which differences between theoretical and empirical Ripley's *K*-functions are minimized. For this paper, I used the one-dimensional version of Ripley's *K*-function because there was not enough spatial information in the perpendicular distances to justify the two-dimensional Ripley's *K*-function. Moreover, perpendicular distances were not recorded in the ROV strip transect surveys.

Estimating Thomas process parameters (ρ, μ, λ) requires estimation of both the empirical *K*-function, $\hat{K}_1(h)$ and theoretical *K*-functions, $K_1(h)$.The empirical *K*-function is based on the observed number of pairs of points that are less than distance *h* away from one another, i.e.,

$$\hat{K}_1(h) = 2n^{-2} L \sum_j \sum_{i \neq j} I(d_{ij}), \tag{B.1}$$

where *L* is the total length of all transects, *n* is the total number of rockfish detected, and d_{ij} is the observed distance between observations *i* and *j.* The indicator function $I(d_{ij}) = 1$ if $d_{ij} < h$ and zero otherwise.

The theoretical *K*-function, $K_1(h)$ is given by:

$$K_1(h \mid \lambda, \rho, \sigma) = \left(2\varphi\left(h / \rho\sqrt{2}\right) - 1\right) \Big/ \left(2\lambda\sqrt{\pi}\sqrt{\sigma^2 + \rho^2}\right), \tag{B.2}$$

where σ is the half-width of the line transect detection function (see below), $\varphi\left(h / \sqrt{2\rho^2}\right)$ is the standard normal distribution function, λ is the parent intensity of the Thomas process, and ρ^2 is the variance of the isotropic bivariate normal distribution. Parameters (λ, ρ) are estimated by minimizing the sum of squared differences between the empirical and theoretical *K*-functions, summed over values of *h*, i.e.,

$$\sum_{h=0}^{h=h_f}\left[\left\{\hat{K}_1(h)\right\}^{c_0}-\left\{K_1(h)\right\}^{c_0}\right]^2, \tag{B.3}$$

where $c_0 = 0.25$ and $h_f = 50$ m are the power transformation and the maximum separation distance, respectively. Diggle (1983) recommends the above c_0 value to reduce the influence of large values of *h*. Because the optimal tuning parameters depend on the scale of interest, choosing a single parameter value for all scales is not possible when patterns are complex (Batista and Maguire 1998). Therefore, I investigated the effect of using h_f = 25 m, and h_f = 100 m on Thomas process parameter estimates. I provide sensitivity analyses of tuning parameter effects on performance of simulated sampling-estimator combinations in the main text.

Finally, the expected number of individuals per parent μ of the Thomas process is (Aldrin et al. 2003)

$$\mu = E[n]/E[clust],$$

where the actual number of fish observed *n* is substituted for the expected value $E[n]$ and the expected number of observed clusters in a transect is:

$$E[clust] = 2\lambda\sigma g(0)L\sqrt{\pi},$$

where the product $2\sigma g(0)L\sqrt{\pi}$ is the average strip area for a line transect (Buckland et al. 2001). The intensity of points γ in the Poisson process was estimated by multiplying the intensity of the clusters by the mean number of individuals in a cluster $\lambda\mu$.

Perpendicular distances from the manned submersible line transect surveys were used to estimate the half-width σ of the detection function. I assumed that observed fish locations arose from a combination of their true spatial distribution and the probability of observing them, which is a decreasing function, $g(x)$, of the perpendicular distance x between each fish in the population and the transect line (Cowling 1998; Aldrin et al. 2003). I used the following half-normal detection function to model this probability (Buckland et al. 2001)

$$g\left(x\right) = g(0)\exp\left(-x^2 / 2\sigma^2\right),$$

where $g(0)$ is the probability of detection at $x = 0$ (i.e., on the transect line) and σ is the half-width or the distance at which as many fish are observed as are missed (Buckland et al. 2001). Here I assumed that $g(0) = 1$. The half-width parameter σ was estimated as

$$\hat{\sigma} = \sqrt{\sum X_i^2 / n},$$

where the X_i are the observed perpendicular distances between each fish and the transect line and n is the total number of fish observed (Buckland et al. 2001).

The spatial model fit to the observed spatial pattern was tested using a Monte Carlo approach, in which the empirical *K*-function from simulated ROV surveys was compared with the empirical *K*-function from the actual ROV surveys from SJC and CHR sites, as well as the manned submersible sites. A simulation approach for this fitting procedure is required because the fit of the pattern cannot be observed directly. The simulation proceeded by randomly generating 500 values each of the three Thomas process parameters from a bias-corrected lognormal distribution, i.e.,

$$\theta_R = \exp(\overline{\theta}) \cdot \exp\left(\delta\tau - \frac{\tau^2}{2}\right), \tag{B.4}$$

where $\overline{\theta}$ is the mean of the log-parameter value over all sites, $\delta \sim N(0,1)$ and τ is the standard deviation of the log-parameter values across sites. For each simulated parameter set, I generated rockfish point locations (using the *rThoma*s function in the R package *spatstat*; Baddeley and Turner 2005) within 500 simulated survey sites each with dimensions 1 km x 25. One strip transect survey 1.58 m in width was placed down the centre of each site and the location of each rockfish within the strip was recorded. I then calculated the power transformed empirical *K*-function $\{\hat{K}_1(h)\}^{c_0}$ from these simulated locations in the manner described above. I obtained the distribution of empirical *K*-function values by pooling the estimates across sites and parameter combinations. This procedure was repeated over a range of the separation distance *h.* The 2.5th and 97.5th percentiles of the empirical *K*-function distribution values for each value of *h* formed a 95% confidence envelope (Diggle 1983).

This simulation procedure was repeated using the Poisson process (using the *rpoispp* function in the R package *spatstat*; Baddeley and Turner 2005) to simulate a similar confidence envelope. I then plotted the transformed empirical *K*-function values from the actual ROV and submersible survey data on top of the two simulated confidence envelopes to determine whether either of the spatial models accurately characterized the empirical data. I refer to the Thomas process as the "clustered rockfish model" and the Poisson process as the "random rockfish model".

Spatial model results

The mean Thomas process parameter estimates for the five manned submersibles sites suggest that rockfish are distributed in dense overlapping clusters with a mean diameter of approximately 50 m and 270 fish per cluster. Approximately 80 such clusters could be expected in one square km. Although similar numbers of fish per cluster were observed at all sites, other aspects of distribution varied considerably (Table 2). For example, in the SEO sites, cluster intensity (λ) ranged from 2.2 x 10^{-5} to 5.5 x 10^{-5} clusters/m^2 and the cluster radius (ρ) was 23 m to 51 m. In contrast, cluster intensity was an order of magnitude higher in QCI at 2.5 x 10^{-4} clusters/m^2 while cluster radius was only 0.8 m, indicating that fish were densely packed within clusters and clusters were much more numerous. Because the number of individuals per cluster was approximately equal in QCI and SEO, the intensity of fish was also at least six times higher in QCI at 0.062 fish/m^2 compared to 0.0056 – 0.010 fish/m^2 in SEO.

Sensitivity analyses showed that increasing the maximum separation distance (h_f) caused increases in the estimated mean number of fish per cluster and cluster radius, and decreases in estimated cluster intensity (Figure 11). Relationships between maximum separation distance and estimated cluster size could occur if, for example, clusters occur at multiple scales in rockfish populations. The Thomas process is only able to fit clusters of a single size rather than a complex pattern of clusters at different scales; therefore, changing the scale causes changes to the pattern observed. Another, more likely explanation is that most of the sites exhibit only marginal clustering. In the absence of strong clusters at other scales, least squares estimates based on the observed data may have occurred when the fitted spatial pattern consisted of one large cluster about the size of the maximum separation distance. The results of the sensitivity analysis on the maximum separation distance support this hypothesis because cluster radius is generally approximately half the maximum separation distance h_f. An exception to this finding occurs in the QCI where cluster radius is quite different from $h_f/2$ and remains small at all values of h_f. Therefore, the value of maximum separation distance may have affected parameter estimates where clustering was weak, but in sites with stronger clustering, such as QCI, the value of h_f appeared to be less important.

The Monte Carlo testing procedure suggested that the Poisson spatial process adequately characterized three of the submersible surveys (NSEO, CSEO, and EYKT; Figure 12a) and one of the ROV surveys (CHR; Figure 12b), while the clustered Thomas process better characterized two submersible sites (QCI and SSEO) and one ROV site (SJC). The empirical Ripley's *K*-function from the CHR site and the majority of the SEO sites fell within the range of *K*-functions expected from the Poisson process. The empirical *K*-functions from QCI and SJC were above the Poisson envelope for their entire range, indicating that there were more pairs of fish observed at small distances than one would expect from a random distribution. The difference between these *K*-functions and the Poisson envelope was greatest close to zero and then decreased, suggesting that small clusters may be the most dominant feature of the distribution (Figure 12). In contrast, the *K*-function from SSEO was within the Poisson envelope at a scale of 20 m or less, indicating that the distribution was not significantly different from random; however, at larger scales, rockfish were more clustered that would be expected by chance. These results support the hypothesis that some of the sites exhibit defined clustering at small (< 5 m) scales, while the remaining sites exhibit poorly defined clusters at larger (20 m or more) scales. Because clusters may exist at multiple scales and be poorly or well defined, either the Poisson process or the Thomas process could be used to characterize the distribution of rockfish from different sites.

Confidence envelope width for the Poisson process was sensitive to whether parameter values were drawn at random (as I show here) or fixed at their mean values across sites. Both Thomas and Poisson process confidence envelopes were wide because parameters were drawn at random from a distribution of values. Occasionally, low values of cluster intensity drawn by chance resulted in simulated sites where no fish occurred within the 1.58 m ROV transect or even within the entire study region. This explains why the lower confidence envelope was essentially zero. If, on the other hand, process parameters had been fixed at their mean value, the Poisson envelope would have been much narrower because most, if not all, simulated transects would have observed fish at most spatial scales. In contrast, the Thomas envelope would likely have remained wide because the probability is much higher that a single transect will have either many fish or no fish in a clustered distribution compared to a random distribution. Simulating longer transects would have allowed us to obtain more precise confidence envelopes that excluded zero; however, longer transects would have required simulating larger-scale spatial patterns than our data could support and would also have increased computation time.

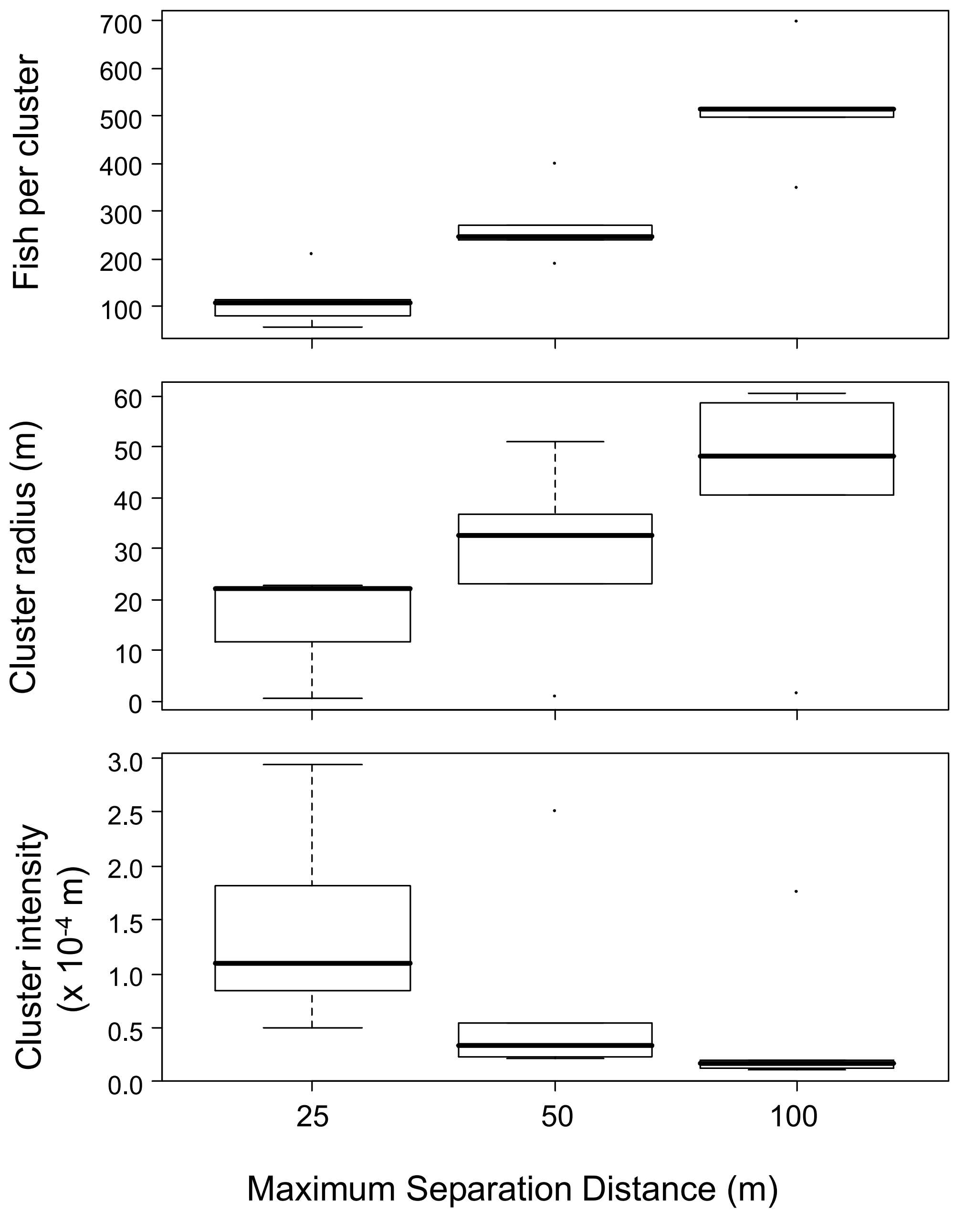

Figure 11 **The effect of the maximum separation distance used in the parameter estimation procedure (h_f) on Thomas process parameter estimates: fish per cluster (μ), cluster radius (ρ), and cluster intensity (λ). The baseline value used was h_f = 50. The boxplots show the distribution of parameter values for the five sites surveyed with manned submersibles.**

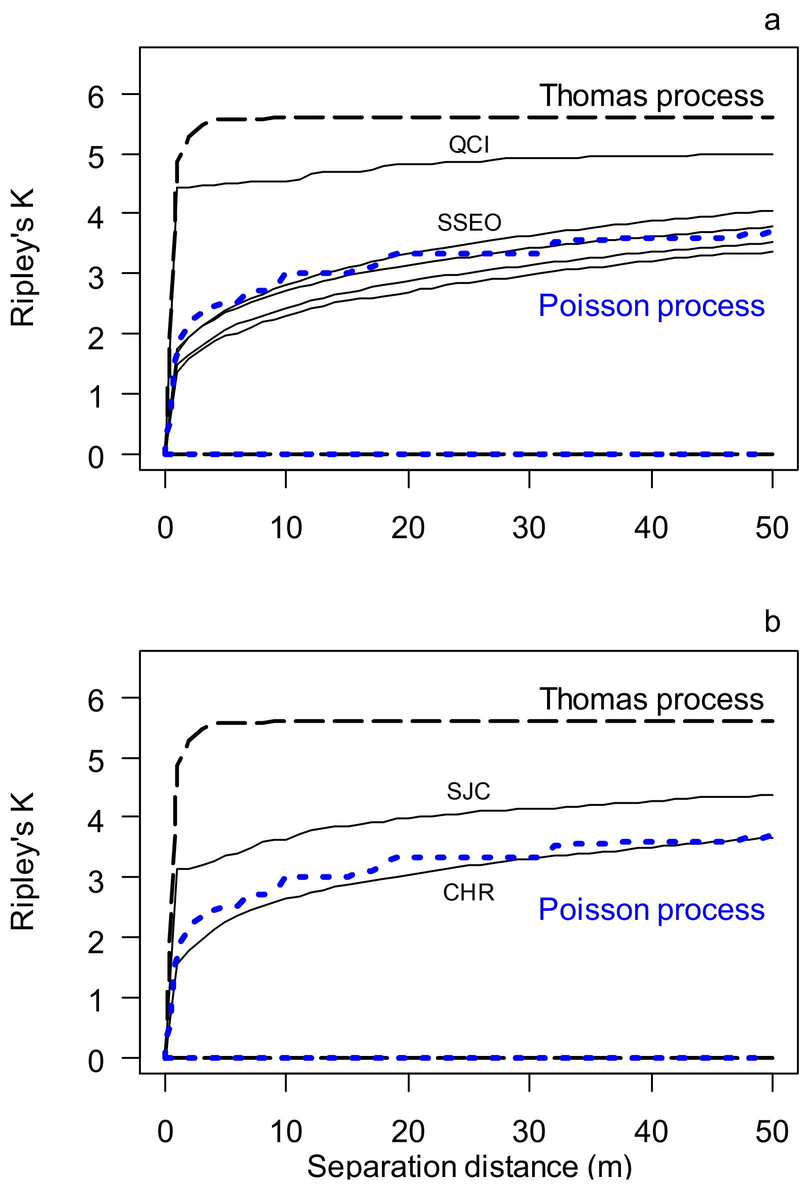

Figure 12 **The empirical *K*-functions for ROV data simulated from spatial models compared to the empirical *K*-functions from the actual underwater visual survey data: a. submersible data and b. ROV data. The empirical *K*-functions for the real survey data are shown by thin solid lines: Queen Charlotte Islands (QCI), Southern Southeast Outside (SSEO), San Juan Channel (SJC), and Chiswell Ridge (CHR). The submersible surveys for the remainder of the Southeast Outside sites are show by not labelled (NSEO, CSEO, and EYKT). The dashed and dotted lines represent the upper confidence envelope of the empirical *K*-functions for the simulated data from the Thomas process and Poisson process, respectively. The lower confidence envelope for both processes was a horizontal line at Ripley's *K* equal to 0.**