Mar Biol (2007) 151:1343–1348
DOI 10.1007/s00227-006-0571-2

RESEARCH ARTICLE

The application of adaptive cluster sampling for rare subtidal macroalgae

Nisse A. Goldberg · John N. Heine · Jennifer A. Brown

Received: 25 July 2006 / Accepted: 23 November 2006 / Published online: 15 December 2006

Abstract Adaptive cluster sampling (ACS) is a targeting sampling method that provides unbiased abundance estimators for populations of rare species that may be inadequately sampled with simple random sampling (SRS). ACS has been used successfully to estimate abundances of rockfish and sardine larvae from shipboard surveys. In this study, we describe the application of ACS for subtidal macroalgae. Using SCUBA, we measured abundances of *Codium mamillosum*, *C. pomoides*, and *Halimeda cuneata* at three islands and two levels of wave exposure. The three species were relatively patchy and could be sampled with ACS at one site per dive. Their distributions differed among islands and with exposure to wave energy, with *H. cuneata* found at only one island. ACS is a useful tool for understanding the spatial distribution and abundance of populations of rare benthic species, but, as was the case in this study, may not be as efficient as sampling with SRS with comparable replication.

Communicated by S.D. Connell, Adelaide.

N. A. Goldberg (✉)
School of Plant Biology, University of Western Australia,
35 Stirling Highway, Crawley, WA, Australia
e-mail: lumilumi@yahoo.com

J. N. Heine
CSIRO Marine and Atmospheric Research,
Strategic Research Fund for the Marine Environment,
Private Bag No. 5, Wembley, WA 6913, Australia

J. A. Brown
Biomathematics Research Centre,
University of Canterbury, Private Bag 4800,
Christchurch, New Zealand

Introduction

Rare species typically outnumber common species (Cao et al. 1998; Magurran and Henderson 2003) in terrestrial and marine systems and yet, most ecological studies are conducted on common species (Chapman 1999). Rare species have been removed from diversity analyses and treated as statistical noise on account of their patchy distributions that translate into large variances (Cao et al. 1998). However, the questions that have been addressed for common species remain to be investigated for many rare marine species. Such questions include: what are their habitat requirements, what processes contribute to their persistence and rarity, and are species rare throughout their range (Chapman 1999)? Answers to these questions may help identify the status of rare marine taxa (Chapman 1999).

Investigators that aim to describe the distribution of rare species are faced with challenges in estimating abundances of rare species that have different distributions from those of common species. Species-abundance curves for assemblages are generally skewed to the right (Magurran and Henderson 2003) with most species collected in a survey being rare (Cao et al. 1998). Adaptive cluster sampling (ACS) is a sampling design that targets rare species that are spatially patchy, and can be used to collect valuable information with respect to temporal and spatial persistence of individuals and patches, patch distributions, and a species' habitat use (Thompson 1990; reviewed in Turk and Brokowski 2005).

Adaptive cluster sampling is based on 'network sampling' in which the probability of detecting a rare species is greater in the vicinity where the species has already been recorded (Thompson 1990). The advantage of ACS is in its flexibility as the researcher determines the initial sample size, the size of the sampling unit, and the sampling condition upon which adaptive sampling is initiated (Turk and Brokowski 2005). First, an initial set of units (e.g., quadrats) is sampled randomly. If the target species is present, then neighboring units are sampled to form a network or patch. Conditions are defined as to whether a unit is included in a network. For example, a unit would be included if there is ≥1 individuals present in the quadrat. By definition, edge units surround a network but do not meet the condition requirement (Thompson 1990; Philippi 2005).

Adaptive cluster sampling is a targeted sampling methodology that utilizes unbiased estimators to calculate the mean and variance of clustered and rare species (Thompson 1990). These estimators give unbiased estimates because they do not require assumptions regarding the population. However, whether the design yields a good estimate (i.e., one with low variance) will depend on the distribution of the population and high within-network variance (Thompson 1990). Two estimators were proposed by Thompson (1990): the Horvitz–Thompson estimator incorporates the probability that a unit is included in the final sample, and the Hansen–Hurwitz estimator incorporates the probability that a unit's network is intersected by the initial sample.

The application of ACS is described using subtidal algae sampled from the Recherche Archipelago, Western Australia. ACS had not been assessed for subtidal benthic organisms, but has been used with mobile taxa such as fish (Hanselman and Quinn II 2004; Smith et al. 2004). We applied ACS to estimate densities of three algal species whose distributions were presumed to be patchy and differ with exposure to wave energy as indicated by previous surveys (Goldberg, unpublished data). Of the 216 quadrats that were sampled in the Recherche Archipelago, 63% of the species were recorded in ≤10% of the quadrats (Fig. 1; Goldberg, unpublished data). These infrequent species can have widespread distributions throughout southern Australian waters (Womersley 1990; Phillips 2001) in contrast to the distributions of rare and endemic decapods, numerous species of echinoderms, and molluscan fauna that were found to be restricted to vulnerable habitats or with relatively short ranges (O'Hara and Poore 2000; O'Hara 2002). The low frequency of occurrences for numerous algal species contributed to large estimated variances, a finding that motivated the authors to test an alternative method in estimating the abundances of rare and patchy species. The primary aim of this study was to apply ACS to describe the size of patches and to estimate densities and variances of subtidal benthic algae. The relative efficiency of ACS versus simple random sampling (SRS) was evaluated by comparing standard errors based on comparable sample sizes (Noon et al. 2006).

Materials and methods

The Recherche Archipelago is located along the southern coast of Australia and is exposed to Southern Ocean wind and swells. We sampled Black, Woody, and Thomas Islands, which are located in the relatively protected waters of Esperance Bay. Subtidal reef platforms are granitic and characterized with diverse subtidal macroalgal assemblages that differ between wave-exposed and wave-sheltered sites (Goldberg and Kendrick 2004).

Three species of algae were sampled using ACS: *Codium mamillosum*, *C. pomoides*, and *Halimeda cuneata*. *C. mamillosum* and *C. pomoides* have globose thalli, and *H. cuneata* is an erect, calcified alga. *C. pomoides* is endemic to Australia (Womersley 1984). These species were selected on the basis that they were easy to identify and have been recorded previously at Black, Woody, and Thomas Islands (Goldberg and Kendrick 2004). *C. mamillosum* was the most common of the three species. The alga was recorded at Black, Woody, and Thomas Islands in depths ≤20 m, with a frequency of occurrence (total sample size = 12–0.25 m^2 quadrats per exposure per island) ranging between 50 and 83% at wave-exposed sites, and 42 and 58% at wave-sheltered sites (Goldberg, unpublished data). *C. pomoides* was the second most common of the three species among the islands and was recorded at Woody and Thomas Islands with a sampling frequency of 8 and 33% at wave-exposed sites, and 8 and 50% at wave-sheltered sites. *H. cuneata* was rare among the islands and was recorded only at Black Island with a frequency of occurrence of 50% at wave-exposed sites and 8% at wave-sheltered sites.

To describe their spatial distribution, *C. mamillosum*, *C. pomoides*, and *H. cuneata* were sampled in depths between 7 and 13 m at wave-exposed and wave-sheltered sites at Black, Woody and Thomas Islands. Each species was sampled separately within a study area of 48 m^2 at each of the sites. The sampling condition (i.e., the number of individuals per 0.25 m^2 to trigger ACS) was determined by swimming over the area to evaluate general abundance and patchiness. To maximize sampling efficiency, we aimed to keep the number of units in a network relatively small (2–4 units, Brown 2003) by adjusting the sampling condition.

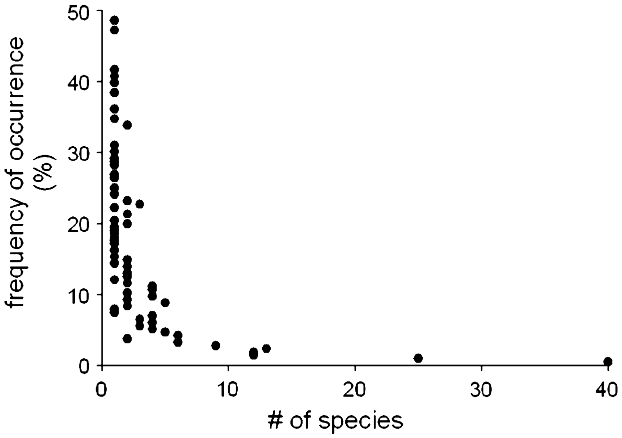

Fig. 1 Frequency of occurrence for subtidal macroalgal species recorded in the Recherche Archipelago, Western Australia. Total number of samples was 216, collected from 11 islands. Sample area = 0.25 m^2 (Goldberg, unpublished data)

Although the sampling condition for each species varied among the different islands and between exposures, the results still provided unbiased density estimates.

Given the logistical challenges of sampling networks in subtidal, turbulent waters within a specified area and working in buddy-pairs, we used a semi-systematic sampling design to start ACS. Systematic designs guarantee sampling coverage throughout a study area despite the potential of overestimated variances (Thompson 1992; Smith et al. 2003). Sampling coverage within a study area (i.e., replication) is critical as it underpins the precision and accuracy of abundance estimates (Drummond and Connell 2005). Four 10-m transects were placed in the study area, with the first transect placed randomly and the other three transects placed at regular intervals from the initial transect. The transects provided a means to orient the divers for network sampling. Along each transect, four initial quadrats were placed randomly and numbers of individuals per quadrat were recorded. If the initial quadrat satisfied the condition (e.g., ≥1 individuals), then adjacent quadrats were sampled in a contiguous fashion to create a network. Edges of each network were sampled but not included in the analyses to estimate means and variances (Thompson 1990). The configuration of each network was mapped on underwater paper for ease in negotiating the study area. In particular, the number of individuals per 0.25 m^2 unit and number of units within each network were recorded from each study area.

The estimated mean densities and variances were calculated with the Hansen–Hurwitz and the Horvitz–Thompson estimators. For description of estimators refer to Thompson (1990). The total number of sample quadrats at each site (N) was 768. Estimators were calculated using the SAS macro published in Philippi (2005). The ratio of standard errors of estimates from SRS and ACS was used to evaluate the efficiency of ACS (Noon et al. 2006). A ratio larger than one indicates greater efficiency for ACS. The estimated variance from the initial randomly sampled set was adjusted to account for the greater ACS sample size:

$$\hat{\sigma}^2_{\mathrm{SRS-adj}} = \left(1 - \frac{n'_{\mathrm{ACS}}}{N}\right)\frac{\hat{\sigma}^2}{n'_{\mathrm{ACS}}},$$

where n'_{ACS} is the number of units in the ACS sample (including edge units) and $\hat{\sigma}^2$ is the estimate of the population variance for SRS (Noon et al. 2006).

Results

The distribution of the three species differed among islands and with exposure to wave energy. Their distributions were aggregated to varying degrees, as indicated by the number of networks within the study areas, with *C. mamillosum* more evenly distributed (only 1 cluster consisting of >1 unit, Table 1) than *C. pomoides* (0–3 clusters, Table 2) and *H. cuneata* (4–6 clusters, Table 3). Ratios of the sampling errors from SRS and ACS ranged between 0.35 and 0.77, indicating that SRS was more efficient at estimating densities for the three algal species (Tables 1, 2, 3).

The distribution of each species highlighted the strengths and limitations of ACS. *C. mamillosum* was less abundant at wave-sheltered sites with mean densities five to tenfold lower than recorded at wave-exposed Woody Island (Table 1). At wave-sheltered sites, a sampling condition of ≥1 individuals was appropriate at Thomas and Woody Islands and resulted in reasonable network sizes of 3–5 units (Table 1). *C. mamillosum* was distributed throughout the area (present in practically every quadrat sampled randomly) at the wave-exposed sites at Thomas and Black Islands. Even with the sampling condition set at a value far larger than one, ACS resulted in excessively large networks and it was not feasible to sample these sites for *C. mammilosum*. At the wave-exposed side of Woody Island, the number of individuals per quadrat from the initial set (i.e., randomly sampled) ranged between 0 and 9 per 0.25 m^2. The condition for inclusion in the network was ≥6 in order to gain maximum sample efficiency from ACS while minimizing the risk of no adaptive selection of units. In general, the distribution of *C. mamillosum* was not clustered and may be better suited for SRS as indicated by the greater efficiency of SRS estimates (Table 2).

Table 1 Abundance estimates of *Codium mamillosum* per 0.25 m^2 for sites where adaptive cluster sampling (ACS) was possible

	Thomas Island sheltered	Woody Island sheltered	Woody Island exposed
Condition for inclusion in network	≥1 individuals	≥1 individuals	≥6 individuals
Range in initial sample (number of individuals)	0–4	0–3	0–6
Number of networks >1 unit in size (maximum = 16)	1	1	1
Network size	5 units	3 units	3 units
H–H[a] estimator mean (variance)	0.39 (0.065)	0.27 (0.025)	1.96 (0.254)
H–T[b] estimator mean (variance)	0.44 (0.065)	0.27 (0.025)	1.97 (0.254)
Ratio of standard errors for simple random sampling (SRS) and ACS (number of units in ACS sample, including edge units)	0.63 (38 units)	0.46 (38 units)	0.57 (34 units)

[a] Hansen–Hurwitz estimator

[b] Horvitz–Thompson estimator

Similar to *C. mamillosum*, the spatial distribution of *C. pomoides* was not typically clustered and ACS estimates were not more efficient than those of SRS with comparable sample sizes (Table 2). *C. pomoides* was more abundant at Black Island's wave-sheltered site with four times the mean densities compared with the wave-exposed site (Table 2). At wave-exposed sites, the alga was recorded in two quadrats at Black Island, in one quadrat at Thomas Island, and in no quadrats sampled at Woody Island. Density estimates for *C. pomoides* are reported only for Black Island (Table 2). Conditions to initiate ACS were adjusted in order to minimize the potential of excessively large networks. *C. pomoides* had relatively low abundances in the initial sample (up to five individuals per 0.25 m^2) with some networks being fairly large (up to 9 units at Thomas Island) at wave-sheltered sites.

Sampling *H. cuneata* at Black Island with ACS was less efficient than sampling with SRS (Table 3). *H. cuneata* was relatively more abundant at the wave-sheltered site. Setting the sampling condition for ACS was challenging because the numbers of individuals per quadrat in the initial sample set were highly variable, ranging between 0 and 11. Even with the condition set at ≥3 individuals at the exposed site some networks were large (Table 3).

Discussion and conclusions

With the targeted sampling design, the distributions of each species as a function of wave exposure became apparent. *C. mamillosum* was more abundant at wave-exposed sites; *C. pomoides* was more abundant at wave-sheltered sites; and *H. cuneata* had a distribution limited more by island location than by exposure to wave energy. ACS estimates were not lower than those estimated from SRS of equivalent replication for populations that were characterized with aggregated spatial distributions (i.e., *H. cuneata*). The poor performance of ACS may have been a function of low within-network variance with a large number of networks of size one (Thompson and Seber 1996; Noon et al. 2006). In general, networks of >1 unit were not common.

Table 2 Abundance estimates of *C. pomoides* per 0.25 m^2 for sites where ACS was possible

	Thomas Island sheltered	Woody Island sheltered	Black Island sheltered	Black Island exposed
Condition for inclusion in network	≥1 individuals	≥3 individuals	≥3 individuals	≥1 individuals
Range in initial sample (no. of individuals)	0–2	0–5	0–3	0–1
Total number of networks >1 unit in size (maximum = 16)	3	2	0	1
Range in network sizes	2–9 units	3 units	1 unit	2 units
H–H[a] estimator mean (variance)	0.39 (0.032)	1.00 (0.145)	0.79 (0.106)	0.19 (0.018)
H–T[b] estimator mean (variance)	0.41 (0.032)	1.00 (0.145)	0.81 (0.106)	0.19 (0.018)
Ratio of standard errors for SRS and ACS (number of units in ACS sample, including edge units)	0.60 (54 units)	0.35 (47 units)	0.46 (42 units)	0.49 (27 units)

[a] Hansen–Hurwitz estimator

[b] Horvitz–Thompson estimator

Table 3 Abundance estimates of *Halimeda cuneata* per 0.25 m^2 for sites where ACS was possible

	Black Island sheltered	Black Island exposed
Condition for inclusion in network	≥1 individuals	≥3 individuals
Range in initial sample (number of individuals)	0–9	0–11
Number of networks >1 unit in size (maximum = 16)	6	4
Range in network sizes	2–7 units	2–6 units
H–H[a] estimator mean (variance)	2.14 (0.311)	1.71 (0.366)
H–T[b] estimator mean (variance)	2.19 (0.310)	1.70 (0.365)
Ratio of variances for SRS and ACS (number of units in ACS sample, including edge units)	0.44 (96 units)	0.77 (58 units)

[a] Hansen–Hurwitz estimator

[b] Horvitz–Thompson estimator

Because the sampling was conducted subtidally, limitations of ACS became apparent. The extent of the study areas was limited by the time it would take for each diver to sample one species within the allowable dive time in depths <15 m. We used a semi-systematic sampling design to initiate ACS, rather than SRS. This design was successful for divers to sample throughout each study area, a property of a good sampling design as advocated by Christman (2000). In addition, the use of transects gave divers a reference to orient their quadrats in turbulent waters. Any potential gains (or losses) in statistical efficiency from the use of a semi-systematic design were incidental and were not directly measured.

Adaptive cluster sampling is less efficient than SRS for populations that have low abundances and are not clustered (Christman 1997; Brown 2003). We had purposely chosen the three species based on the assumption that they were rare and clustered and yet their densities proved to be less efficiently sampled with ACS than if they had been sampled with SRS. We support Smith et al. (2003) in their call for research in identifying populations that are conducive to ACS and do not require a priori knowledge of the distribution and density of target species.

For all sites where ACS was used the Horvitz–Thompson estimator either gave the same, or smaller variance than the Hansen–Hurwitz estimator. We recommend the use of the Horvitz–Thompson estimator in favor of the Hansen–Hurwitz estimator (Salehi 2003). We also recommend the use of the published SAS macro (Philippi 2005) which easily solved the more computationally-complex Horvitz–Thompson estimator.

The results from this study illustrate the need to be “adaptive” in designing an efficient ACS for multi-species surveys, and for surveys where species’ distributions vary with physical factors such as exposure to wave energy. Smith et al. (1995) observed that different species may not be sampled equally well using the same ACS design and sampling conditions on account of the organisms having different spatial distributions. Understanding the relationship among the different species becomes important in designing efficient ACS strategies (Dryver 2003). The options are to either have multiple designs and hence logistical complexities, or to have one design which overall is optimal but not necessarily for individual species.

We overcame issues of different distributions among the species by adjusting the sampling condition (i.e., the number of individuals per 0.25 m^2 to trigger ACS) for each site. Another parameter of the ACS design that can be altered for sampling the different species at the various sites is the quadrat size. Different quadrat sizes will influence network sizes, sample efficiency, and sampling effort (Christman 1997; Philippi 2005). Philippi (2005) suggests that a quadrat should be larger than the distance between individuals within a cluster. Smaller quadrat size may not have resulted in greater ACS efficiency in our study due to the low number of individuals per 0.25 m^2, suggesting that the populations consisted of a relatively few individuals that were not clustered. Sample units need not be quadrats and line transects (e.g., tows) can be used for marine surveys.

A third parameter that can be altered in ACS design to improve efficiency is the definition of the neighborhood of the adaptive selection. Here we used a neighborhood definition of the four surrounding quadrats but other definitions can be used, e.g., the two or eight surrounding quadrats (Brown 2003), or even non-contiguous quadrats (Smith et al. 2004).

Adaptive cluster sampling provides a means to begin to understand the spatial distribution of rare subtidal species. ACS estimates were not more efficient than SRS of equivalent replication. Although SRS may provide more spatial coverage with substantial increase in effort, the distribution and size of clusters would not be described. For a more comprehensive site-specific investigation of the spatial distribution of rare and clustered populations, a non-adaptive, spatially balanced sampling design such as systematic sampling might provide more information for understanding spatial distribution because sampling effort would be better distributed throughout the site (Christman 2000). Systematic sampling has the advantage of being easier to implement in the field than ACS and does not require prior knowledge of the population of interest (Christman 2000). Future directions of research include studying the persistence and processes that help maintain clusters, investigating the possible role of source–sink dynamics (Dias 1996), role of dispersal and asexual propagation in maintaining populations, and the genetic structure of populations within and among networks, with increasing spatial scale.

Acknowledgments We thank P. Rose for captaining the vessel. We also thank K. Murray and M. Feng who assisted with carrying out the statistics, and G. Kendrick for thoughtful discussions. We thank T. Philippi for publishing the ACS macro. We thank G. Brown for providing support in the field. Previous versions were greatly improved from the insightful comments of three anonymous reviewers and S. Connell. NAG was supported by University of Western Australia Faculty Grant. Sampling conducted in this study complies with the laws of Australia.

References

Brown JA (2003) Designing an efficient adaptive cluster sample. Environ Ecol Stat 10:95–105

Cao Y, Williams DD, Williams NE (1998) How important are rare species in aquatic community ecology and bioassessment? Limnol Oceanogr 43:1403–1409

Chapman MG (1999) Are there adequate data to assess how well theories of rarity apply to marine invertebrates? Biodivers Conserv 8:1295–1318

Christman MC (1997) Efficiency of some sampling designs for spatially clustered populations. Environmetrics 8:145–166

Christman MC (2000) A review of quadrat-based sampling of rare, geographically clustered populations. J Agric Biol Environ Stat 5:168–201

Dias PC (1996) Sources and sinks in population biology. TREE 11:326–330

Drummond SP, Connell SD (2005) Quantifying percentage cover of subtidal organisms on rocky coasts: a comparison of the costs and benefits of standard methods. Mar Freshw Res 56:865–876

Dryver AL (2003) Performance of adaptive cluster sampling estimators in a multivariate setting. Environ Ecol Stat 10:107–113

Goldberg NA, Kendrick GA (2004) Effects of island groups, depth, and exposure to ocean waves on subtidal macroalgal assemblages in the Recherche Archipelago, Western Australia. J Phycol 40:631–641

Hanselman DH, Quinn II TJ (2004) Sampling rockfish populations: adaptive sampling and hydroacoustics. In: Thompson WL (ed) Sampling for rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Washington, pp 271–296

Magurran AE, Henderson PA (2003) Explaining the excess of rare species in natural species abundance distributions. Nature 422:714–716

Noon BR, Ishware NM, Vasudevan K (2006) Efficiency of adaptive cluster and random sampling in detecting terrestrial herpetofauna in a tropical rainforest. Wildl Soc Bull 34:59–68

O'Hara TD (2002) Endemism, rarity and vulnerability of marine species along a temperate coastline. Invertebr Syst 16:671–684

O'Hara TD, Poore GCB (2000) Distribution and origin of southern Australian echinoderms and decapods. J Biogeogr 27:1321–1335

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

Phillips JA (2001) Marine macroalgal biodiversity hotspots: why is there high species richness and endemism in southern Australian marine benthic flora? Biodivers Conserv 10:1555–1577

Salehi M.M (2003) Comparison between Hansen–Hurwitz and Horvitz–Thompson estimators for adaptive cluster sampling. Environ Ecol Stat 10:115–128

Smith DR, Conroy MJ, Brakbage DH (1995) Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. Biometrics 51:777–788

Smith DR, Villella RF, Lemarie DP (2003) Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environ Ecol Stat 10:7–15

Smith DR, Brown JA, Lo NCH (2004) Applications of adaptive sampling to biological populations. In: Thompson WL (ed) Sampling for rare or elusive species: concepts, designs, and techniques for estimating population parameters. Island Press, Washington, pp 77–122

Thompson SK (1990) Adaptive cluster sampling. J Am Stat Assoc 85:1050–1059

Thompson SK (1992) Sampling. Wiley, New York

Thompson SK, Seber GAF (1996) Adaptive sampling. Wiley, New York

Turk P, Brokowski JJ (2005) A review of adaptive cluster sampling: 1990–2003. Environ Ecol Stat 12:55–94

Womersley HBS (1984) The marine benthic flora of Southern Australia. Part I. South Australian Government Printing Division, Adelaide

Womersley HBS (1990) Biogeography of Australasian marine macroalgae. In: Clayton MN, King RJ (eds) Biology of marine plants. Longman Cheshire, Melbourne, pp 368–381