DEVELOPMENT OF SAMPLING REGIMES FOR ESTIMATING DENSITY OF RED SEA URCHINS
(*Strongylocentrotus franciscanus*)

by

Karen M. Skibo
B.Sc., University of Winnipeg, 2001

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

Report No. 388

SIMON FRASER UNIVERSITY

Fall 2005

APPROVAL

Name: **Karen M. Skibo**

Degree: **Master of Resource Management**

Title of Research Project: **Development of Sampling Regimes for Estimating Density of Red Sea Urchins (*Strongylocentrotus franciscanus*)**

Report No.: **388**

Examining Committee:

____

Dr. Randall M. Peterman
Senior Supervisor
Professor
School of Resource and Environmental Management
Simon Fraser University

____

Dr. Carl J. Schwarz
Committee Member
Professor
Department of Statistics and Actuarial Science
Simon Fraser University

Date Defended/Approved: Nov. 25/2005

ABSTRACT

Uncertainty in stock biomass estimates of marine invertebrates arises, in part, due to difficulties in obtaining accurate and precise density estimates. In aggregated populations, such as with red sea urchins (*Strongylocentrotus franciscanus*), a precise estimate of density may be quite challenging to obtain. I examined different survey designs using simulated urchin populations with the objective of improving density estimation. I evaluated alternative designs in terms of precision, cost, and bias. Designs under consideration were a simple version of random transect sampling, and more complex random transect sampling designs, including restricted adaptive cluster sampling and a design stratifying by substrate within the transect. The complex designs were more precise and cost-efficient than the current sampling method used in British Columbia but produced slightly biased estimates. Choice of sampling designs for use in field surveys, which are conducted by SCUBA divers, may be influenced by the ability to implement complex designs.

Keywords:
survey design, red sea urchin, sampling, adaptive cluster sampling, simulation

ACKNOWLEDGEMENTS

I thank my supervisory committee, Randall Peterman and Carl Schwarz, for providing their expertise and insightful advice. I would also like to thank the members of the Fisheries Research Group for all their suggestions in the development of this research project. For providing data, I thank Alan Campbell and Dimitri Tzotzos of Fisheries and Oceans Canada, and the Pacific Urchin Harvesters Association. Special thanks to Ryan Galbraith for his encouragement and the many hours he spent editing my initial drafts. Thanks to my friends in REM, in particular Jaclyn Cleary and Stacy Webb for their invaluable friendship and support.

Financial support for my research was provided by a C.D. Nelson Memorial Scholarship, an SFU Graduate Entrance Scholarship, a SFU Graduate Fellowship, as well as a scholarship from Coastal Zone Canada. In addition, I received a Research Assistantship from the National Science and Engineering Research Council through a grant to Randall Peterman.

TABLE OF CONTENTS

Approval ii
Abstract iii
Acknowledgements iv
Table of Contents v
List of Tables vii
List of Figures viii
Introduction 1
Methods 10
Basis for the Spatial Model 10
Simulation of a Population 13
Description of the Simulated Survey Designs 17
Simple Random Survey (SRS) 17
Random Transect Surveys (RTS) 18
Stratification of the Survey Area 23
Monte-Carlo Simulation 25
Cost Considerations 28
Results 30
Simulated "True" Populations 30
Comparison of Sampling Designs without Stratification of the Survey Area 31
Relative Efficiency of Designs 31
Relative Bias of Designs 32
Precision of Designs 33
Consideration of Costs 33
Comparison of the Sampling Designs with Stratification of the Survey Area 34
Relative Efficiency of Designs 34
Relative Bias of Designs 35
Precision of Designs 36
Consideration of Costs 36
Discussion 37
Comparison of Alternative Designs to Current Sampling Design 37
Sampling Effort 37
Relative Efficiency 37
Relative Bias 38
Precision 40
Consideration of Cost 41

Effect of Stratification 42
Other Research 43
Limitations 45
Conclusion 48

References 51

Tables 54

Figures 59

Appendix 76

LIST OF TABLES

Table 1. Results of a mixed model fit to the Laredo Channel 2000 survey data 54

Table 2. Transition matrix of probabilities of a given substrate type in quadrat *a+1* following a given substrate type in quadrat *a*, while moving from the shallow to the deep end of the transect. Based on the Laredo Channel 2000 survey data 55

Table 3. Transition matrices for each substrate type with three categories of counts for urchins: zero count, low count (1 - 10 urchins), and high count (>10 urchins). Based on the Laredo Channel 2000 survey data 56

Table 4. Mean number of urchins in the low- and high-count categories and the maximum urchin count values ("Max. val.") for the substrate types in the Laredo Channel 2000 survey data.. 57

Table 5. Relative sampling effort for each substrate type in the simulation based on the Laredo Channel 2000 survey data.. 58

LIST OF FIGURES

Figure 1. Map of Laredo Channel, British Columbia displaying the systematic layout of transects for the survey conducted in 2000 59

Figure 2. Mean number of urchins (± SE) per 1 m^2 quadrat for nine different substrate types (open bars) and associated substrate composition of the survey area (shaded bars). Based on Laredo Channel 2000 survey data. 60

Figure 3. Mean proportion of a transect (± SE) with quadrats occupied by red sea urchins for different ranges of transect length (open bars). Mean proportion of a transect (± SE) with quadrats composed of bedrock and/or boulders for different ranges of transect length (shaded bars). Based on Laredo Channel 2000 survey data. 61

Figure 4. An example of the bootstrapping procedure using a subset of the Laredo Channel 2000 survey data. 62

Figure 5. An example of the restricted adaptive cluster design (RTSacs) (not drawn to scale) 63

Figure 6. Flowchart outlining the basic procedure of the simulation model. 64

Figure 7. (a) Distribution of mean densities of the 200 realizations of "true" urchin populations. (b) Distribution of the coefficient of variation of the 200 realizations of "true" urchin populations 65

Figure 8. Distribution of the relative efficiency (RE) of each design without stratification of the survey area relative to DFO's current random transect survey (RTScurrent) for 200 urchin populations. 66

Figure 9. Distribution of the relative efficiency (RE) of each sampling design without stratification of the survey area relative to a simple random sampling design (SRS) for 200 urchin populations 67

Figure 10. Distribution of the relative bias of each design without stratification of the survey area for 200 urchin populations 68

Figure 11. Distribution of the standard error of the mean density for the sampling designs without stratification of the survey area for 200 urchin populations 69

Figure 12. Distribution of the inefficiency (marginal cost x variance) of each design without stratification of the survey area for 200 urchin populations 70

Figure 13. Distribution of the relative efficiency (RE) of each sampling design in a stratified survey area relative to DFO's current random transect survey

(RTScurrent without stratification of the survey area) for 200 urchin populations ([variance of RTScurrent]/[variance of estimator of interest]) 71

Figure 14. Distribution of the relative efficiency (RE) of each sampling design in a stratified survey area relative to simple random sampling (SRS) for 200 urchin populations 72

Figure 15. Distribution of the relative bias of the each design in a stratified survey area for 200 populations....73

Figure 16. Distribution of the standard error of the mean density for the sampling designs in a stratified survey area for 200 urchin populations....74

Figure 17. Distribution of the inefficiency (marginal cost x variance) of each sampling design in a stratified survey area for 200 urchin populations....75

INTRODUCTION

Stock assessment programs face the challenge of reliably estimating abundance of aquatic species. Some of the greatest challenges are with invertebrates, which, up until the last decade, have generally received less research attention than finfish. In new or emerging fisheries where management is required to follow a precautionary and risk-averse approach (Fisheries and Oceans Canada (DFO) 2003), biological and exploitation rate information collected via population sampling is vital to decision-making. In developing marine invertebrate fisheries, managers are often faced with making decisions when limited information is available on the biology and exploitation of the species in question (Perry et al. 1999). A precautionary and risk-averse management approach increases the probability of maintaining a sustainable fishery while managers and scientists improve their understanding of the population dynamics of target species.

The red sea urchin (*Strongylocentrotus franciscanus*) is one such species where the population dynamics are not well understood. Research suggests that population parameters such as recruitment and mortality vary greatly by location. Sloan et al. (1987) estimated recruitment of red sea urchins in British Columbia at 9.5% of the total number of sea urchins but noted that recruitment varied substantially in study sites within close proximity, as well as between sites sampled in a broad-scale survey. Density-dependent mechanisms or Allee effects (Allee 1931) potentially influence recruitment of urchins. The Allee effect is the positive feedback between population density and reproductive success of individuals that tends to drive populations toward extinction if density is below some threshold. Urchins are broadcast spawners and as a result require some minimum density for successful reproduction (Levitan et al. 1992). Aggregations of adults are also important for the protection of juveniles (Tegner and Dayton 1977). Juveniles will settle under the cover of adults' spines to avoid predation. Reported estimates of mortality rates for urchin populations vary greatly. For instance, instantaneous natural mortality rates in northern California were estimated to reach 0.21/year (Morgan et al. 2000), whereas Ebert et al. (1999) report study sites with much lower estimates (0.016/year) of mortality in the same region. Scientific understanding is better for the habitat preferences than for mortality rates. Red sea urchins feed on kelp and drift algae, and play an important role in kelp forest communities (Tegner 2001). Red sea urchins also have a preference for certain substrates and are commonly found on rocky substrates from the intertidal zone to depths of 20 m (Campbell et al. 1999a).

In British Columbia, DFO is responsible for maintaining a balance between healthy red sea urchin populations and harvest levels that provide reasonable economic returns to the urchin fishery (Campbell et al. 1999a). Accurately estimating biomass is an important factor in achieving a sustainable harvest. In the management of the red sea urchin fishery, fishing quotas are in part based on a modified surplus production model for which biomass estimates are a key input (Campbell et al. 1999b, 2001). The model calculates the maximum sustainable yield from a stock. Scientists acknowledge the uncertainty in the model by applying a correction factor to the quota estimate. The correction factor is a means of preventing overharvesting due to errors in the biomass estimate. The upper and lower 90% confidence intervals of the current biomass estimate are also used in additional model scenarios to provide a range of quota options for managers (Campbell et al. 1999b, 2001; DFO 2004).

There are two main approaches to estimating biomass of red sea urchins in British Columbia. One approach is to use fishery-dependent data, such as catch-per-unit-effort (CPUE), which are an index of abundance and can help estimate fishery trends. However, Campbell et al. (1999a) identify several reasons why this approach is unreliable for red sea urchins. First, CPUE can remain relatively unchanged if sequential depletion of populations is occurring across space. That is, fishermen may continually move from one fished area to an unfished one, preventing a significant reduction in CPUE until all the populations in a region have been overfished. Second, market demand can also affect CPUE. Fishermen may modify effort to meet a buyer's requests. Third, preferences of fishermen to harvest only in optimal areas to obtain high-quality roe may also influence the CPUE. Harvesting these areas subsequently allows red sea urchins occupying lower quality habitat to move to the more optimal areas. Consequently, the CPUE may remain unchanged in successive visits to a site and abundance indices can become distorted. The resulting lack of reliability of fishery-dependent data creates a necessity for high quality fishery-independent surveys that can provide precise estimates of density for use in harvest planning.

The other approach to estimating biomass of red sea urchins is fishery-independent population surveys, which is the primary stock assessment tool used by DFO. Two general types of fishery-independent surveys are currently used in British Columbia to estimate density of red sea urchins, "broadbrush" and "targeted" surveys. Both types of surveys are conducted by SCUBA divers. Few fishery-independent surveys had been carried out prior to the 1990s, even though a commercial fishery in British Columbia has been ongoing since the 1970s (Campbell et al. 1999a). Surveys within a given area may also be infrequent. DFO strives to survey Pacific Fishery Management sub-areas at least every five years or more frequently in locations that have experienced heavy fishing or depredation by otters (Campbell et al. 2001). Broadbrush surveys, first conducted in 1993 (Jamieson and Schwarz 1998), provide a general estimate of the density of red sea urchin for large areas such as a sub-area of a Pacific Fishery Management Area (Campbell et al. 1999a). An example of the size of a survey area is a survey of Laredo Channel in British Columbia, which covered approximately 111 km of total shore length. In a typical broadbrush survey, the location of the starting transect is randomly selected and the other transects are systematically placed relative to the first one. Surveys generally occur in water depths no greater than 15 m below chart datum (chart datum refers to lowest normal tide) which corresponds to the majority of fishing activity (Campbell et al. 1999a). Advantages of the broadbrush survey include its simplicity to plan and its flexibility for estimating density on a g/m^2 basis or g/m of shoreline basis (A. Campbell, personal communication, Fisheries and Oceans Canada, Nanaimo, B.C., 2004).

The other type of fishery-independent survey, targeted survey, began in 2003 and is similar in sampling design to the broadbrush survey but differs in that sampling effort is not evenly spaced across the survey area (A. Campbell, personal communication, 2004). In these surveys, transect placement is also perpendicular to the shoreline but 75% of the transects are concentrated randomly within harvest areas, and 25% outside of harvest areas. Surveys provide an estimate of red sea urchin density and size-frequency distribution, which are used to estimate biomass for the given area.

Uncertainty in biomass estimates arises from several sources (Campbell et al. 1999b, 2001). First, surveys in a given area may occur once every five years, so that past data may no longer reflect the current state. Density estimates are also not available for all sub-areas of a Pacific Fisheries Management Area (PFMA), so an estimate is based on an average estimate from the entire PFMA for all years surveyed. In PFMAs where no data is available from fishery-independent surveys, then an estimate of density is based on the data of an adjacent PFMA. Second, in addition to requiring the density estimate and size-frequency data, an estimate of the total urchin bed area is necessary to estimate the total biomass. Bed area refers to an area of fishable density of red sea urchins (e.g. $\geq$ 1 urchin per m^2). The estimate of urchin bed area, though, is highly uncertain. The accuracy of harvest logbooks used to determine the location and size of a bed area is unknown. Furthermore, biologists are unsure of how many years of harvest logbook data to include within a cumulative total of bed area because it is unknown how closely annual changes in harvest locations reflect changes in urchin bed area. Third, precise density estimation of spatially clustered species, such as urchins, can be difficult to obtain (Smith et al. 1995). My research will focus on reducing this third source of uncertainty by improving density estimation through a survey's sampling design.

Difficulties associated with surveying red sea urchin populations are typical of many species with an aggregated distribution. The aggregated distribution of red sea urchins results in large areas without any urchins and a few areas with many urchins, so traditional sampling regimes (e.g. simple random sampling) will yield a dataset with many zero-counts (Thompson and Seber 1996). For instance, in a 1994 DFO red sea urchin survey using a two-stage cluster design, urchins were present in only one-third of the total quadrats measured (Jamieson and Schwarz 1998). Such situations suggest adaptive cluster sampling (ACS) strategies should be used because they can improve precision of density estimates over conventional designs for rare and/or highly aggregated populations (Thompson 1990, 1991a, 1991b, 1992, 1996; Seber and Thompson 1994; Thompson and Seber 1996). ACS designs focus sampling effort in areas where non-zero-counts occur and still provide an unbiased estimate of density using either the Hansen-Hurwitz (HH) or the Horvitz-Thompson (HT) estimators. ACS essentially has two stages; the initial random selection of n sample units and the adaptive sample. If a unit within the initial random selection satisfies the density condition, then sampling occurs in the neighbourhood of the selected unit. A neighbourhood is commonly defined as the units adjacent to the selected unit; however, complex neighbourhood patterns are possible (Christman 1997; Lo et al. 1997). If a unit within the neighbourhood satisfies the condition, then the neighbourhood of that unit is also sampled and so forth, until no further units satisfy the condition. The result of the sampling is a contiguous group of sampled units, called a cluster, made up of edge units and a network. The network includes those units in the cluster that satisfy the density condition, whereas edge units refer to observations that fall short of the condition. Since the sampling of observational units continues until all units of a neighbourhood fail to satisfy the density condition, the final sample size is unknown prior to the completion of the survey. This is problematic because in most cases, time and budget constraints limit the sample size of a survey, potentially making ACS a less feasible option for field applications (Brown and Manly 1998; Conners and Schwager 2002; Salehi and Seber 2002; Smith et al. 2003; Su and Quinn 2003).

This problem can be dealt with by modifying adaptive cluster sampling to limit the total sampling effort. I illustrate this approach using several examples. First, Lo et al. (1997) limited sampling effort by using a restricted adaptive cluster sampling design that imposed neighbourhood restrictions in a survey of hake larval abundance. They restricted the adaptive sampling stage to a maximum number of units beyond the initial sampling unit. Similarly, Su and Quinn (2002) imposed a stopping rule restricting the number of adaptively sampled units in the sampling of simulated fish and bird populations. This sampling design biased the Hansen-Hurwitz (HH) and the Horvitz-Thompson (HT) estimators by using the stopping rule and further biased the HT estimator by using a design where the density condition value changed with each sample. Another variation of restricted adaptive cluster sampling design (Brown and Manly 1998) used a defined limit for the final sample size. A unit within the initial sample was selected and adaptive sampling took place (if the pre-defined condition was satisfied) before the next initial unit was randomly selected. As part of this restricted design, when the cumulative sample size met or surpassed the pre-specified limit on the final sample size, sampling ceased after the completion of the cluster even if all units within the expected initial sample had not yet been sampled. Sample size was not fixed, so the error in estimating the expected value of the initial sample size biased the density estimate. In other research, Woodby (1998) constrained sample size in adaptive cluster sampling surveys of simulated populations of red sea urchins by adaptively sampling only those units midway between secondary sampling units (i.e. quadrats). Thompson and Seber (1996) describe additional methods of limiting sampling effort in ACS.

My research focuses on improving density estimates of red sea urchins in British Columbian waters by comparing alternative designs, including ACS, with the current sampling method used by DFO. The purpose of this research is to identify potential designs that are both practical in the field and improve the precision of density estimates over what is achieved through current survey methods. Several other benefits will also result from this research. As previously mentioned, improving the precision will reduce the uncertainty of biomass estimates that are used in making harvest policy decisions, which should reduce the chance of both overharvesting and harvesting less than could be sustained. Second, cost and time savings could result from identifying designs that require less sampling effort than what is currently used while still providing a similar level of precision.

Little research has been published that investigates the use of different sampling designs for estimating density of red sea urchins. Currently, the only published simulation experiment assessing sampling designs for red sea urchins is Woodby (1998), which compares ACS to simple random sampling. I expand upon Woodby (1998) by using simulated urchin populations to assess the performance of other designs, in addition to adaptive cluster sampling. Specifically, I simulate simple random sampling and variations of random transect sampling designs, including a modified version of Woodby's (1998) ACS method of constraining sample size in red sea urchin populations. I primarily evaluate the designs relative to the current method by determining which one(s) provides the greatest precision and statistical efficiency, and minimize the bias in the density estimates of red sea urchins. As well, the cost of a design is often a primary consideration that managers cannot ignore because they are faced with budget constraints. Therefore, I also assess cost efficiency of each design. Simple random sampling is included in the simulation primarily as a means of determining a design's statistical efficiency, but it is not considered a realistic alternative to the current sampling method.

METHODS

To compare the potential field performance of various sampling designs for surveys of red sea urchins, I tested the designs on simulated populations of urchins that were generated using a spatial model. The spatial model simulated populations using data from the red sea urchin population in Laredo Channel in British Columbia (Pacific Fishery Management Area 6) (Figure 1). I bootstrapped the field survey data collected from that location to generate one spatial configuration ("simulated realization") of a survey area. The resulting simulated survey area represented the spatial configuration of red sea urchin distribution and associated habitat. For my comparison of sampling methods, I assumed that the simulated survey area was the "true" spatially heterogeneous distribution of red sea urchins. By simulating sampling designs in a survey area with a known density, I was able to compare the estimates of density for each design against that known "true" value. I then re-bootstrapped the field survey data to create 200 "true" populations to compare the sampling designs over a wide variety of potential spatial configurations. I programmed the simulation in Excel Visual Basic for Applications (2002). To generate uniform random deviates, I used an algorithm, function "ran3" (Press et al. 1992) that was independent from Excel's built-in RAND function.

Basis for the Spatial Model

The basis for the spatial model that I used to generate realizations of urchin populations was survey data collected by DFO to estimate red sea urchin density in the area of Laredo Channel in 2000 (Tzotzos et al. 2003). These data were collected via a broadbrush survey, which extended over 111 km of coastline. Using a systematic transect survey design, 86 transects were placed systematically along the coastline from a randomly selected starting point. Transects were laid out perpendicular to the coastline to a depth of 10 m below chart datum; thus, the slope of the ocean floor determined the actual length of each transect. Transects ranged in length from 17 m to 341 m and had a mean length of 53 m. Within a transect, SCUBA divers counted the number of urchins in every second quadrat (quadrat = 1 m^2) and collected information on urchin size, substrate, algae, and water depth. In total, 2303 quadrats were sampled. Density estimates among transects varied widely, ranging from zero to 19.8 urchins/m^2. The density estimate of the survey area in 2000 was 1.09 (standard error = 0.22) urchins/m^2. Further details of the survey are found in Tzotzos et al. (2003).

To parameterize the spatial model, I first analyzed the field data in several ways. Red sea urchins have a preference for rocky substrate and urchin distribution in the 2000 survey data show this strong association between substrate and density. In particular, substrates with bedrock and boulders have a much greater mean density of red sea urchins than substrates of cobble, gravel, sand, shells, and mud (Figure 2). A mixed effects model was fit to the survey data. The model in simplified syntax that represents the treatment, experimental, and randomization structure is:

*count = substrate transect(R) substrate*transect(R)* (1)

where count is the response variable, *R* indicates that an effect was random and the asterisk indicates interaction between the blocking factor (*transect*) and the treatment (*substrate*). Equation 1 indicates that the count (response variable) is affected by substrate and transect. I ran the model assuming a covariance structure of compound symmetry (i.e. the correlation between quadrats in a transect is the same) (Table 1). Results confirmed that there was a significant difference in counts of red sea urchins between types of substrate and this difference was consistent across transects. The variation in density among transects was small when the collection of substrates within the transects was the same. The large residual value from the model (Table 1) signifies high natural variation in the counts of urchins within a particular substrate of a transect. I fit the same model again but assumed a spatially auto-regressive covariance structure (where the (i,j) element of the covariance matrix between urchin counts in successive quadrats is $\sigma^2 \rho^{|i-j|}$). The results were similar to the model fit that assumed a covariance of compound symmetry and the correlation value between urchin counts in successive quadrats of a transect was low (0.18). In the survey area, sand substrate occurred most frequently (38%) and pea gravel occurring least frequently (0.1%) (Figure 2).

Another feature of the survey data is that only approximately 16% of quadrats had red sea urchins present. From a transect perspective, shorter transects had a higher proportion of quadrats occupied by urchins than the longer transects (Figure 3). This appears to be a function of substrate type. Shorter transects generally have a larger proportion of area composed of bedrock/boulders whereas longer transects have a greater proportion of the area with less urchin-desirable substrate types (e.g. sand, mud, cobble) (Figure 3). Bedrock and boulder substrates likely occur on sharper gradients than sand or mud substrates and as a result, transects composed mainly of bedrock and/or boulders are generally shorter in length because transect length is constrained by water depth. Using these general features of the population and survey results as basic characteristics, I simulated urchin populations and their associated habitat.

Simulation of a Population

Two hundred realizations of a red sea urchin population were simulated by bootstrapping the Laredo Channel transect survey data with replacement. I assumed the field survey to be representative of the substrate types and urchin numbers in the area. I also assumed that transects in the survey data were independent. The length of the transect and the proportion of each substrate within the transect were bootstrapped together (Figure 4). However, the order of substrates occurring along a transect, as well as the urchin counts within the quadrats, changed from the field survey data, as described below. Each simulated survey area was composed of 107 bootstrapped transects (primary units) to enable a range of sample sizes to be tested. Each transect was made up of quadrats (secondary units) 1 m^2 in size. Because lengths of transects varied, the total number of possible quadrats in the simulated survey area depended on which transects were randomly selected in the bootstrap procedure.

The order in which substrates occurred within a transect in the simulated survey area was based on two factors. The first reflected the observation that most transects in the field survey have bedrock occurring near the shallow end while substrates such as sand and mud are more likely to occur on the deep end of the transect. Either smooth bedrock or bedrock with crevices occurred in the first quadrat from the shallow end of the transect 93% of the time in the field data. Although depth was not explicitly considered in the model, the order of substrate types may be important, so the model was designed to give bedrock substrates a 93% probability of being selected as the shallowest quadrat of the transect if they actually occurred in the randomly selected transect.

The second factor affecting the order of substrate type along a transect in the model was the observation in the field survey data that quadrats of the same substrate type have a high probability of being adjacent to each other. Note that the order of quadrats and substrate types in the modelling process is discussed here in terms of moving along a transect from the shallow to the deep end. For example, moving from shallow to deep water, a sandy quadrat is more likely to follow a sandy quadrat within a transect than a quadrat of a differing substrate type. All substrates types had a high probability (≥ 0.74) of the next quadrat in the transect being of the same substrate type in the field survey data (highlighted diagonal of Table 2), except pea gravel. However, less than 1% of the quadrats in the field survey data consisted of pea gravel (Figure 2). The simulation used a transition matrix of probabilities (Table 2) to control the transition of substrate type between adjacent quadrats in a transect. For each new quadrat, a uniform random number was generated. Random numbers that were less than or equal to the transition probability resulted in the selection of a quadrat with the same substrate type as the previous quadrat, whereas a random number above the transition probability meant that all substrate types present in the transect had an equal chance of selection for that next quadrat. I used an equal probability of selection in this case instead of using the transition probabilities (non-highlighted values of Table 2) to determine the next substrate type in the transect because transition probability values between differing substrate types were generally much lower than those between the same substrate types. The exception to this was pea gravel. The observed probabilities for this substrate type were likely due to its rare occurrence in the Laredo Channel survey data. Furthermore, the number of substrates present in the transect was often small (approximately three on average) and the order of substrate type within a transect was already influenced by the weighted probability that bedrock substrate types would occur more frequently at the beginning of a transect. Bootstrapped transects did not change from the field survey transects in terms of length or the proportion of substrate types present. Once the proportion of a substrate was met in a transect, I determined the substrate type of the next quadrat using equal probability of selection (as described above) of the remaining substrate types that were not yet proportionally represented in the transect. I did not alter the length of the transects or substrate proportions within them because a wide range of lengths and substrate proportions were already present in the data. I therefore assumed the outcome of the simulation would not be affected.

Abundance values of urchins for each quadrat in the simulated survey area were then generated using the urchin count data from the Laredo Channel field survey. For the first quadrat of a simulated transect, the abundance value was randomly selected from the urchin counts in the field survey data associated with the given substrate type of the simulated quadrat. The count values for the remaining quadrats in a simulated transect were generated using transition matrices and Poisson distributions based on the field survey data. I produced a 3 x 3 transition matrix for each substrate type with three count categories: zero-count, low-count (1 - 10 urchins), and high-count (>10 urchins). Each matrix contained the probabilities of a quadrat of a given count category following another quadrat of a given count category (Table 3) as the surveyor moved from the shallow to the deep end of the transect. Using these transition matrices, I attempted to capture any patterns of clustering or aggregation of urchins within a transect that may exist in the field survey data. In the model, I used these probabilities to assign a count category to a quadrat based on the count of the previous quadrat. For instance, if the urchin count in quadrat *a* is 4 and it has a sandy substrate, then if quadrat *a*+1 is also composed of sand, it has a 0.67 probability of a zero-count, 0.33 probability of a low-count, and zero probability of a high-count (Table 3). For simplicity, the rule remained the same when the substrate type changed within a transect (e.g. sand to cobble). For example, if the urchin count in quadrat *a* is 4 and has a sandy substrate, then quadrat *a*+1, composed of cobble, still has a 0.67 probability of falling in the zero-count, 0.33 probability of a low-count, and zero probability of a high-count (Table 3). These transition probability values are associated with sand, not cobble. If quadrat *a*+2 is also composed of cobble, then its selected count category would then be based on transition probabilities associated with cobble.

The actual count value of a quadrat was randomly selected from a Poisson distribution based on these count categories. Poisson distributions were generated for the low-count category and the high-count category of each substrate using the observed mean count of each category from the field survey (Table 4). The largest count value that was possible for a quadrat in a low-count category was ten for any substrate, whereas the largest count value that was possible for a quadrat in a high-count category differed according to substrate type (Table 4). Substrates of gravel, pea gravel, and mud did not have any count values greater than zero in the field survey. Thus, quadrats in the simulated population composed of one of these three substrates types always had zero urchins.

Description of the Simulated Survey Designs

I examined the statistical performance and cost efficiency of five different sampling regimes to estimate density of the simulated urchin populations. Statistical performance was measured in terms of bias, relative efficiency, and precision. Cost efficiency was assessed using survey time and variance of the density estimate, where survey time was defined as the total time to count each of the sampled quadrats. The five different sampling regimes under evaluation were a simple random sampling design, a simple random transect sampling design, the current random transect sampling design (approximating DFO's current survey method), a within-transect stratification sampling design, and a restricted adaptive cluster sampling design. With the exception of the simple random sample design, each of the designs was a variation of random transect sampling.

Simple Random Survey (SRS)

In the simple random sample (SRS) survey, quadrats were randomly selected with replacement from the simulated survey area. The SRS estimator of the mean density of urchins is:

$$\hat{\mu}_{SRS} = \frac{1}{n}\sum_{i=1}^{n} y_i \tag{2}$$

where y_i denotes the observed density on the i^{th} quadrat and n is the number quadrats selected.

Random Transect Surveys (RTS)

In the random transect sampling (RTS) design, transects were randomly selected with replacement from the simulated survey area. Transects were perpendicular to the shoreline, as is currently the practice in the field. They were also of different length, so it was necessary to weight them in the calculation of the overall mean density of the survey area. Thus, for each of the RTS designs described below, the estimator of the mean density of the population is (Campbell et al. 1999b):

$$\hat{\mu}_{RTS} = \frac{\sum_{T=1}^{K} \hat{\mu}_T L_T}{\sum_{T=1}^{K} L_T} \qquad (3)$$

where L_T is the length of transect T in metres, K is the total number of transects in the sample, and $\hat{\mu}_T$ is the estimated density of transect T. The calculation of $\hat{\mu}_T$ is specific to each survey design and is described below for each variation of the random transect survey.

Simple Random Transect Sampling Design ("RTSsimple")

In the "RTSsimple" design, the number of urchins in all quadrats of a transect were counted. The estimated density of sea urchins for a given transect is:

$$\hat{\mu}_T = \frac{\sum_{i=1}^{n_T} y_i}{a_T} \qquad (4)$$

where n_T is the number of quadrats sampled in transect T, and a_T is the surface area of all quadrats surveyed on transect T. Note that $a_T = n_T$ because the area of a quadrat in the simulation was equal to 1 m^2.

Current Random Transect Sampling Design ("RTScurrent")

The "RTScurrent" design approximated the survey regime that DFO currently employs by counting urchins in every second quadrat of a randomly selected transect. In DFO surveys, the transects are placed systematically as opposed to randomly; however, I considered systematic and random placement of transects to be equivalent because it is unlikely that the systematic placement of transects matches any trends in the distribution of the urchin population (Jamieson and Schwarz 1998). The starting point for sampling within a transect was randomly selected to be quadrat 1 or 2. The estimated density of sea urchins for a given transect is calculated using Equation 4.

Within-Transect Stratification Sampling Design ("RTSstrat")

In the "RTSstrat" survey, the field survey data were used as prior knowledge to estimate the level of sampling effort necessary within each substrate type. Sampling effort within a randomly selected transect varied according to substrate type. This design may be useful to the biologist planning a survey if the area is not physically suitable to stratify into homogenous areas and/or data are scarce such that available information is inadequate to stratify. The estimated density of sea urchins for a given transect (Cochran 1977) is:

$$\hat{\mu}_T = \sum_{h=1}^{H} \hat{\mu}_h W_h \tag{5}$$

where $\hat{\mu}_h$ is the mean density of stratum h, H is the total number of strata, and W_h is the weighting factor calculated by (Cochran 1977):

$$\mathrm{W_h} = \frac{m_h}{L_T} \tag{6}$$

where m_h is the number of quadrats in stratum (substrate) h of the transect.

Using the field survey data, I determined the allocation of sampling effort for each substrate type using Neyman allocation (Neyman 1934). Allocation of effort was calculated by multiplying the proportion of total quadrats occupied with urchins by a given substrate type with the standard deviation of urchin counts in the same substrate type (Table 5). For the simulation, sampling effort in each stratum (substrate type) was made relative to sampling effort in the smooth bedrock substrate,

$$Effort = \frac{N_h s_h}{N_1 s_1} \tag{7}$$

where N_h is the proportion of stratum h in the survey area, s is the standard deviation of urchins per quadrat in stratum h, N_1 is the proportion of smooth bedrock substrate in the survey area, and s_1 is the standard deviation of urchins per quadrat in the smooth bedrock substrate. The result was that sampling effort in the simulation varied greatly between the strata (substrate types) within a transect (Table 5). Twice as much sampling occurred in quadrats of smooth bedrock than in quadrats of bedrock with crevices and one-fourth the sampling effort in quadrats of boulders as that of quadrats in smooth bedrock (Table 5). For some substrate types (gravel, pea gravel, shell, mud), the Neyman allocation scheme indicated that very little or no sampling effort was required (Table 5). To maintain some simplicity in the sampling design, quadrats of these latter four substrate types were sampled every 20^{th} time they occurred in a transect (Table 5).

Restricted Adaptive Cluster Sampling Design ("RTSacs")

In the "RTSacs" survey design, the initial sample of quadrats was selected systematically within a transect and the quadrat next to a sampled quadrat was adaptively sampled when a density condition was satisfied (Figure 5). The neighbourhood definition was the next quadrat adjacent to the observed quadrat along the transect. Adjacent quadrats preceding the initially observed one or areas outside of the transect were not sampled as part of the neighbourhood, resulting in an asymmetric neighbourhood pattern.

In the simulation, the initial sample was the systematic sampling of every third quadrat in a randomly chosen transect. The starting point for the initial sample within the transect was also randomly selected (possible starting points were quadrats 1, 2, and 3) (Figure 5). Each possible starting point for the initial sample represents a different possible systematic sample within the given transect, so there were three possible systematic samples (N) within a transect. Only one initial sample of the three possible initial samples was taken per randomly selected transect, so $n = 1$. The number of quadrats measured in the initial sample varied depending on the length of the transect (e.g. if transect length = 12 m and every third metre is measured then four quadrats are sampled as part of the initial sample). The criterion for sampling the neighbourhood of y_{ij} was based on the number of urchins per quadrat ($y_{ij} \geq c$ where y_{ij} is the number of urchins in the quadrat of interest and c is equal to some condition). A condition of $c = 1$ urchin per quadrat was used in all simulated scenarios because of the high frequency of zeros in the data. Thus, the detection of one urchin was treated as a possible cluster. A cluster is a contiguous group of sampled units consisting of edge units and a network. The network includes those units in the cluster that satisfy the density condition, whereas edge units are observations that do not satisfy the condition (Figure 5).

The restricted adaptive cluster sampling design (RTSacs) constrained the adaptive sampling within the randomly chosen transect as opposed to potentially sampling across the whole survey area. This constraint required the calculation of an estimator of mean density for each sampled transect. The Hansen-Hurwitz (HH) estimator and Horvitz-Thompson (HT) estimator (Thompson 1991) normally produce different, unbiased estimates for adaptive cluster designs. In this case of restricted adaptive cluster design, both produce the same estimate (Woodby 1998) (Appendix 1). However, that estimate is biased because of the asymmetric neighbourhood pattern used in the design. I will examine only the Hansen-Hurwitz (HH) estimator. Using the Hansen-Hurwitz (HH) estimator (Thompson 1991), the mean density of urchins in a randomly chosen transect is:

$$\hat{\mu}_{T_{HH}} = \frac{1}{n}\sum_{i=1}^{n} w_i \tag{8}$$

where w_i is the mean urchin density over all networks detected by the i^{th} transect and n = 1, as described above. The variable, w_i is calculated by:

$$w_i = \frac{1}{M}\sum_{k=1}^{K} \frac{y_k^*}{x_k} \tag{9}$$

where x_k is the number of possible samples that intersect the k^{th} network, y_k^* is the total number of urchins in all quadrats of network k, and M is the (L_T)/(step size of the systematic sample).

Recall that the total possible number of initial samples in a transect (N) was three because there were three possible starting quadrats (1, 2, or 3) from which to start sampling. Using Figure 5 as an example, the transect has two networks detected when sampling begins at quadrat 1. Network 1 intersects all three of the possible initial samples ($x_1 = 3$) but extends over four quadrats, so $y_1 = 8$. Network 2 intersects two of the possible initial samples ($x_2 = 2$) and y_2 totals to 4 urchins. Transect length (L_T) is 12 m long and the systematic step is three so, M = 4 as defined above.

In summary, five different sampling regimes were evaluated using simulated urchins populations. A simple random sampling design (SRS) selected quadrats randomly from the entire survey area. The four other sampling designs selected a random sample of transects from the entire survey area but the sub-sampling within each transect varied among the designs. Specifically, simple random transect sampling regime (RTSsimple) sampled all quadrats within a transect. The current transect sampling design (RTScurrent), which approximated DFO's current survey method, sampled every second quadrat in a transect. In the within-transect stratification regime (RTSstrat), sampling effort within a transect depended on substrate type. Finally, the restricted adaptive cluster sampling design (RTSacs), took an initial sample of every third quadrat in a transect and an adaptive sample when the density condition was satisfied.

Stratification of the Survey Area

To determine the extent to which the performance of each survey design (except SRS) would improve if stratification was used, I next stratified the simulated survey area based on survey type by dividing the simulated shoreline equally into 11 strata. In each stratum, I determined the ratio of quadrats with "good" substrate (smooth bedrock, bedrock with crevices, and boulders) to the total number of quadrats. Allocation of sampling effort among strata was proportional to the ratio of good-to-total substrate. As mentioned previously, good substrate was defined as those substrates for which urchins show a strong preference (bedrock and boulders). I assumed perfect information on the number of quadrats with good substrate and total number of quadrats in each stratum.

Although this is of course not realistic in the field situation, my analysis would indicate the maximum potential improvement possible by adding stratification to a survey design. The number of sampling units in a given stratum was determined by (Cochran 1977):

$$n_h = \frac{nP_h}{P} \tag{10}$$

where n is total sample size for the survey area, P_h is the proportion of good substrate relative to the total substrate in stratum h, and $P = P_1+P_2+\ldots+P_H$. The sum of the sample sizes across all strata should equal the total sample size for the survey area. In some cases, this did not occur, so I adjusted the sample sizes of the strata until their sum was equal to the total sample size using the following rules. When $\sum_1^H n_h > n$, then the sample size of the stratum with the lowest ratio of good-to-total substrate was reduced by one. If $\sum_1^H n_h$ was still greater than n, then the stratum with the next lowest ratio of good-to-total substrate was reduced in sample size. Conversely, when $\sum_1^H n_h < n$, then the sample size of the stratum with the greatest ratio of good-to-total substrate was increased such that $\sum_1^H n_h = n$. The estimate of the population mean ($\hat{\mu}_{st}$) when using a stratified sampling design was calculated using Equation 4. The weighting factor (W_h), in this case, is the number of possible transects in stratum h relative to the total number of possible transects in the entire survey area.

I applied each of the survey designs (except the simple random sampling survey) described in the previous section, "Description of the Simulated Survey Designs", to the stratified survey area, yielding a total of nine survey designs that I investigated.

Monte-Carlo Simulation

I conducted repeated Monte Carlo simulations of the sampling procedure (Figure 6) to ensure that results were not due to some chance selection of a small number of sampling units. The overall spatial model was parameterized based on Laredo Channel survey data, as discussed above. Using the spatial model, 200 realizations (simulated spatial configurations of sea urchin populations) were generated using the methods described earlier. For each realization, 1200 Monte Carlo trials were run for each sampling design scenario. This number of Monte Carlo trials was set after conducting several numerical experiments that increased the number of trials by 100 until the relative standard error of mean urchin density was different by $\leq 7\%$ from simulations with fewer trials. At least two successive incremental increases in the number of trials had to occur where the relative standard error fell within this criterion. The sample size for each design ranged from 5 to 45 transects with a step size of 10, except in the scenarios where the survey area was stratified. Only sample sizes from 25 to 45 transects were used in scenarios involving stratification of the survey area to enable the opportunity to sample in each of the 11 strata. For each Monte Carlo trial of the random transect (RTS) type surveys, I estimated the urchin density of the simulated survey area and then simulated a corresponding simple random sampling (SRS) survey using the equivalent number of sampled quadrats. Each survey design applied a different amount of sampling effort, yet equivalent sample sizes were required for a fair comparison of the designs to SRS. Thus, a corresponding SRS survey was simulated for each trial of a design using the equivalent sample size.

Calculations of the summary statistics for a given population were based on Su and Quinn (2003). Average mean density over a set of Monte Carlo trials is:

$$\bar{\mu} = \sum_{i}^{R} \frac{\hat{\mu}_i}{R} \tag{11}$$

where $\hat{\mu}_i$ is the i^{th} estimate of an estimator and R is the number of Monte Carlo trials. Absolute bias is:

$$B(\hat{\mu}) = \bar{\mu} - \mu \tag{12}$$

where μ is the true density of the population. Relative bias is:

$$RB(\hat{\mu}) = \frac{B(\hat{\mu})}{\mu} * 100 \tag{13}$$

Variance was estimated via Monte Carlo simulation, as opposed to using variance equations developed by statistical theory. The variance of the estimator is:

$$V(\hat{\mu}) = \sum_{i}^{R} \frac{(\hat{\mu} - \bar{\mu})^2}{R} \tag{14}$$

Relative efficiency is a means of comparing one estimator to another, given equivalent sample size. The relative efficiency of an estimator relative to simple random sampling (SRS), as defined by Woodby (1998), is the ratio of the variance of SRS to the variance of the estimator of interest:

$$RE(\hat{\mu}) = \frac{V(\hat{\mu}_{SRS})}{V(\hat{\mu})} \tag{15}$$

The density estimate ($\hat{\mu}$) for SRS and the design of interest was calculated using equivalent sample sizes. The total number of quadrats sampled is the equivalent SRS sample size. In the case of adaptive sampling, the total number of number of quadrats sampled is the sum of edge units and units that are part of a network. The equivalent sample size for SRS is the expected final sample size of a given design, which, based on the methods of Salehi and Seber (2002), is:

$$E(\nu) \approx \frac{1}{R}\sum_{i=1}^{R} \nu_i \tag{16}$$

where ν is the final sample size and R and i are as denoted above.

I provide an example for clarification. If I simulate three survey replications using the RTSacs design, then $R = 3$ and final sample sizes in secondary units (edge and network cells) are $\nu_1 = 1585$, $\nu_2 = 1560$, and $\nu_3 = 1591$. The expected final sample size of SRS is then the mean number of the replications, 1578.7. Instead of using the expected final sample size of SRS when simulating the SRS surveys, I used the final sample size for each Monte Carlo trial of the sampling design under consideration. The standard error of the mean urchin density for SRS using the expected final sample size, $E(\nu)$, was approximately equal (numerical tests show this was a good approximation) to the average of the standard error of the mean urchin density for SRS over all Monte Carlo trials. In other words,

$$\frac{\sigma}{\sqrt{E(\nu)}} \approx \frac{\sum_{i=1}^{R} \frac{\sigma_i}{\sqrt{\nu_i}}}{R} \tag{17}$$

Efficiency of the estimators was also determined relative to the RTScurrent design. In this case, I substituted $V(\hat{\mu}_{RTScurrent})$ for $V(\hat{\mu}_{SRS})$ in Equation 15.

Cost Considerations

Cost was incorporated into the model by determining the cost per unit of information (Snedecor and Cochran 1967; Swallow 1987; Kosmelj et al. 2001). In the case of random transect sampling survey designs for red sea urchins, base costs (e.g. equipment and fuel) and daily costs (e.g. salaries and travel expenses) for each of the designs are approximately the same because the primary sampling unit in all variations of the random transect sampling designs is the strip transect. Thus, I excluded base and daily costs and included only marginal (or variable) costs in the study. I did not consider the costs of simple random sampling because it is not a practical alternative sampling design. Marginal costs are those that are directly affected by the number of quadrats sampled and urchin density. However, determining the actual cost per quadrat is a complex task. I used sampling time as a proxy for cost because divers record the time it takes to complete the sampling of urchins for each quadrat along a transect. The average sampling time to complete a quadrat was generalized into three urchin count groups: zero-count time, low-count time (1-10 urchins/quadrat), and high-count time (>10 urchins/quadrat) using the Laredo Channel survey data and a second DFO survey completed in 2003 in the Dundas Island Group. For each urchin count group, I chose the larger sampling time value of the two data sets.

An optimal sampling design is the one that has the lowest cost per unit of information, where information is defined as the reciprocal of the variance of the mean (Snedecor and Cochran 1967; Swallow 1987; Kosmelj et al. 2001). Thus, inefficiency is the product of cost (or time in this case) and variance of the mean density:

$$IE = \frac{t1\sum c1 + t2\sum c2 + t3\sum c3}{(V(\hat{\mu}))^{-1}} \tag{18}$$

where $t1$ is the time it takes to sample a quadrat that contains no sea urchins (zero-count time = 52 seconds from the above data), $c1$ is the number of sampled quadrats with a zero-count, $t2$ is the time it takes to sample a low-count quadrat (low-count time = 90 seconds), $c2$ is the number of sampled quadrats with a low-count (1-10), $t3$ is the time it takes to sample a high-count quadrat (high-count time = 152 seconds), and $c3$ is the number of sampled quadrats with a high-count (>10).

RESULTS

Simulated "True" Populations

The 200 simulated realizations of "true" sea urchin populations resulted in a large range in the degree of urchin aggregation, a desirable situation for testing the performance of the sampling designs. The mean population density ranged widely from 0.74 to 1.79 urchins/quadrat (Figure 7a); however, the median value (1.11 urchins/quadrat) well-approximated the field survey mean of 1.09 urchins/quadrat. The distribution of the coefficient of variation (CV) for the 200 populations represents their range of aggregation (Figure 7b). Note that a CV = 1 means the standard deviation is equal to the mean. Of the 200 realizations, 160 populations had a CV < 2 and were classified as "weakly-aggregated", eight populations were "moderately-aggregated" with a CV ranging between 2 and 3, and 32 populations could be classified as "highly-aggregated" with a CV greater than three. A large gap in CV values occurred between the weakly-aggregated populations with a maximum CV value of 1.78 and the moderately-aggregated populations with a minimum CV value of 2.45, so I combined the small, moderately-aggregated group with the larger, highly-aggregated group to create two aggregation categories: weakly-aggregated with 160 populations and highly-aggregated with 40 populations.

Comparison of Sampling Designs without Stratification of the Survey Area

Relative Efficiency of Designs

The sampling designs in the scenarios without stratification of the survey area were generally equally efficient to, or more efficient than, DFO's current random transect sampling design (RTScurrent) (Figure 8). Recall, in this case, that relative efficiency (RE) is the ratio of the variance of RTScurrent to the variance of the estimator of interest. A RE value greater than one (dotted line) indicates that the alternative design was more efficient than RTScurrent. The simple random transect sampling design (RTSsimple) had a median efficiency relative to RTScurrent of 1.21 to 1.18 for sample sizes of 5 to 45 transects. Within-transect stratification design (RTSstrat) had the lowest relative efficiency values of the three designs and was roughly equal in efficiency to the RTScurrent design with values ranging from 1.01 to 0.95 (median values) for sample sizes of 5 to 45. The restricted adaptive cluster sampling design (RTSacs) performed similarly to RTSsimple with median efficiency values relative to RTScurrent ranging from 1.17 to 1.21 for sample sizes of 5 to 45 transects.

Relative efficiency (RE) of the designs in the scenarios without stratification of the survey area was lower when I compared the designs to simple random sampling (SRS) instead of RTScurrent (Figure 9). In this case, relative efficiency is the ratio of the variance of SRS to the variance of the estimator of interest. Most simulated scenarios were less efficient than SRS. Relative efficiency values less than one (dotted line) indicate that the design was less efficient than its set of corresponding SRS surveys. RTScurrent (median RE values ranged from 0.43 to 0.45 for sample sizes of 5 to 45 transects) fared better relative to its corresponding set of SRS surveys than RTSsimple

(RE = 0.25 to 0.26 for sample sizes 5 to 45) did relative to its set of SRS surveys. Within-transect stratification (RTSstrat) and restricted adaptive cluster sampling (RTSacs) designs performed very similarly relative to their respective sets of SRS surveys. The median relative efficiency of RTSstrat ranged from 0.57 (n = 5) to 0.62 (n = 45) and RTSacs ranged from 0.58 (n = 5) to 0.65 (n = 45). For some highly-aggregated populations, RTSacs and RTSstrat had relative efficiency values greater than one, indicating that they were more efficient than SRS in those scenarios. For many others that were categorized as highly-aggregated, efficiency was just below one. The rank order of the designs remained the same for each sample size because sample size has no effect on relative efficiency and any differences were an artefact of the simulation.

Relative Bias of Designs

The bias ([absolute bias/true population density] x 100) of RTScurrent, RTSsimple, and RTSstrat decreased with sample size, indicating that bias in the density estimates from smaller sample sizes was due to the ratio estimator used to estimate density (Figure 10). Overall, when sample size was sufficient to eliminate the bias from the ratio estimator, the RTScurrent and RTSsimple designs without stratification in the survey area were unbiased (Figure 10, e.g. n = 45 transects). The distribution of relative bias over the 200 simulated populations was the narrowest for RTScurrent and RTSsimple. The RTScurrent design had a median bias of 14.3% for n = 5 and declined to 1.8% for n = 45. RTSsimple had a bias similar to RTScurrent with a median value of 14.1% for n = 5 and declined to 1.7% for n = 45. The bias in the estimates of RTSstrat were the largest of the four designs, declining from 22.1% (n = 5) to 9.7% (n = 45). This design also produced the widest distribution for each sample size. Density estimates for

RTSacs were least biased at n = 5 (1.7%) but increased in bias to approximately -8% for n = 25, 35, and 45 transects. When I switched the direction of sampling in the simulation so that sampling began at the deep end of the transect instead of the shallow end of the transect, I observed that the bias of the RTSacs design was positive and it approached zero as sample size increased. These results are discussed further in the "Discussion".

Precision of Designs

Within each sample size, the distribution of the standard error of the mean density (precision) for the 200 simulated urchin populations was very similar across sampling designs in scenarios without stratification in the survey area (Figure 11). The median standard error of the RTScurrent design decreased from 0.89 for n = 5 transects to 0.27 for n = 45 transects (i.e. precision increased with sample size). RTSsimple improved in precision with median standard error values decreasing from 0.80 (n = 5) to 0.24 (n = 45) and median values of RTSstrat design went from 0.88 (n = 5) to 0.27 (n = 45). The standard error for RTSacs had a similar range as RTSsimple, decreasing from 0.80 (n = 5) to 0.24 (n = 45).

Consideration of Costs

The distribution of the inefficiency (marginal cost x the variance of the density estimate) of the four sampling designs without stratification of the survey area indicates the marginal cost per unit of information that each design incurred sampling the 200 simulated urchin populations (Figure 12). The marginal cost per unit of information when using the RTScurrent design (median values) ranged from 0.84 (n = 5) to 0.69 (n = 45) and the values for RTSsimple ranged from 1.39 (n = 5) to 1.16 (n = 45). The marginal cost per unit of information when using the RTSstrat design was less than RTScurrent and RTS simple, ranging from 0.63 (n = 5) to 0.51 (n = 45). RTSacs performed similar to RTSstrat with values ranging from 0.63 (n = 5) to 0.50 (n = 45). The rank order of designs remained the same for each sample size because sample size has no effect on inefficiency. Thus, RTSstrat and RTSacs designs had the lowest marginal cost per unit of information and RTSsimple regime had the highest cost per unit of information.

Comparison of the Sampling Designs with Stratification of the Survey Area

Relative Efficiency of Designs

In sampling scenarios with stratification of the survey area, each of the sampling designs were evaluated in a survey area that was divided into 11 strata, as described in "Stratification of the Survey Area" of the "Methods" section, whereas scenarios without stratification of the survey area did not divide the survey area into strata.

Efficiency of the designs in a stratified survey area relative to DFO's current random transect design (RTScurrent without stratification of the survey area) was generally less than one (Figure 13), indicating that the RTScurrent design was more efficient than the alternative designs for scenarios in a stratified survey area. Recall, in this case, that relative efficiency (RE) is the ratio of the variance of RTScurrent to the variance of the estimator of interest. Note that only sample sizes of 25 to 45 transects were tested in scenarios with stratification of the survey area to enable the opportunity to sample in each of the 11 strata.

The rank order of the sampling designs in terms of efficiency relative to SRS (i.e. the ratio of the variance of simple random sampling to the variance of the estimator of interest) did not change when the survey area was stratified (Figure 14) compared to when the survey area was not stratified (Figure 9). Relative efficiency values less than one (dotted line) indicate that the design was less efficient than its corresponding set of SRS surveys of equivalent sample size. Relative efficiency is not dependent on sample size, so I report the relative efficiency values for n = 45, which are similar to the values of the other sample sizes evaluated (Figure 14). RTScurrent and RTSsimple had a median relative efficiency of 0.34 and 0.22, respectively. RTSstrat and RTSacs designs had higher (best) median efficiency values of 0.48 and 0.47, respectively. In general, the relative efficiency was lower for all designs when the survey area was stratified compared to no stratification. As a result, fewer scenarios occurred where RTSacs and RTSstrat had a relative efficiency greater than one. These results are contrary to expectation and are discussed further in the "Discussion".

Relative Bias of Designs

When the survey area was stratified, the relative bias of all sampling designs decreased with increasing sample size (Figure 15), but only by small amounts. For the most part, the designs performed poorly, with much larger relative biases and wider distributions compared to their counterparts without stratification (Figure 10). The stratified versions of RTScurrent and RTSsimple had the same amount of bias decreasing from 28% to 21% (median values) for sample sizes of 25 to 45 transects. Median relative bias of RTSstrat decreased from 34% (n = 25) to 29% (n = 45) and RTSacs decreased from 13% (n = 25) to 8% (n = 45). The large bias in all designs is contrary to expectation and is discussed further in the "Discussion".

Precision of Designs

Within each of the sample sizes, the standard error was approximately the same between sampling designs in a stratified survey area, suggesting that the designs had similar levels of precision (Figure 16). The median standard error of RTScurrent declined from 0.43 (n = 25) to 0.31(n = 45) and for RTSsimple the standard error decreased from 0.39 (n = 25) to 0.28 (n = 45). Similarly, the median standard error of RTSstrat design declined from 0.42 to 0.31 and the standard error of RTSacs ranged from 0.40 to 0.29 for sample sizes of 25 to 45. For all designs, the standard error of the mean was slightly larger compared to when the survey area was not stratified (Figure 11).

Consideration of Costs

The distribution of the inefficiency (marginal cost x the variance of the estimator of interest) of the designs in a stratified survey area indicates the marginal cost per unit of information that each design incurred sampling the 200 simulated urchin populations (Figure 17). Inefficiency is independent of sample size, so I report the inefficiency values for n = 45, which are similar to the inefficiency values of the other sample sizes evaluated (Figure 17). The RTScurrent design had a median marginal cost per unit of information of 0.89 and RTSsimple had a value of 1.41. RTSstrat and RTSacs were the designs in the scenarios of a stratified survey area with the lowest (median) marginal cost per unit of information (0.70 and 0.69, respectively) (Figure 17).

DISCUSSION

Comparison of Alternative Designs to Current Sampling Design

Sampling Effort

The sampling designs varied considerably in the number of quadrats measured. Compared to DFO's current random transect sampling design (RTScurrent), restricted adaptive cluster sampling (RTSacs) surveyed 18% fewer quadrats, the within-transect stratification (RTSstrat) regime sample size was smaller by 33%, and simple random transect sampling (RTSsimple), as expected, was double the number of quadrats.

The neighbourhood restrictions in the RTSacs design constrained the increase in sampled quadrats from the initial to final sample size. The change from initial to final sample size for RTSacs could not exceed three times the initial sample size because the initial sample included every third quadrat in a transect and the neighbourhood pattern restricted adaptive sampling to within the transect. The final sample size increased on average by 24% of the initial sample. This value was consistent across all sample sizes ($n = 5$ to 45 transects, step size = 10).

Relative Efficiency

The RTScurrent design did not perform well in the first performance measure, relative efficiency. The RTScurrent design ranked third of four survey methods when the efficiency of each design was measured relative to its corresponding set of simple random sample (SRS) surveys (Figure 9). Recall, in this case, that relative efficiency is the ratio of the variance of SRS to the variance of the estimator of interest. RTSacs and RTSstrat both had a higher relative efficiency than RTScurrent; however, neither had a relative efficiency greater than one on average across all 200 simulated populations. In other words, the respective corresponding SRS surveys of RTSacs and RTSstrat were more efficient for most of the simulated populations. Restricted adaptive cluster sampling and the within-transect random transect design were at their most efficient in the highly-aggregated urchin populations because both designs placed more sampling effort in areas of higher urchin density, which have the greatest variability in urchin counts.

When the designs are compared directly to DFO's current design (RTScurrent), RTSsimple and RTSacs had a relative efficiency above one and RTSstrat had an efficiency of about one (Figure 8). Relative efficiency, in this case, is the ratio of the variance of RTScurrent to the variance of the estimator of interest. However, in the scenarios of the RTSsimple design, twice the amount of sampling effort was used as RTScurrent to achieve a relative efficiency greater than one, while RTSacs actually sampled 18% fewer quadrats than RTScurrent to achieve a high efficiency (i.e. RE > 1). Similarly, RTSstrat performed about the same as RTScurrent (i.e. RE $\approx$ 1) but sampled 33% fewer quadrats. Both the restricted adaptive cluster sampling design and the within-transect random transect sampling design performed better than DFO's current design when one accounts for the differences in sampling effort.

Relative Bias

For the second performance measure, relative bias, the RTSstrat design had a wide distribution and a fairly substantial positive bias occurred in some realizations (Figure 10). The tail behaviour of a distribution for a sampling design is important for cases in which estimation of stock size influences management decisions (Conners and Schwager 2002). This is the case with red sea urchins, where biomass estimates are a factor determining the harvest quota. Thus, the occurrence of large overestimates in density produced by RTSstrat is undesirable. The positive bias in the density estimates of RTSstrat was likely a result of the ratio estimator that was used to estimate the density within each substrate (or stratum) of the transect. The bias of the ratio estimator is of the order $1/\sqrt{n}$, becoming negligible at a large sample size (Cochran 1977). It is likely that within some stratum, the sample size was not large enough for the bias to become negligible. Thus, the RTSsimple, RTScurrent, and RTSacs designs may be preferable from a long-term biological conservation standpoint because they had a narrower distribution that was not positively biased, meaning a reduced probability of large overestimates. From the standpoint of the fishing industry, a large positive bias in the density estimates is undesirable for the long-term sustainability of the fishery and a negative bias in the density may result in a fishing quota lower than what an unbiased estimate would provide.

The restricted adaptive cluster sampling design (RTSacs) also produced biased estimates of density. The asymmetric neighbourhood pattern constrained adaptive sampling to the next quadrat from the quadrat of interest moving from shallow to deep along the transect, whereas a symmetric pattern would adaptively sample the next and previous quadrat from the quadrat of interest. This asymmetric neighbourhood pattern biased the Hansen-Hurwitz (HH) and Horvitz-Thompson (HT) estimators. Bias of the estimate also changed direction based on the direction of sampling (shallow vs. deep end as starting point). This directional change in bias was due to the asymmetry in the neighbourhood pattern, and was exacerbated by the spatial model, which had a realistic, but weighted probability of urchin-preferred substrate occurring at the shallow end of a transect. Thus, sampling quadrats from shallow to deep resulted in networks averaging out counts of urchins lower than expected, yielding an underestimate of the density of the transect (negative bias). The network averages were more likely to decrease as sampling moved to deeper water because quadrats were less likely to have favourable substrate and thus, fewer urchins. For instance, if one started sampling with the first quadrat in a transect and formed a network, the network would likely average out to a lower value as the next quadrats along the transect were sampled because they are more likely to have fewer urchins. If the starting point for sampling was the third shallowest quadrat in the transect, then the first two quadrats, which are generally more likely to be occupied by urchins than quadrats deeper along the transect, have zero probability of being sampled. The network, in this case, would average out to an even lower value than if the starting point was at quadrat 1. Thus, the negative bias in the estimate was a result of the networks averaging lower values than the true density and generally not averaging higher values. Conversely, sampling from deep to shallow resulted in networks averaging out higher than the true density more often than expected, yielding an overestimate of the density of the transect (positive bias). Networks average out higher than expected because quadrats at the shallow end of the transect have a greater probability of being included in a network regardless of the starting point at the deep end.

Precision

The third performance measure, precision, was measured using standard error of the mean density. In general, the designs had a similar level of precision (Figure 11) for a given sample size, despite the different amount of sampling effort used in each design. RTScurrent, RTSstrat, and RTSacs measured fewer quadrats in the transect than RTSsimple, which measured all quadrats in the transect, and they all achieved a similar level of precision in the density estimate of the survey area for a given sample size. Therefore, each design measured a sufficient number of quadrats within a transect to result in a negligible sampling error within the transect. It is the number of transects, though, that influenced the overall precision in the density estimates of the survey area. Precision of the density estimates was poor when only a small number of transects were measured, but it improved greatly as sample size increased. Thus, when transect-to-transect variation is high, such as in this simulation study, the greatest gains in precision are achieved by increasing the number of transects to sample.

Consideration of Cost

The fourth performance measure, marginal cost, was assessed using marginal cost per unit of information (marginal cost x variance of density estimate) or inefficiency. The RTScurrent design had the second highest marginal cost per unit of information as a result of its relatively large variance of the mean density and the large number of quadrats sampled (Figure 12). RTSstrat and RTSacs both had a lower marginal cost per unit of information. However, ranking the inefficiency of the designs based on marginal costs may be irrelevant because by definition, only the marginal costs (and not the larger base or fixed costs) of a survey are affected by the different survey designs. The total cost of a survey of 10 days is generally about $15,000 which includes base costs, such as boat and fuel, and daily costs, such as travel expenses and the salary of the biologist and divers

(Mike Featherstone, Pacific Urchin Harvesters Association, New Westminster, B.C., 2005). Base costs are essentially fixed, but if the reduction in quadrats was large enough to increase time savings by a day, then cost savings could potentially add up from reduced spending on travel expenses and salaries. A pilot study would provide an indication of just how much time could be saved by switching to either an RTSstrat or RTSacs design from the current DFO design. In any case, the results of the simulation suggest that neither of these designs would increase the amount of time required to complete a survey.

In summary, the restricted adaptive cluster sampling design (RTSacs) and within-transect stratification (RTSstrat) outperformed the other designs (with the exception of simple random sampling) in efficiency, precision, and marginal cost per unit of information. In practice though, simple random sampling is not a realistic option because it would incur other costs such as greater fuel usage, which would result from travelling between the many randomly selected quadrats. Furthermore, the simple random sampling design is not logistically possible due to SCUBA diving constraints. Divers are limited in the number of dives that they can do in a single day, which would then increase the total number of days required to complete a survey.

Effect of Stratification

Performance of all designs was worse when the survey area was stratified than when it was not. The poor performance, in particular the large positive bias in the density estimates of urchins, was likely due to a combination of the small sample size in each stratum and use of a ratio estimator to estimate density. A small sample size in each stratum resulted from dividing the survey area into many strata in an attempt to make the areas as homogenous as possible, while using the same total sample size (n = 25, 35, 45) as in the non-stratified designs. This, in combination with using a ratio estimator to calculate the density of each stratum, caused an increase in the bias from the non-stratified design. As described earlier, the bias of the ratio estimator is of the order $1/\sqrt{n}$, becoming negligible at a large sample size (Cochran 1977). However, sample size in the strata never got large enough for this bias to become negligible. The small sample size within each stratum also increased the variance of the density estimate. The survey area required many strata because of the low level of substrate homogeneity between simulated transects. Simulated transects were bootstrapped from the transects in the Laredo Channel field survey data, which were independent from each other, so two adjacent transects would not necessarily share similar characteristics of substrate type and urchin abundance. In an actual field survey, stratification may be more successful if the strata are fairly homogenous and the sample size within each stratum was larger than used here in the simulation. However, less-than-perfect information would be available in a field survey, necessitating decision-making based on previous survey data and scientific expertise to determine the location of strata. This may or may not be feasible in most situations.

Other Research

Previous simulation studies that have tested adaptive cluster sampling (ACS) have generally compared it to simple random sampling (SRS). Su and Quinn (2002) found the performance of adaptive cluster sampling using a restricted design was dependent on sample size, stopping rule, and degree of aggregation in the population. In particular, ACS was generally less efficient than SRS for populations that the authors defined as low-aggregated (CV = 1.92), while ACS designs in the intermediate- and highly-aggregated populations were more efficient than SRS. My results correspond to some degree with these findings; the efficiency of my restricted adaptive sampling design was less than one for weakly-aggregated populations and more than one (i.e. more efficient than SRS) for some of the most highly-aggregated populations (not all populations categorized as highly-aggregated had a relative efficiency greater than one). Woodby (1998), who applied a similar restricted adaptive design to what I used, also had comparable efficiencies to mine for highly-aggregated populations with values averaging somewhat less than one. Woodby's (1998) results differed in that his design achieved a relative efficiency approximating one for spatially random populations, whereas my results for similar populations were much less than one. The difference in my results from Woodby's may be related to the modifications I made in the restricted adaptive cluster sampling design and/or differences in the characteristics of the simulated populations.

Among the few published attempts at using adaptive cluster sampling (ACS) in biological field applications, success of the design has been mixed. In a study of freshwater mussels, ACS was less efficient than simple random sampling (SRS) but increased the detection of uncommon species (Smith et al. 2003). Adaptive sampling designs were less precise and had a larger bias in their estimates compared to traditional sampling regimes in a survey of benthic invertebrates (Cabral and Murta 2004). Sufficient survey time has also been problematic in field trials when a restricted design was not in place to control final sample size. Adaptive sampling in a hydroacoustic survey for Lake Erie smelt was not completed because the researchers ran out of vessel time (Conners and Schwager 2002). A restricted adaptive cluster design proved to be more precise and time efficient than SRS for estimating density of Pacific Ocean perch, which are considered highly clustered (Hanselman et al. 2003). For populations of shortraker and rougheye rockfish, which are more uniformly dispersed, the same authors found ACS was less precise and time efficient than SRS. Lo et al. (1997) found that using restricted adaptive sampling in a simple stratified design was more efficient than their proportional stratified sampling but they did not do a direct comparison with SRS. To my knowledge, a field application of adaptive sampling for red sea urchin populations has not yet been attempted.

Limitations

The degree of aggregation of red sea urchins on the BC coast varies widely, as exemplified by the wide range of urchin densities even in the transects of the Laredo Channel 2000 survey data. The simulated realizations of urchin populations characterized aggregations on a transect level. Aggregation across adjacent transects in the simulation was not necessary because all designs tested here were strip transects. As mentioned above, this method of simulating urchin populations likely affected the performance of the designs in the simulated stratified survey area. Using independent transects created a very heterogeneous survey area which I attempted to compensate for by creating smaller strata to increase the homogeneity of each; however, this resulted in a small sample size within each stratum. Another limitation of my results is that they are based on simulated populations which were parameterized by data from a single urchin survey. However, this limitation was somewhat compensated for by generating 200 realizations of urchin populations that varied widely in their degree of aggregation (Figure 7b), which may represent, in part, the diversity of spatial configurations that red sea urchins populations form. Future research should look at survey data from other locations.

In the simulation, restricted adaptive cluster sampling (RTSacs) and within-transect stratification (RTSstrat) design performed better than DFO's current random transect sampling (RTScurrent) design in terms of efficiency, precision, and marginal cost per unit of information. However, it is unclear how realistic it is to use these designs in field surveys of red sea urchins. A potential problem with both RTSacs and RTSstrat designs is that they may be so complex that divers may make frequent errors while sampling underwater.

RTSacs, which is likely the simpler of the two designs, creates one major additional complexity over the current DFO sampling design. The divers would need to know when to sample adaptively and which quadrats to include as part of the initial sampling stage. For instance in the simulation of RTSacs, every 3rd quadrat in a randomly selected transect was included as part of the initial sample. A second issue that could deter its use in the field is the design's sensitivity to choices that affect its efficiency, such as the neighbourhood pattern, density condition for sampling adaptively, and quadrat size (Christman 1997). In the case of surveying red sea urchins, the neighbourhood pattern and the density condition are the most relevant concerns. A restricted neighbourhood pattern is necessary because it is a dive survey and the amount of time spent underwater needs to be controlled both for the safety of the diver and cost. This type of restricted design would reduce the uncertainty in time required to complete the transect because divers know the approximate length of the transect before going underwater. Expanding the neighbourhood definition to include adjacent areas beyond the bounds of the transect would substantially increase the uncertainty of the finishing time for a transect. The neighbourhood pattern proposed here is fairly logical and would be a manageable task for divers. It enables divers to move forward along a transect and never swim back towards the beginning or beyond the transect's bounds. This pattern did result in a biased estimate in the simulations. In a field survey estimate of density, the amount of bias can be estimated using simulation (Hanselman et al. 2003) and then the density estimate can be adjusted accordingly.

Setting an appropriate density condition may be the largest barrier in the use of adaptive sampling for red sea urchins. In most areas of a survey, the density of urchins is fairly low, so a criterion value equal to one urchin would likely be suitable. However, in areas where the density is high for a large portion of the transect or for numerous transects, the number of quadrats measured could increase dramatically and increase the survey time considerably due to a large adaptive sample (Salehi and Seber 1997). Setting a large critical value can increase the precision of an estimate (Brown 1996 *in* Lo et al. 1997), but a critical value that is set too high will result in a survey consisting of all edge units (Salehi and Seber 1997). Lo et al. (1997) address this problem by stratifying the survey area into areas with a high density condition and areas with a low density condition. Using a stratified design also enabled those authors to revise the density condition midway through their survey.

Within-transect stratification (RTSstrat) design may also not perform as well in the field as the simulation results suggest because of complexities in the design. First, the design is more technically complex than RTScurrent and potentially more complicated than RTSacs. Divers may have difficulty in keeping track of the differing levels of sampling effort required for each substrate within a transect. I simulated the ideal sampling strategy based on optimal allocation of effort (Table 5); however, a simpler strategy could make this design a more practical option. For instance, quadrats with any substrate that is not bedrock or boulders (i.e. cobble, gravel, pea gravel, sand, shells, and mud) could be sampled at every 20^{th} quadrat, while bedrock and boulder substrates could be sampled according to the optimal allocation. A second issue of the design that may cause poorer performance in the field than the simulation is the necessity of prior information. The planning of a RTSstrat survey requires previous survey information to estimate the level of sampling effort required for each substrate type. Unfortunately, in the case of red sea urchins on the British Columbia coast, the time between surveys in a given area may be several years. The urchin biologist planning a survey would be forced to assume that little or no change has occurred in the area since the last urchin survey, a gap that could be as long as five years (Campbell et al. 2001). Thus, the quality of the survey will depend on the accuracy of the prior information relative to the current state.

Conclusion

A number of factors are responsible for the uncertainty in estimating biomass of red sea urchins, including the estimates in total bed area, the infrequency of population surveys, and the patchy distribution of urchins, which affects the quality of survey density estimates. This research focused on this third source of uncertainty by testing different sampling designs to determine whether the uncertainty of density estimates could be reduced, as well as whether a more cost-effective design is available as an alternative to the current survey method used by DFO. The research suggests that the within-transect stratification design and a restricted adaptive cluster design offer improvement over the current DFO survey design in terms of efficiency, precision, and marginal cost. Nevertheless, the complexity of the within-transect stratification design may be a barrier to its use in field applications. A more practical approach could be to use simpler versions of this design. With the restricted adaptive cluster design, factors such as the potential difficulty in selecting an appropriate density condition upon which to base adaptive sampling may deter its use. The results of the simulation indicate that the designs had a similar levels of precision because sampling error of all the designs was negligible within the transect. The key to improving precision of estimates then, is to increase the number of sampled transects because variability among transects is high. Time savings gained by sampling fewer quadrats within a transect could be used to sample additional transects. Alternative sampling methods such as towed video camera surveys could potentially provide significant time savings over dive surveys.

Future research should investigate the robustness of these results by evaluating the designs using field testing and also further simulations. In the latter case, the survey designs may perform differently relative to each other when tested on simulated populations based on another spatial model of the underlying spatial distribution of sea urchins. Using field surveys from additional locations besides Laredo Channel will likely provide more insight about the spatial configuration of substrates and urchins. The results of this study can also serve as a starting point from which to test the restricted adaptive cluster sampling design, the within-transect stratification design and other potentially time-efficient methods using pilot surveys. Pilot surveys can assess the viability of the designs in the field and provide feedback for modifications to the design. Modifications suggested from pilot surveys can then be incorporated into simulations to assess the performance of the survey designs for future use in the field.

REFERENCES

Allee, W.C. 1931. Animal aggregations. The University of Chicago Press, Chicago.

Brown J.A. and Manly, B.F.J. 1998. Restricted adaptive cluster sampling. Environ. Ecol. Stat. 5: 49-63.

Cabral, H.N. and Murta, A.G. 2004. Effect of sampling design on abundance estimates of benthic invertebrates in environmental monitoring studies. Mar. Ecol. Prog. Ser. 276: 19-24.

Campbell, A., Boutillier, J., and Rogers, J. 1999a. Discussion on a precautionary approach for management of the red sea urchin fishery in British Columbia. Can. Stock Assess. Sec. Res. Doc. 99/094.

Campbell, A., Hajas, W., and Bureau, D. 1999b. Quota options for the red sea urchin fishery in British Columbia for fishing season 2000/2001. Can. Stock Assess. Sec. Res. Doc. 99/201.

Campbell, A., Tzotzos, D., Hajas, W.C., and Barton, L.L. 2001. Quota options for the red sea urchin fishery in British Columbia for fishing season 2002/2003. Can. Scientific Advice Secretariat Res. Doc. 2001/141.

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

Cochran, W.G. 1977. Sampling Techniques. 3rd edition. John Wiley & Sons, Inc., New York, p.89-91, 107.

Conners, M.E. and Schwager, S.J. 2002. The use of adaptive cluster sampling for hydroacoustic surveys. ICES J. Mar. Sci. 59: 1314-1325.

Ebert, T.A., Dixon, J.D., Schroeter, S.C., Kalvass, P.E., Richmond, N.T., Bradbury, W.A., and Woodby, D.A. 1999. Growth and mortality of red sea urchins (*Strongylocentrotus franciscanus*). Mar. Ecol. Prog. Ser. 190: 189–209.

Fisheries and Oceans Canada. Fisheries development program: new and emerging fisheries. http://www-comm.pac.dfo-mpo.gc.ca/publications/fishdev_e.htm. Accessed: July 6, 2005.

Fisheries and Oceans Canada. Pacific Region Integrated Fisheries Management Plan Red Sea Urchin by Dive, August 1, 2004 to July 31, 2005. http://www-ops2.pac.dfo-mpo.gc.ca/xnet/content/MPLANS/MPlans.htm. Accessed: July 6, 2005.

Jamieson, G.S. and Schwarz, C.J. 1998. Survey protocol considerations for 1995 red sea urchin surveys. *In*: B.J. Waddell, G.E. Gillespie, and L.C. Walthers (eds.). Invertebrate Working Papers reviewed by the Pacific Stock Assessment Review Committee (PSARC) in 1995. Part 2. Echinoderms. Can. Tech. Rep. Fish. Aquat. Sci. 2215: 69-81.

Kosmelj, K., Cedilnik, A., Kalan, P. 2001. Comparison of a two-stage sampling design and its composite sample alternative: an application to soil studies. Environ. Ecol. Stat. 8: 109-119.

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

Levitan, D.R., Sewall, M.A., and Chia, F. 1992. How distribution and abundance influence fertilization success in the sea urchin *Strongylocentrotus franciscanus*. Ecology. 73: 248-254.

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

Morgan, L.E., Botsford, L.W., Wing, S.R., and Smith, B. D. 2000. Spatial variability in growth and mortality of the red sea urchin, *Strongylocentrotus franciscanus*, in northern California. Can. J. Fish. Aquat. Sci. 57: 980–992.

Neyman, J. 1934. On the two different aspects of the representative methods: the method of stratified sampling and the method of purposive selection. J. Roy. Stat. Soc. 97: 558-625.

Perry, R.I., Walters, C.J., and Boutillier, J.A. 1999. A framework for providing scientific advice for the management of new and developing invertebrate fisheries. Rev. Fish Biol. Fish. 9: 125-150.

Press, W.H., Teukolsky, S.A., Vetterling, W.T., and Flannery, B.P. 1992. Numerical Recipes in FORTRAN Sampling. 2nd edition. Cambridge University Press, Cambridge, pp.273-274.

Salehi, M.M. and Seber, G.A.F. 2002. Unbiased estimators for restricted adaptive cluster sampling. Aust. N.Z. J. Stat. 44: 63-74.

Seber, G.A.F. and Thompson, S.K. Environmental adaptive sampling. 1994. *In*: G.P. Patil and C.R. Rao (eds.). Handbook of Statistics. 12: 201-220. Elsevier Science Publishers, North Holland, New York.

Sloan, N.A., Lauridsen, C.P., and Harbo, R.M. 1987. Recruitment characteristics of the commercially harvested red sea urchin *Strongylocentrotus franciscanus* in southern British Columbia, Canada. Fish. Res. 5: 55-69.

Smith, D.R., Conroy, M.J., and Brakhage, D.H. 1995. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. Biometrics. 51: 777-788.

Smith, D.R., Villella, R.F., and Lemarie, D.P. 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environ. Ecol. Stat. 10: 7-15.

Snedecor, G.W. and Cochran, W.G. 1967. Statistical Methods. The Iowa State University Press, Iowa, pp.531-34.

Sparrow, W.H. 1987. Relative mean squared error and cost considerations in choosing group size for group testing to estimate infection rates and probabilities of disease transmission. Phytopathology. 77: 1376-1381.

Su, Z., and Quinn II, T.J. 2003. Estimator bias and efficiency for adaptive cluster sampling with order statistics and a stopping rule. Environ. Ecol. Stat.10: 17-41.

Tegner, M.J. 2001. The ecology of *Strongylocentrotus franciscanus* and *Strongylocentrotus purpuratus*. *In*: J.M Lawrence (ed.). Edible Sea Urchins: Biology and Ecology. 307-331. Elsevier Science B.V., Amsterdam, Netherlands.

Tegner, M.J. and Dayton, P.K. 1977. Sea urchin recruitment patterns and implications of commercial fishing. Science. 196: 324-326.

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

Thompson, S.K. 1991a. Adaptive cluster sampling: designs with primary and secondary units. Biometrics. 47: 1103-1115.

Thompson, S.K. 1991b. Stratified adaptive cluster sampling: designs with primary and secondary units. Biometrika. 78: 389-397.

Thompson, S.K. 1992. Sampling. John Wiley & Sons, Inc., New York.

Thompson, S.K. 1996. Adaptive cluster sampling based on order statistics. Environmetrics. 7: 123-133.

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

Tzotzos, D., Campbell A., and Norgard, T. 2003. Survey of red sea urchin populations in Laredo Channel, British Columbia, 2000. Can. Manuscr. Rep. Fish. Aquat. Sci. 2629.

Woodby, D. 1998. Adaptive cluster sampling: efficiency, fixed sample sizes, and an application to red sea urchins (*Strongylocentrotus franciscanus*) in southeast Alaska. *In*: G.S. Jamieson and A. Campbell (eds.). Proceedings of the North Pacific Symposium on Invertebrate Stock Assessment and Management. Can. Spec. Pub. Fish. Aquat. Sci. 125: 15-20.

TABLES

Table 1. Results of a mixed model fit to the Laredo Channel 2000 survey data. The model uses a covariance structure of compound symmetry, which has constant variance and constant covariance (where the (i,j) element of the residual covariance matrix is $\sigma_1 + \sigma^2 1(i = j)$). Transects and substrate*transect were set as random effects and quadrats as repeated measures within a transect. DF = degrees of freedom and CS = compound symmetry.

(a)

Effect	Numerator DF	Denominator DF	F-value	P
Substrate	8	154	3.75	0.0005

(b)

Covariance parameter	Subject	Estimate
Transect		2.52
Substrate*transect		0.11
CS	Transect	2.04
Residual		27.89

Table 2. Transition matrix of probabilities of a given substrate type in quadrat $a+1$ following a given substrate type in quadrat a, while moving from the shallow to the deep end of the transect. Based on the Laredo Channel 2000 survey data. Note that only 3 of 2303 quadrats within the field survey data were categorized as pea gravel.

		Quadrat *a* +1								
		Smooth bedrock	Bedrock with crevices	Boulders	Cobble	Gravel	Pea gravel	Sand	Shell	Mud
	Smooth bedrock	0.74	0.10	0.09	0.01	0.00	0.00	0.05	0.01	0.00
	Bedrock with crevices	0.07	0.82	0.06	0.02	0.00	0.00	0.02	0.01	0.00
	Boulders	0.02	0.04	0.82	0.04	0.01	0.00	0.05	0.02	0.00
	Cobble	0.03	0.03	0.12	0.76	0.01	0.01	0.04	0.01	0.00
Quadrat *a*	Gravel	0.00	0.00	0.04	0.04	0.85	0.08	0.00	0.00	0.00
	Pea gravel	0.00	0.00	0.67	0.00	0.00	0.00	0.33	0.00	0.00
	Sand	0.01	0.00	0.01	0.00	0.00	0.00	0.97	0.00	0.00
	Shell	0.03	0.07	0.01	0.01	0.00	0.00	0.01	0.85	0.01
	Mud	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.99

Table 3. Transition matrices for each substrate type with three categories of counts for urchins: zero count, low count (1 - 10 urchins), and high count (>10 urchins). Based on the Laredo Channel 2000 survey data. Values in the matrices indicate the probability of a count category in quadrat *a+1* following quadrat *a* of the same substrate type, while moving from the shallow to the deep end of the transect. The values were not calculated between substrate types.

		Quadrat *a*+1		
Substrate type	**Quadrat *a***	zero count	low count	high count
Smooth Bedrock	zero count	0.83	0.12	0.05
	low count	0.54	0.36	0.10
	high count	0.23	0.31	0.46
Bedrock with crevices	zero count	0.85	0.13	0.02
	low count	0.28	0.65	0.06
	high count	0.25	0.67	0.08
Boulders	zero count	0.82	0.15	0.03
	low count	0.38	0.49	0.13
	high count	0.14	0.68	0.18
Cobble	zero count	0.99	0.01	0.00
	low count	0.50	0.50	0.00
	high count	0.00	0.00	1.00
Gravel	zero count	0.99	0.01	0.00
	low count	0.50	0.50	0.00
	high count	0.00	0.00	1.00
Pea gravel	zero count	1.00	0.00	0.00
	low count	0.00	0.00	0.00
	high count	0.00	0.00	0.00
Sand	zero count	1.00	0.0025	0.00
	low count	0.67	0.33	0.00
	high count	0.00	0.00	0.00
Shell	zero count	1.00	0.00	0.00
	low count	0.00	0.00	0.00
	high count	0.00	0.00	0.00
Mud	zero count	1.00	0.00	0.00
	low count	0.00	0.00	0.00
	high count	0.00	0.00	0.00

Table 4. Mean number of urchins in the low- and high-count categories and the maximum urchin count values ("Max. val.") for the substrate types in the Laredo Channel 2000 survey data. The values were used to parameterize substrate-specific Poisson distributions. Count values generated from a Poisson distribution based on a low-count category were between 1 and 10. Count values generated from a Poisson distribution based on high-count category were between 11 and "Max. val.". Gravel, pea gravel, and mud are excluded from the table because they do not have a low- or high-count category; their count data contained only zeros.

	Smooth Bedrock		Bedrock with Crevices		Boulders		Cobble		Sand		Shell	
Category:	**low count**	**high count**	**low count**	**high count**	**low count**	**high count**	**low count**	**high count**	**low count**	**high count**	**low count**	**high count**
Mean	3.1	28.5	3.6	23.4	4.0	18.1	2.0	17.0	3.4	-	3.5	-
Max. val.	9	194*	10	61	10	38	6	21	7	-	4	-

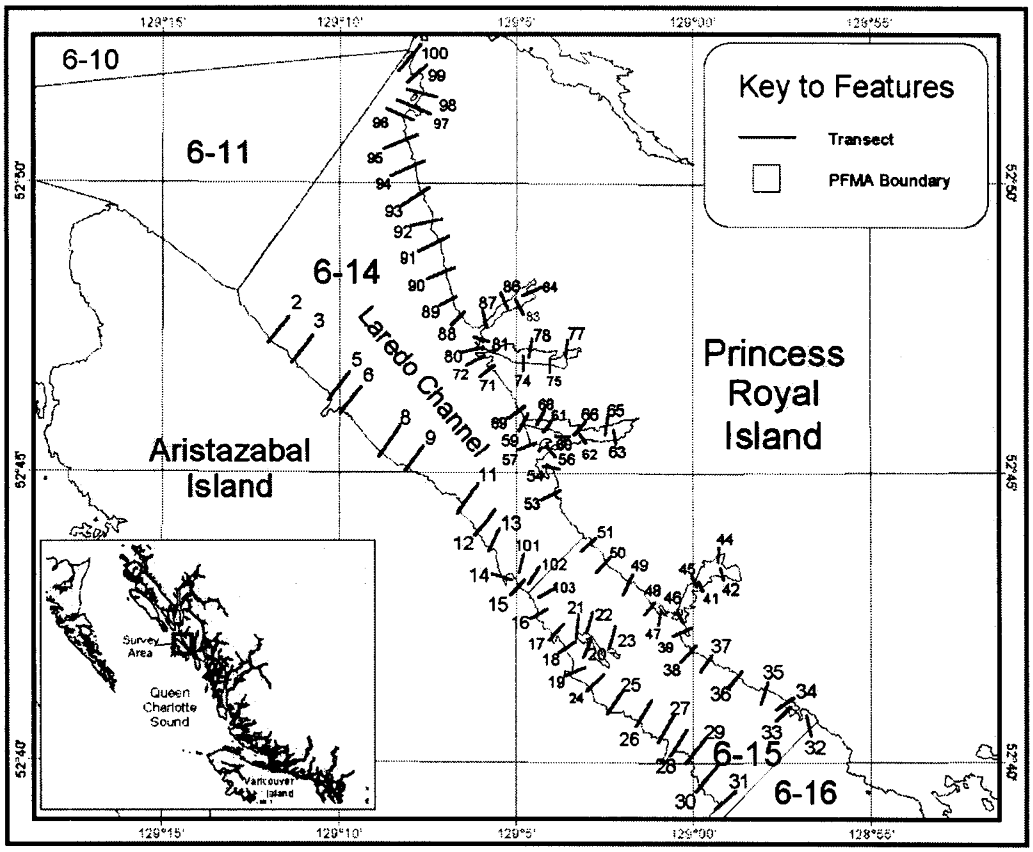
*Actual value used in model was 170. The probability of selecting a value larger than 170 was less than $2x10^{-72}$.

Table 5. Relative sampling effort for each substrate type in the simulation based on the Laredo Channel 2000 survey data. Sampling effort within each substrate type was determined using Neyman optimal allocation and was set relative to smooth bedrock. Sampling in the gravel, pea gravel, shell, and mud substrates was set to every 20th quadrat because the calculated effort relative to smooth bedrock was extremely small or zero. N_h = proportion of stratum h (substrate) in the field survey data; s = standard deviation of the number of urchins per quadrat in the field survey data.

Substrate type (h)	N_h	s_h	$N_h \cdot s_h$	**Effort relative to smooth bedrock**	**Every Xth quadrat sampled**
Smooth bedrock	0.13	12.80	1.61	1.00	1
Bedrock with crevices	0.17	5.33	0.92	0.57	2
Boulders	0.16	2.36	0.39	0.24	4
Cobble	0.05	2.35	0.12	0.08	13
Gravel	0.01	0.00	0.00	0.00	20
Pea gravel	0.001	0.00	0.00	0.00	20
Sand	0.38	0.38	0.14	0.09	11
Shell	0.03	0.56	0.02	0.01	20
Mud	0.06	0.00	0.00	0.00	20
Total	1.00				

FIGURES

Figure 1. Map of Laredo Channel, British Columbia displaying the systematic layout of transects for the survey conducted in 2000. Missing transect numbers are transects that were not surveyed due to logistical difficulties. Hyphenated numbers indicate Pacific Fishery Management sub-areas. Inset map provides location of the survey on a larger scale. PFMA= Pacific Fishery Management Area.



Reprinted with permission from Tzotzos et al. (2003).

Figure 2. Mean number of urchins (± SE) per 1 m^2 quadrat for nine different substrate types (open bars) and associated substrate composition of the survey area (shaded bars). Based on Laredo Channel 2000 survey data.

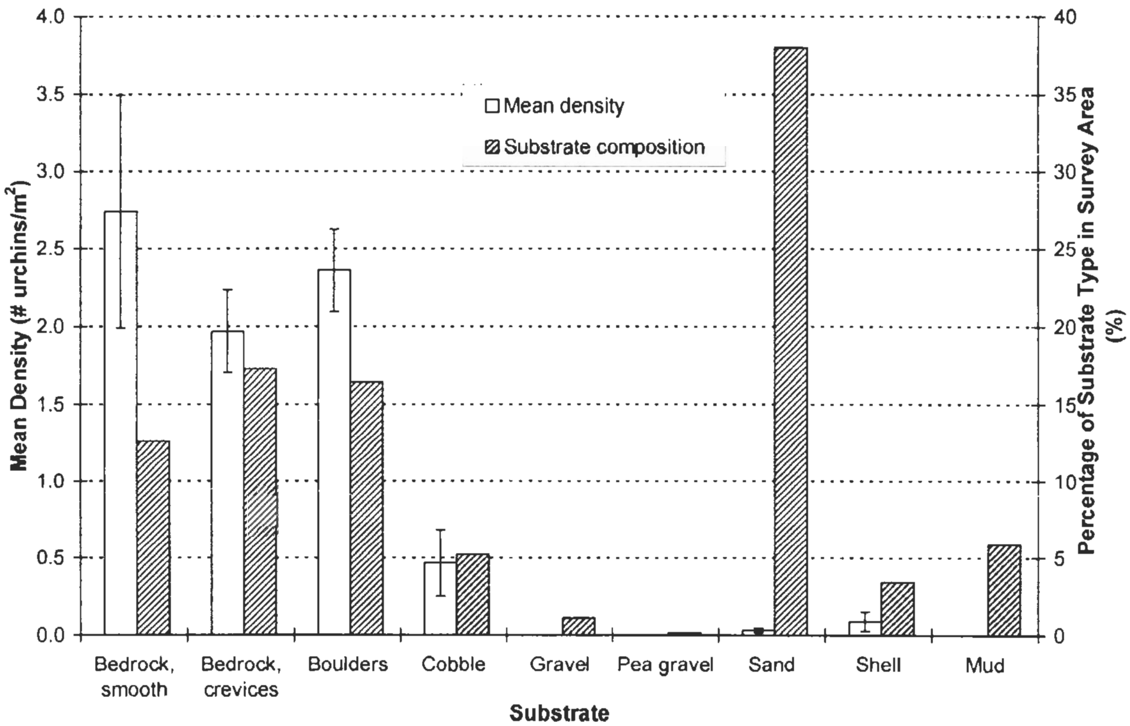


Figure 3. Mean proportion of a transect (± SE) with quadrats occupied by red sea urchins for different ranges of transect length (open bars). Mean proportion of a transect (± SE) with quadrats composed of bedrock and/or boulders for different ranges of transect length (shaded bars). Based on Laredo Channel 2000 survey data.

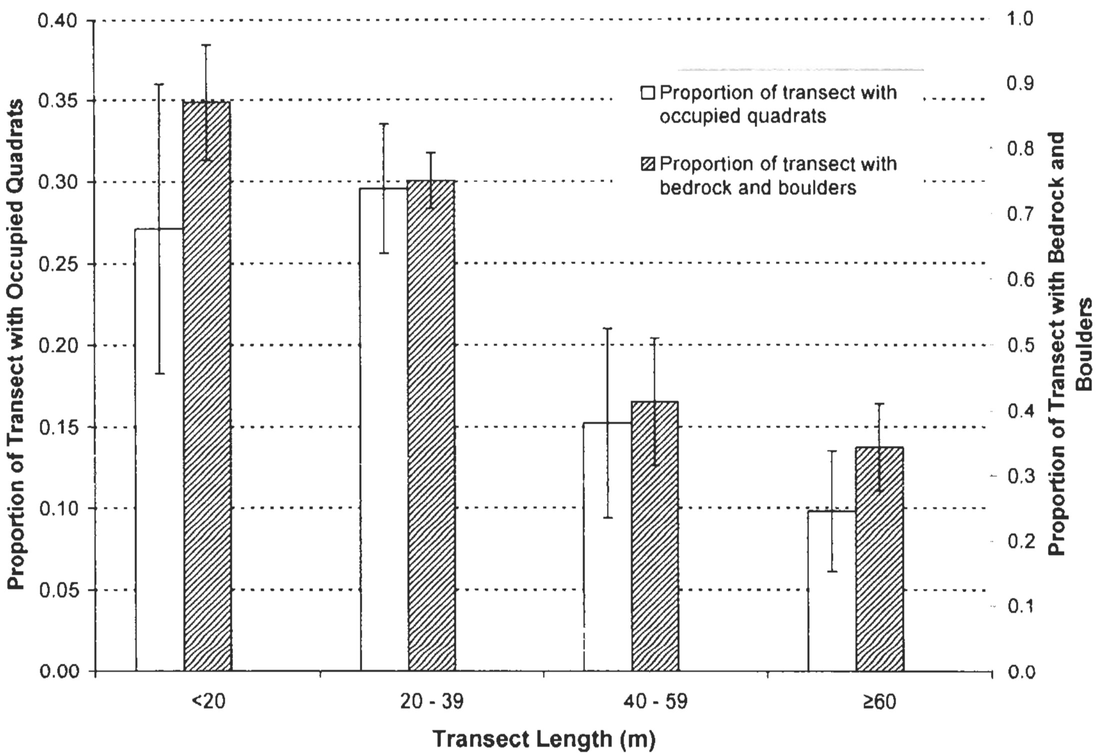


Figure 4. An example of the bootstrapping procedure using a subset of the Laredo Channel 2000 survey data. Length of transect and substrate proportions within a transect are bootstrapped together. Letters A to G represent the transects. Numbers are used to arbitrarily represent the different substrate types. Note that the order of transects and substrate type within each transect is different after bootstrapping. Bootstrapping of urchin count data is not shown here.

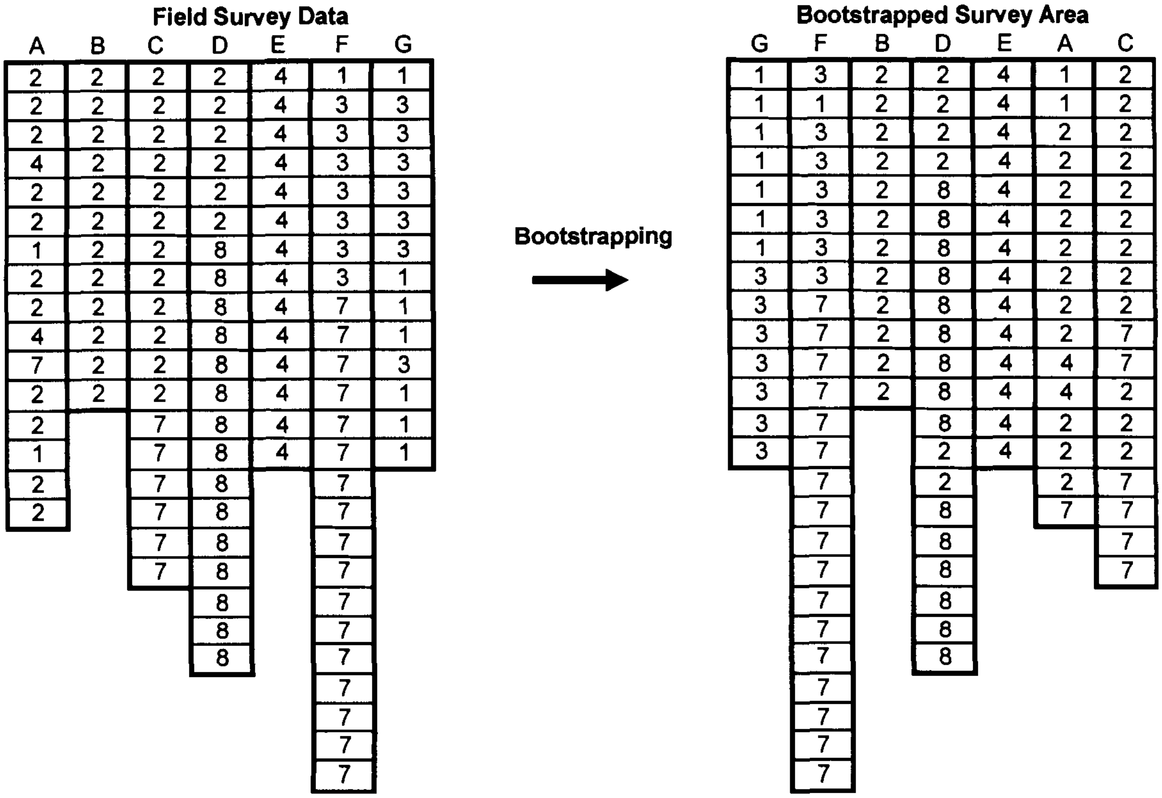


Figure 5. An example of the restricted adaptive cluster design (RTSacs) (not drawn to scale). An initial sample of quadrats is systematically selected (every third quadrat) from a randomly selected transect. The starting point within the initial sample of every third quadrat in the transect is randomly selected; in this example it is quadrat 1 (other possible starting points are quadrats 2 and 3). The number of quadrats sampled in the initial sample of this example is four. Adaptive sampling occurs in the neighbourhood of sampled quadrats that satisfy the density condition of ≥ 1 urchin. The neighbourhood is restricted to the next adjacent quadrat along the transect. A cluster is a contiguous group of sampled units that can be divided into edge units and a network. The network includes those units in the cluster that satisfy the density condition, whereas edge units refer to observations that fall short of the condition. x_k is the number of possible initial samples that a network intersects; in this example there are three possible samples (i.e. maximum value of x_k corresponds to the number of potential starting points).

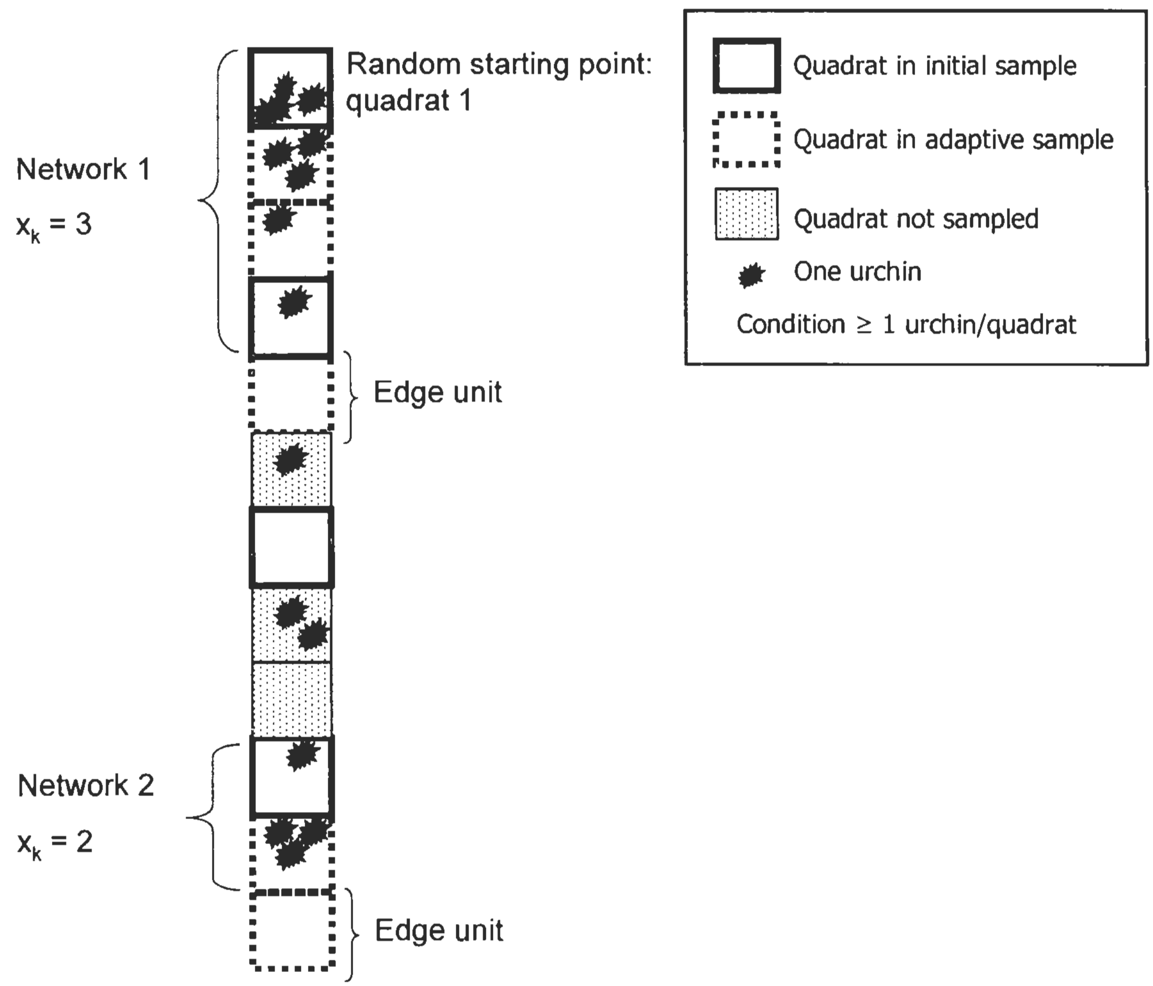


Figure 6. Flowchart outlining the basic procedure of the simulation model. Sampling designs include simple random sampling (SRS), DFO's current survey method (RTScurrent), simple random transect sampling (RTSsimple), restricted adaptive cluster survey (RTSacs), and within-transect stratification survey (RTSstrat). Sample size ranged from 5 to 45 transects with a step size of 10. The same survey designs were also tested using a stratified survey area. In the latter scenarios, sample size ranged from 25 to 45 transects with a step size of 10. MC = Monte Carlo.

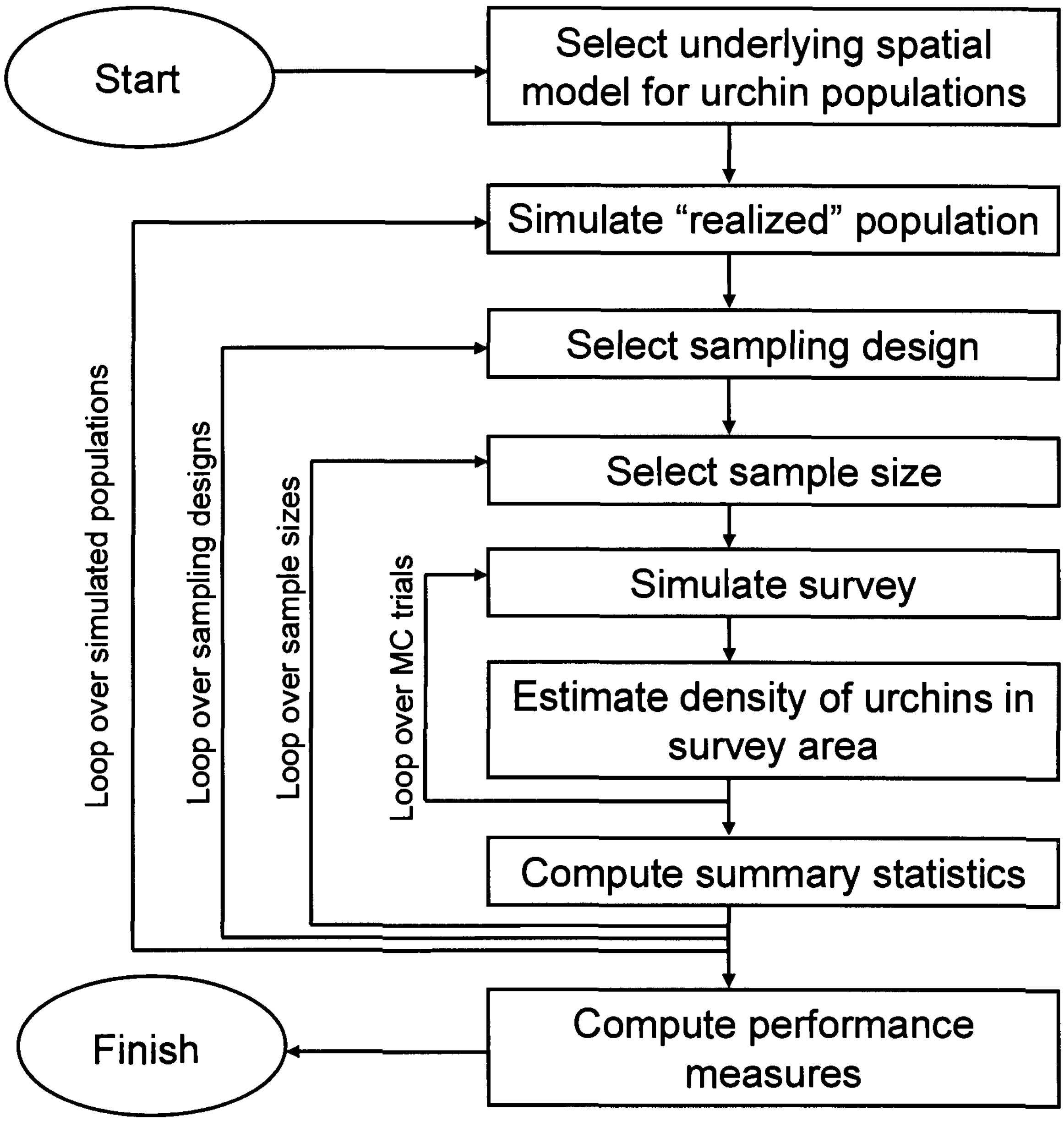


Figure 7. (a) Distribution of mean densities of the 200 realizations of "true" urchin populations. (b) Distribution of the coefficient of variation of the 200 realizations of "true" urchin populations. Each box represents the median and the quartiles. Whiskers extend to the most extreme data point that is no more than 1.5 times the length of the interquartile range. Circles represent values that fall outside the whiskers' range.

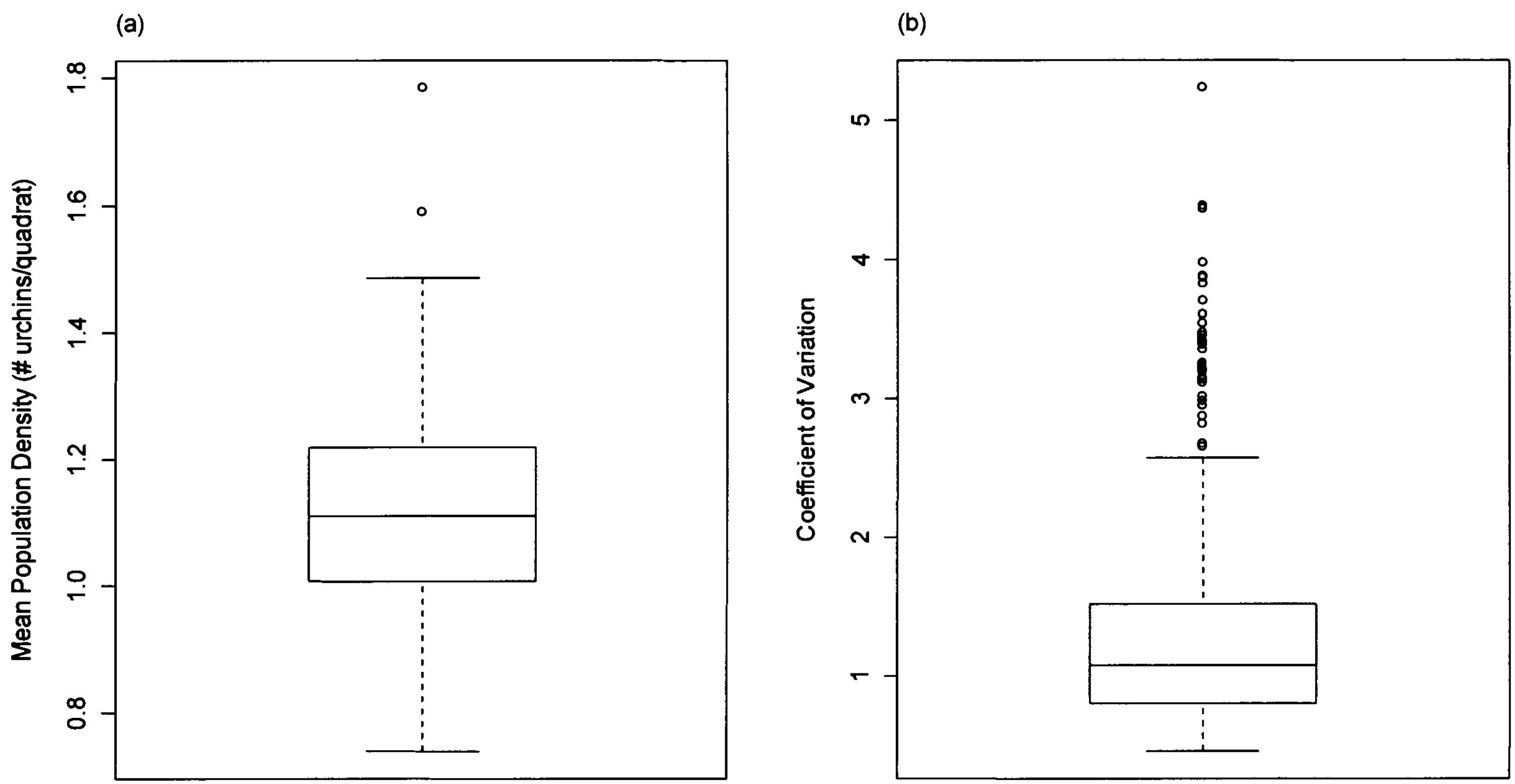


Figure 8. Distribution of the relative efficiency (RE) of each design without stratification of the survey area relative to DFO's current random transect survey (RTScurrent) for 200 urchin populations (i.e. [variance of RTScurrent]/[variance of estimator of interest]). Sampling designs include simple random transect survey (RTSsimple), within-transect stratification survey (RTSstrat), and restricted adaptive cluster survey (RTSacs). Efficiency >1 indicates that the design is more efficient than RTScurrent. n = sample size (# transects). Each box represents the median and the quartiles. Whiskers extend to the most extreme data point which is no more than 1.5 times the length of the interquartile range. Dots represent values that fall outside the whiskers' range.

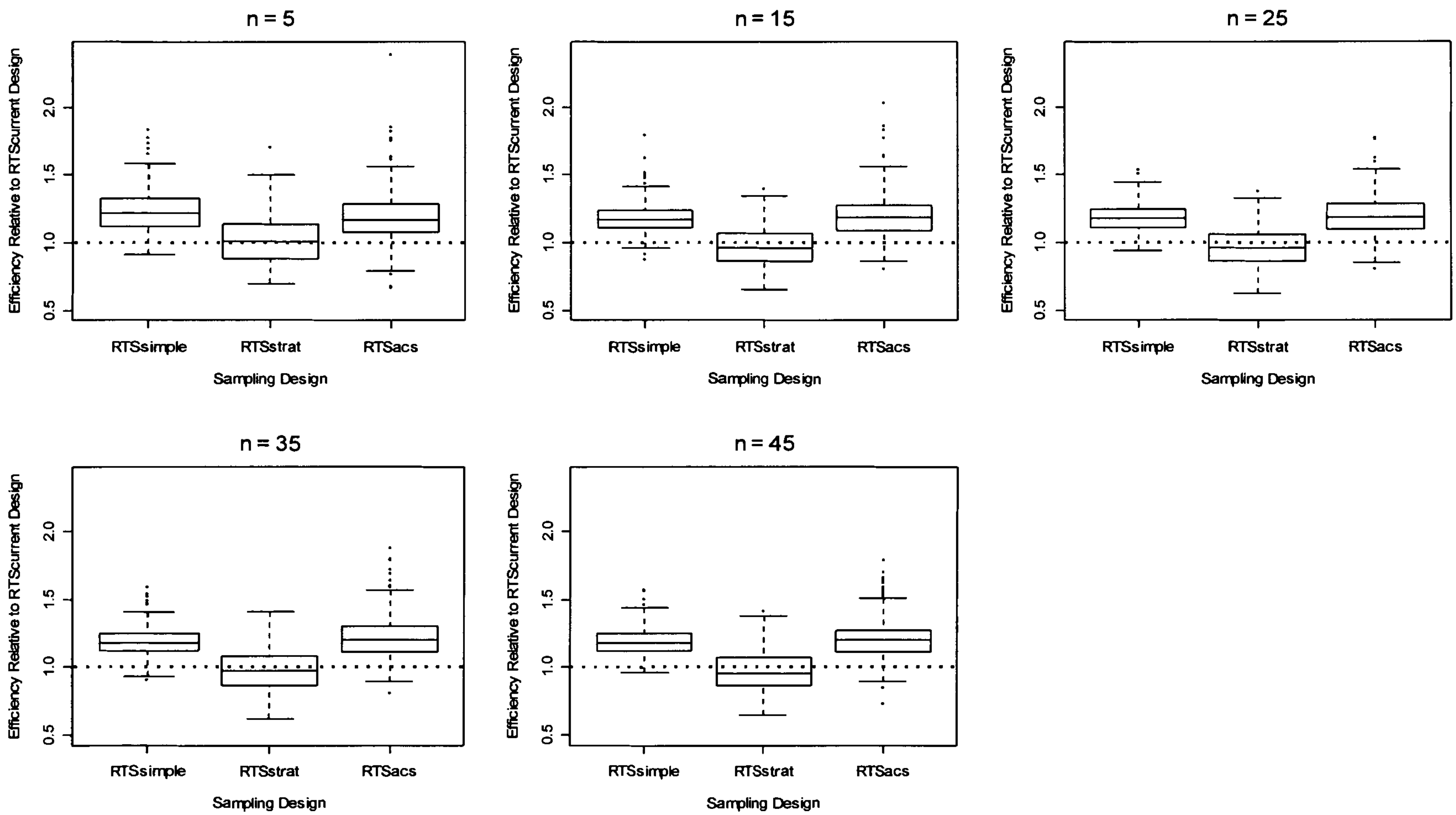


Figure 9. Distribution of the relative efficiency (RE) of each sampling design without stratification of the survey area relative to a simple random sampling design (SRS) for 200 urchin populations (i.e. [variance of SRS]/[variance of estimator of interest]). Efficiency >1 indicates that the design is more efficient than SRS. Sampling designs, sample sizes, and box plots are as defined in Figure 8.

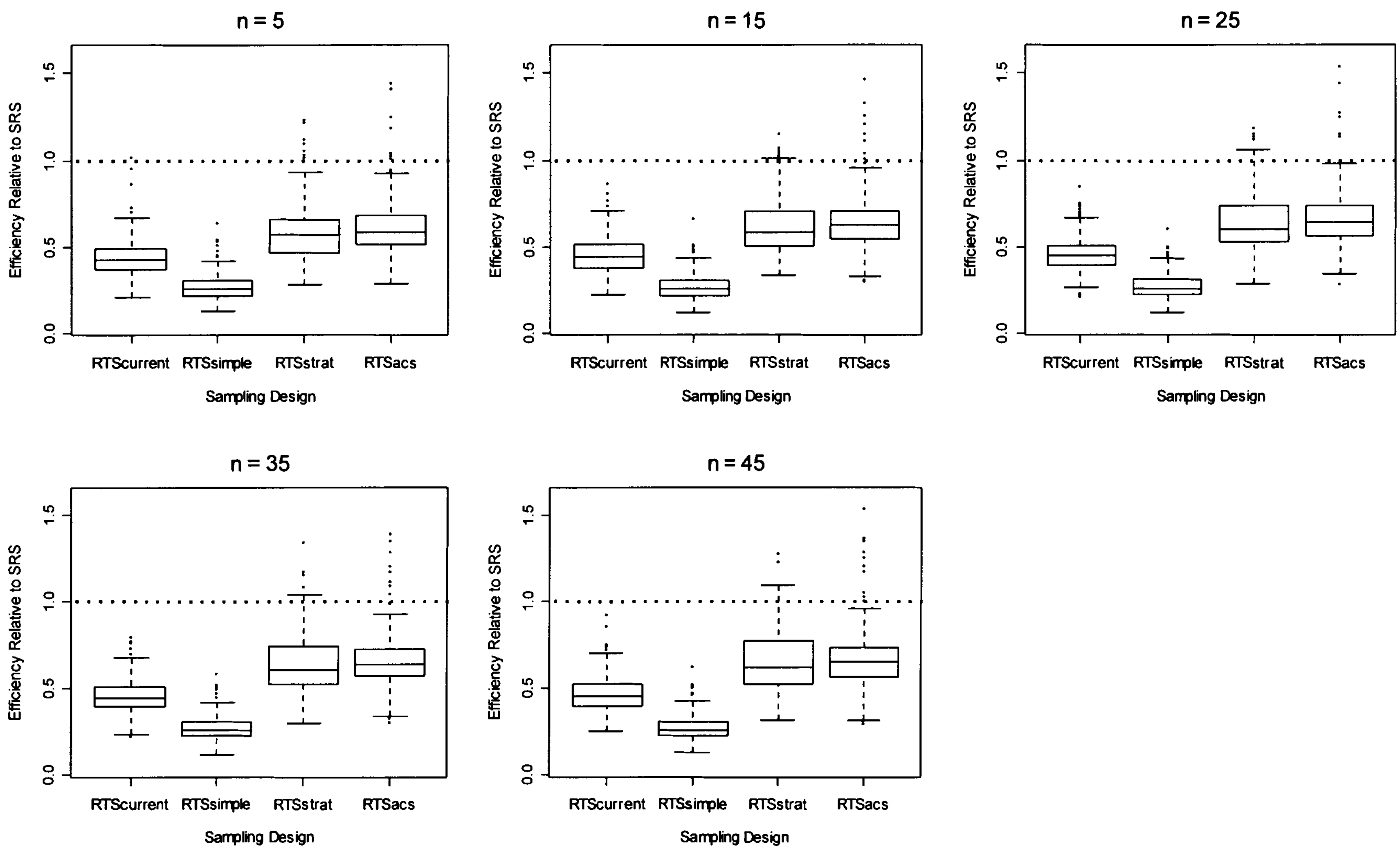


Figure 10. Distribution of the relative bias of each design without stratification of the survey area for 200 urchin populations. Sampling designs, sample sizes, and box plots are as defined in Figure 8.

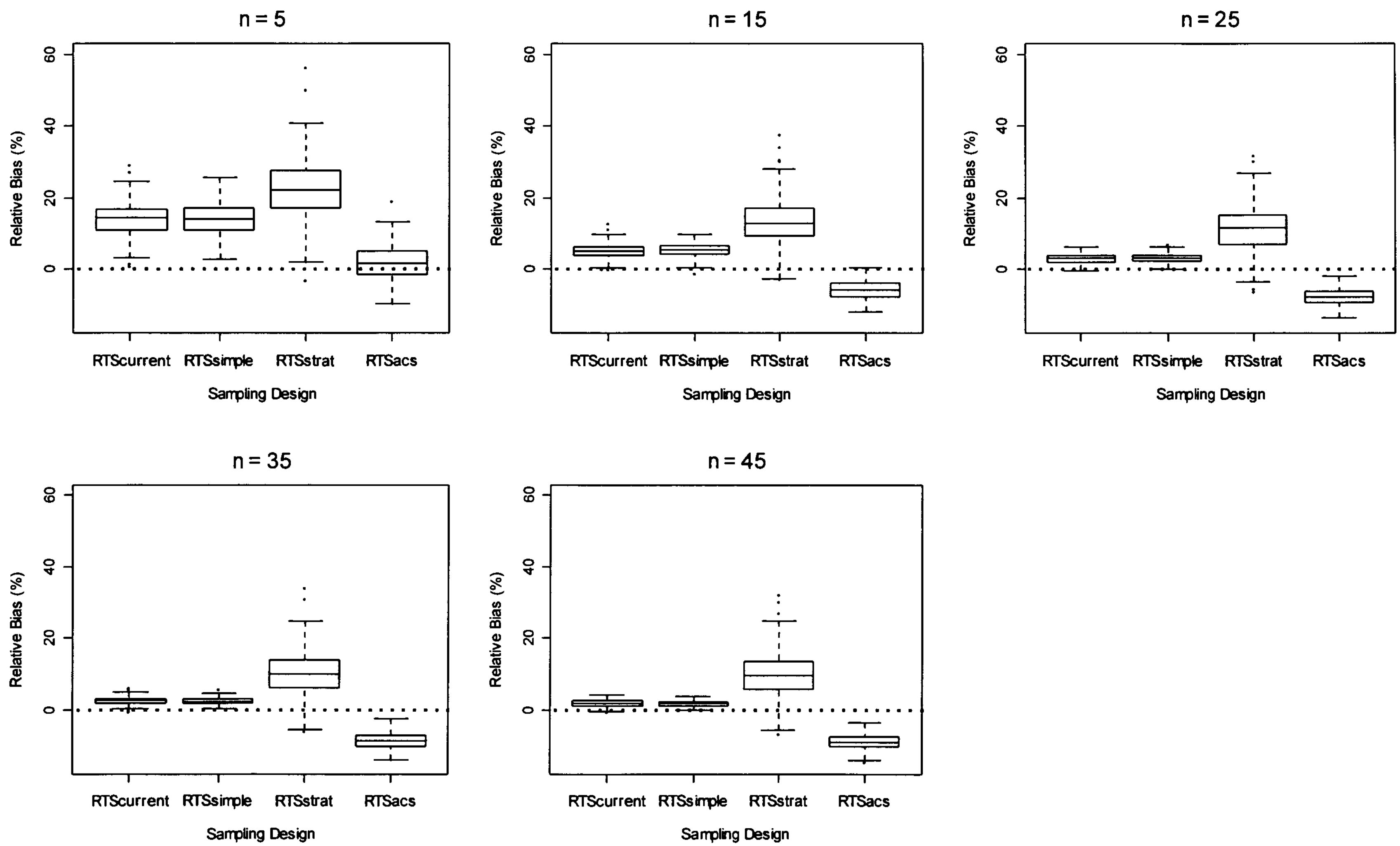


Figure 11. Distribution of the standard error of the mean density for the sampling designs without stratification of the survey area for 200 urchin populations. Sampling designs, sample sizes, and box plots are as defined in Figure 8.

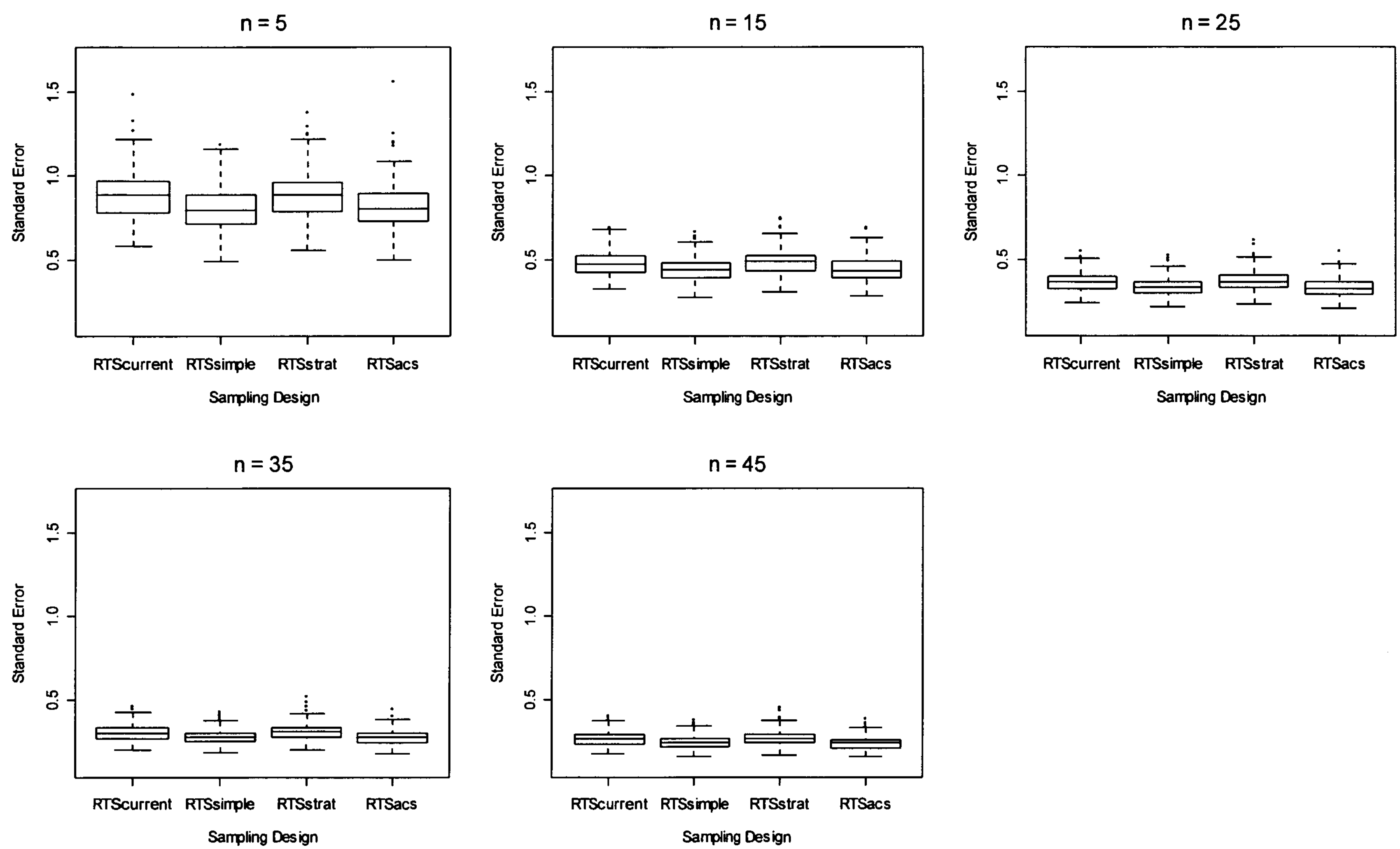


Figure 12. Distribution of the inefficiency (marginal cost x variance) of each design without stratification of the survey area for 200 urchin populations. Sampling designs, sample sizes and box plots are as defined in Figure 8.

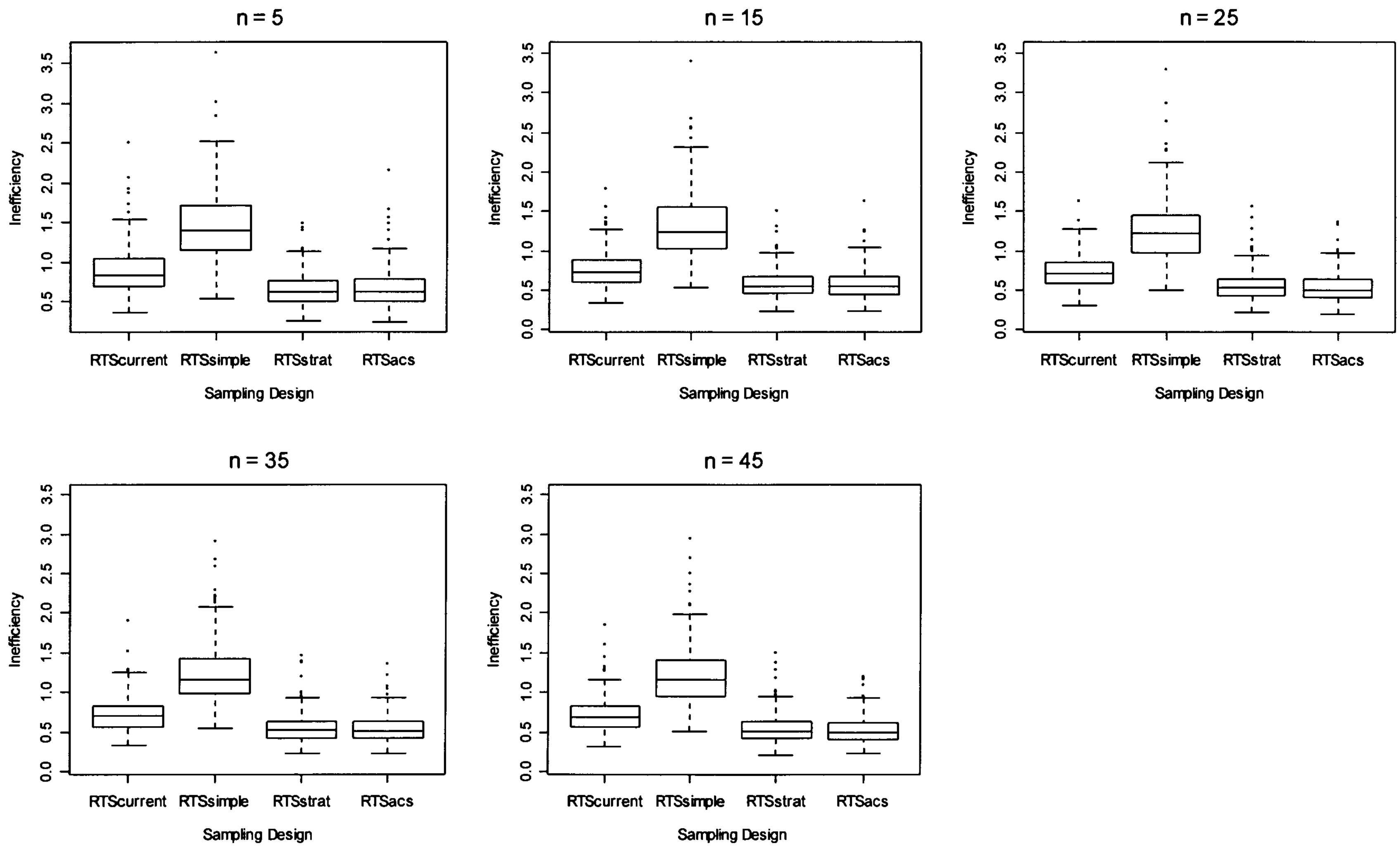


Figure 13. Distribution of the relative efficiency (RE) of each sampling design in a stratified survey area relative to DFO's current random transect survey (RTScurrent without stratification of the survey area) for 200 urchin populations ([variance of RTScurrent]/[variance of estimator of interest]). Efficiency >1 indicates that the design is more efficient than RTScurrent (without stratification). Sampling designs include RTScurrent, simple random transect survey (RTSsimple), within-transect stratification survey (RTSstrat), and restricted adaptive cluster survey (RTSacs). Sample sizes (n) of 25 to 45 transects were used in scenarios involving stratification of the survey area. Each box represents the median and the quartiles. Whiskers extend to the most extreme data point which is no more than 1.5 times the length of the interquartile range. Circles represent values that fall outside the whiskers' range.

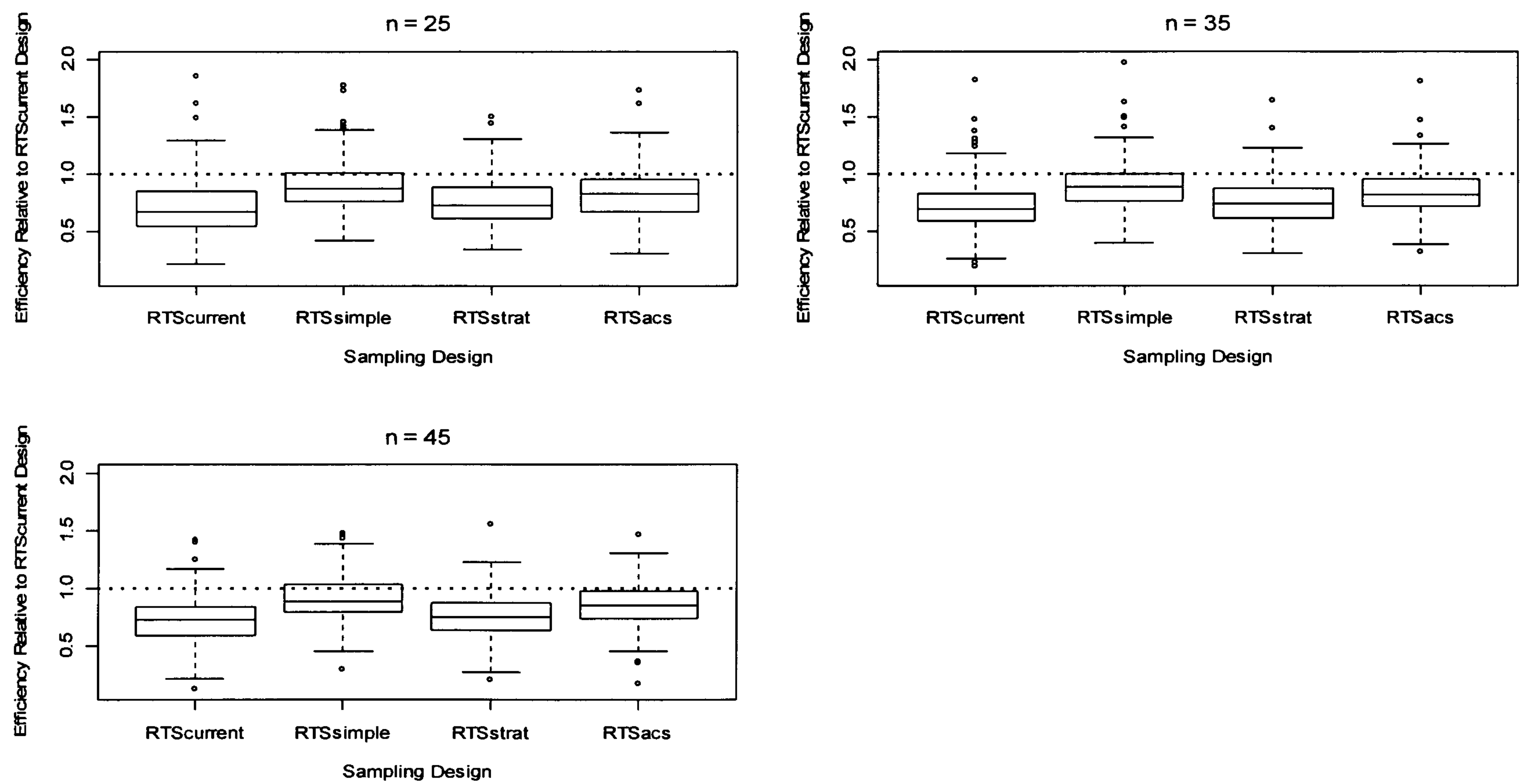


Figure 14. Distribution of the relative efficiency (RE) of each sampling design in a stratified survey area relative to simple random sampling (SRS) for 200 urchin populations (i.e. [variance of SRS]/[variance of estimator of interest]). Efficiency >1 indicates that the design is more efficient than SRS. Sampling designs, sample sizes, and box plots are as defined in Figure 13.

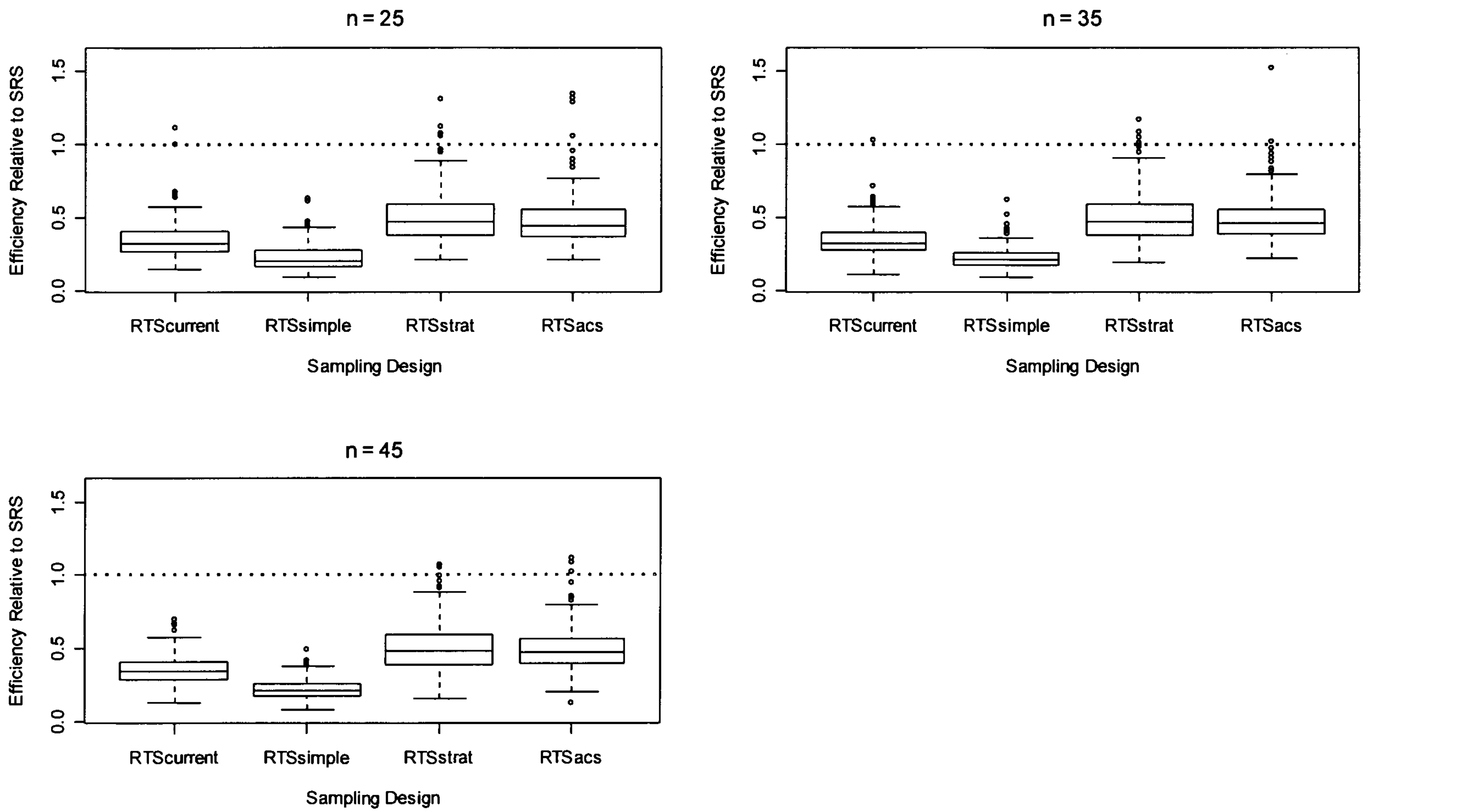


Figure 15. Distribution of the relative bias of the each design in a stratified survey area for 200 populations. Sampling designs, sample sizes, and box plots are as defined in Figure 13.

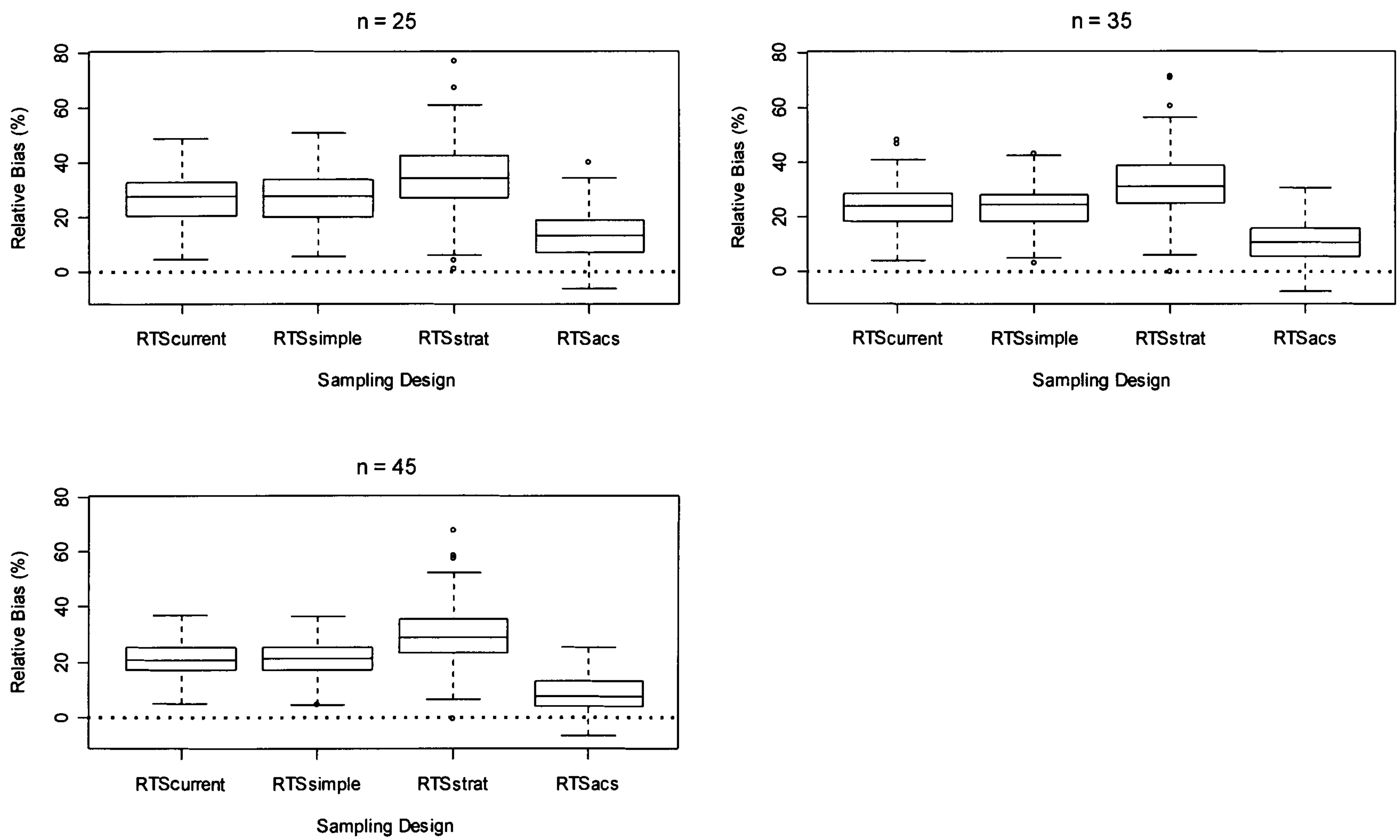


Figure 16. Distribution of the standard error of the mean density for the sampling designs in a stratified survey area for 200 urchin populations. Sampling designs, sample sizes, and box plots are as defined in Figure 13.

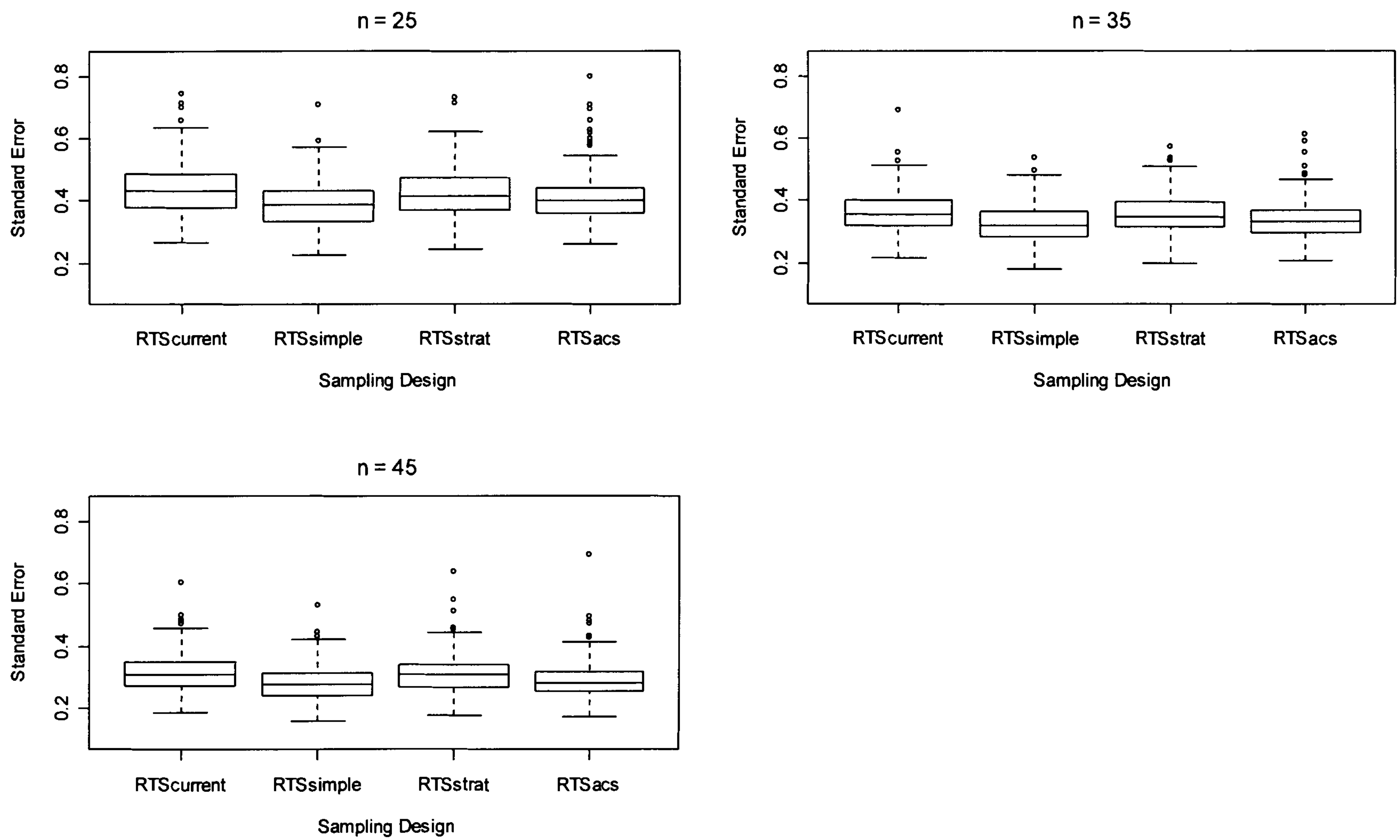
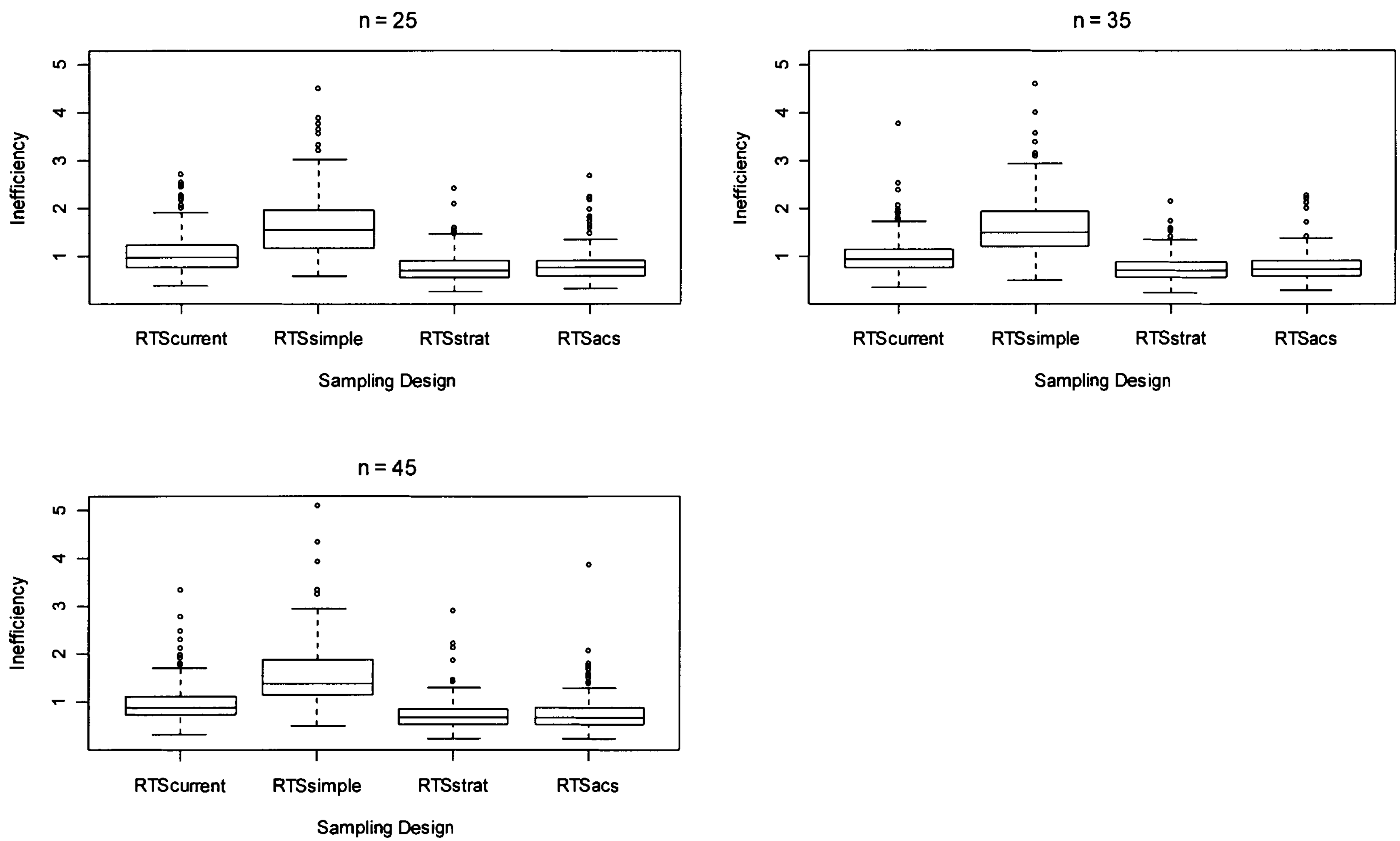


Figure 17. Distribution of the inefficiency (marginal cost x variance) of each sampling design in a stratified survey area for 200 urchin populations. Sampling designs, sample sizes, and box plots are as defined in Figure 13.



APPENDIX

For some variations of adaptive cluster sampling (ACS), the estimators produce the same density estimate. This simulation study used one such variation that resulted in the Hansen-Hurwitz (HH) estimator and the Horvitz-Thompson (HT) estimator providing equivalent estimates. This variation of ACS was a restricted adaptive cluster sampling design (RTSacs) where the neighbourhood for adaptive sampling was restricted to the next adjacent quadrat along the transect. The following example demonstrates that both estimators produce the same density estimate when sampling is based on the RTSacs design:

<u>Assume</u>:

$M = 4$
$N = 3$
$n = 1$
$\sum y_1 = 8$
$\sum y_2 = 1$
$x_1 = 2$
$x_2 = 1$

<u>HH estimator</u>

$$w_i = \frac{1}{M}\sum_{k=1}^{K}\frac{y_k^*}{x_k} \quad = \quad 1/4 * [8/2 + 1/1] = \quad 1.25$$

$$\hat{\mu}_{HH} = \frac{1}{n}\sum_{i=1}^{n} w_i \quad = \quad 1.25$$

HT estimator

$$\alpha_1 = 1 - \left[\binom{N - x_1}{n} \Big/ \binom{N}{n} \right] \quad = \quad 1 - \left[\binom{3-2}{1} \Big/ \binom{3}{1} \right] \quad = \quad .67$$

$$\alpha_2 = 1 - \left[\binom{3-1}{1} \Big/ \binom{3}{1} \right] \quad = \quad .33$$

$$\hat{\mu}_{HT} = \frac{1}{MN} \sum_{k}^{K} \frac{y_k^*}{\alpha_k} \quad = \quad \frac{1}{12} * \left(\frac{8}{2/3} + \frac{1}{1/3} \right) \quad = \quad 1.25$$