Ranked Set Sampling for Ecological Research: Accounting for the Total Costs of Sampling

Nicolle A. Mode Loveday L. Conquest David A. Marker

NRCSE

Technical Report Series

NRCSE-TRS No. 017

Ranked Set Sampling for Ecological Research: Accounting for the Total Costs of Sampling

by

Nicolle A. Mode

Loveday L. Conquest

David A. Marker

TABLE OF CONTENTS

ABSTRACT

Researchers aim to design environmental studies that optimize precision and allow for generalization of results, while keeping the costs of associated field and laboratory work at a reasonable level. Ranked set sampling is one method to potentially increase precision and reduce costs by using "rough but cheap" quantitative or qualitative information to obtain a more representative sample before the real, more expensive sampling is done. In this report, we investigate under what conditions ranked set sampling becomes a cost-effective sampling method for ecological and environmental field studies where the "rough but cheap" measurement has a cost. Ratios of measuring to ranking costs necessary for ranked set sampling to be as cost effective as simple random sampling, for a common precision, are presented for known distributions with and without ranking error. Cost ratios are also presented for a real data set consisting of visually estimated and physically measured stream habitat areas. Results provide specific guidelines for when ranked set sampling is appropriate, and cost effective, for ecological and environmental field sampling.

1. INTRODUCTION

Societal concern for the state of the environment has shifted ecological research from a strictly scientific pursuit to one with significant societal and legal ramifications. Examples of this shift are visible everywhere. The results of research regarding the effects of human activities upon riverine or estuarine habitats and water quality often appear in the news. More and more frequently, the court system finds itself dealing with legal issues of environmental concern. Policy makers and the general public are demanding well-designed, cost-effective, environmental and ecological studies. Such an environmental study endeavors to optimize precision and allow for generalization of results, while minimizing the costs of associated field and laboratory work. A good design will use available expert knowledge (e.g., about a site, a stream, or a species) to its advantage, qualitatively or quantitatively. In the interest of developing good study designs for ecological researchers and environmental managers, we are investigating issues surrounding the application of an existing, but rarely used method for field sampling that increases precision while reducing costs. The method, known as ranked set sampling (RSS), does this by using initial, "rough but cheap" information to obtain a representative sample before the real, more expensive sampling is done. For populations with patchy distributions that are expensive to sample, the ranked set sampling approach can lead to increased precision, decreased sampling costs, or both. In this paper, we investigate under what conditions RSS becomes a cost-effective sampling method for ecological and environmental field studies where the “rough but cheap” measurement has a cost. In order to provide a more appropriate assessment of RSS under these conditions, we present our method of cost comparisons, provide applications to prior results, and apply the method to data. The paper is organized as follows: Section 2 introduces ranked set sampling, Section 3 describes cost analyses using the total costs of sampling, and Section 4 summarizes the findings.

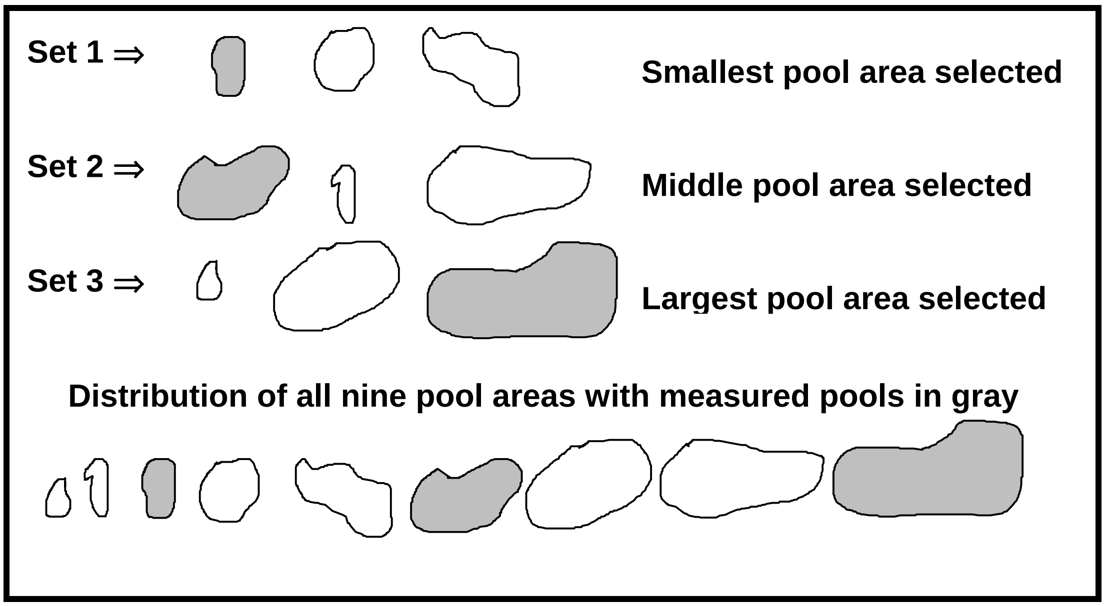
2. RANKED SET SAMPLING

2.1. Description of Ranked Set Sampling

Ranked set sampling (McIntyre 1952) is a two-phase sampling procedure that reduces the number of samples required using a more expensive measurement, termed "costly measurements," by employing expert knowledge or other more economical estimation procedure, termed "frugal measurements," to select samples. The frugal measurement adds information in the form of ranked sets of data. Small sets of samples are ranked using the frugal measurement, and only one sample from each set is measured using the costly measurement. Generally, ranked set sampling involves an initial ranking of n samples of size n (via a frugal measurement), followed by observing (using a more costly measurement) the first order statistic (smallest observation) from the first sample, the second order statistic (second smallest observation) from the second sample, and so on, until the nth order statistic from the nth sample yields a secondary sample of size n from the initial n^2 data points. This process can be repeated to yield large samples from ranking only a few items at a time. Repeating the process m times yields a secondary sample of size nm from an initial n^2m data points.

Estimation of mean pool area for a stream is an ecological study that could possibly benefit from RSS. The U.S. Forest Service measures pool area for Pacific Northwest streams in conjunction with their fish habitat and watershed management programs (U.S.D.A. Forest Service 1997). Pool area can be estimated visually (frugal measurement) or it can be measured more accurately and precisely by a team of field assistants (costly measurement). Ranked set sampling would use the visual estimation to choose a few pools for costly measurement. Suppose a study's budget allows for only three pools to be costly measured (Figure 1). Pools for three sets of size three are randomly chosen from the stream reach. Within each set, the pool areas are ranked using visual examination, and the selected pool is measured by a team of field assistants. The final set of three observations (shaded in gray) is a representative sample of the underlying distribution of pool areas. Ranked set sampling is advantageous for estimating means because it uses information from the nine pool areas to obtain a more representative final sample of three.

Figure 1. Ranked set sample to estimate mean pool area. Pools for three sets of size three are randomly chosen from the stream reach. *Within* each set the pool areas are ranked using visual examination, and the selected pool area is measured accurately. The final set of three observations (shaded in gray) is a representative sample of the underlying distribution of pool areas.



2.2. Background

Ranked set sampling was first described as a method to increase the precision of estimated yield without the bias of researcher-chosen 'representative' samples (McIntyre 1952). The statistical theory was developed, apparently independently, by Takahasi & Wakimoto (1968). RSS has been used to estimate pasture yield (McIntyre 1952; McIntyre 1978), mass herbage in a paddock (Cobby, Ridout, Bassett, & Large 1985), forage yields (Halls & Dell 1966), and shrub phytomass (Martin, Sharik, Oderwald, & Smith 1980; Muttlak & McDonald 1992). More recently it has been recommended for environmental research questions such as estimating plutonium soil concentrations (Gilbert 1995) and quality testing reformulated gasoline (Nussbaum & Sinha 1997). A complete review of applications and theoretical work on RSS will not be made here as many reviews already exist (Patil, Sinha, & Taillie 1994; Kaur, Patil, Sinha, & Taillie 1995; Johnson, Nussbaum, Patil, & Ross 1996).

McIntyre (1952) recognized in his introduction of RSS that the effectiveness of the method was dependent upon the information gained by ranking. In practice, ranking is bound to be performed with some error. Dell & Clutter (1972) showed that the RSS estimator remains unbiased in the presence of unbiased ranking error, and that when ranking is completely random, the RSS estimator has the same precision as the simple random sample estimator. Another interpretation of 'ranking error' is that the ranking is done perfectly but is performed using another variable that is imperfectly correlated with the variable of interest. The correlation between the concomitant variable and the variable of interest is proportional to ranking error. RSS was extended to ranking on a concomitant variable by Stokes (1977).

It is natural to compare RSS to other two-phase or double sampling methods. When restrictive distributional and relational assumptions are satisfied and make the problems tractable, systematic, stratified estimation methods (Patil, Sinha, & Taillie 1993a) and regression estimation methods (Patil, Sinha, & Taillie 1993b) are usually shown to be more efficient than RSS (see Yu & Lam 1997 for counter example). When such assumptions are not satisfied, the robust qualities of RSS make it preferable. Ranked set sampling does not require particular distributional assumptions for an unbiased estimate, and the first-phase variable (ranking) does not even need to be continuous. Thus true comparisons of RSS to other two-phase sampling methods are difficult to present in general terms.

2.3. Notation

The following notation will be used to describe ranked set sampling. Let X_1, X_2, X_3, …, X_n be a random sample of size n from a random variable X with probability density function $f(x)$ and finite mean (μ) and variance (σ^2). Let $X_{i(1)}$ be the first order statistic from the set {X_{i1}, X_{i2},…, X_{in}} which represents the ith random sample of size n. For convenience $X_{i(i)}$ is also written as $X_{(i:n)}$ to denote the ith order statistic from the ith set of n observations with mean $\mu_{(i:n)}$. Let $X_{(i:n)j}$ denote the ith order statistic from the ith sample of size n in the jth cycle (j=1,2, …, m). The unbiased estimator of the population mean is (Takahasi & Wakimoto 1968)

$$\overline{X}_{RSS} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m} X_{(i:n)j}\,.$$

The variance of $\overline{X}_{RSS}$ is

$$\begin{aligned}\operatorname{var}(\overline{X}_{RSS}) &= \frac{1}{n^2 m}\sum_{i=1}^{n} E(X_{(i:n)} - \mu_{(i:n)})^2 \\ &= \frac{\sigma^2}{nm} - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu)^2}{n^2 m}.\end{aligned}$$

Under equal allocation to each order statistic, RSS will always result in as precise an estimate as simple random sampling (SRS) if not better. The degree to which RSS exceeds SRS will depend upon the amount of information gained about the distribution from ranking. One way of comparing RSS to SRS is by relative precision (RP) as defined in survey sampling,

$$\begin{aligned}\mathrm{RP} &= \frac{\operatorname{var}(\overline{X}_{RSS})}{\operatorname{var}(\overline{X}_{SRS})} \\ &= \frac{\dfrac{\sigma^2}{nm} - \dfrac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu)^2}{n^2 m}}{\dfrac{\sigma^2}{k}} \\ &= \frac{k}{nm}\left(1 - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu)^2}{n\sigma^2}\right)\end{aligned}$$

and where k is the sample size under SRS, and in this case k=nm,

(1) $$\mathrm{RP} = \left(1 - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu)^2}{n\sigma^2}\right).$$

This definition of relative precision is consistent with the concept of design effect used in survey sampling (Kish 1965), but is the inverse of RP used in many ranked set sampling papers. As seen in equation (1) the relative precision of RSS to SRS is dependent upon the information gained by ranking relative to the population variance, and the set size *n*. When ranking is completely random and provides no information, the relative precision will be equal to 1. Ranking sets of items that are similar, for example ranking spatially close items when there is a trend on a site, will also drive the RP toward 1, since the variance contained within each set is less than that present in the distribution (Ridout & Cobby 1987).

2.4. Implications of set size on sampling costs

The advantage of ranked set sampling over simple random sampling is either a decrease in sampling costs, an increase in precision, or both. Decreased sampling costs are realized by a decrease in the number of sampling units which are costly measured. Simple random sampling can be viewed in this context as RSS with a set size (n) of 1; where every unit randomly selected is accurately measured.

To give a context for this comparison, sample sizes have been calculated for estimating the mean under several conditions using both RSS and SRS (Table 1). When the variance is large relative to the mean, more samples are needed to attain a given level of relative error for both sampling methods. Since the relative error of the RSS estimate is a function of the information attained by ranking, it is not surprising that the number of samples required decreases with larger set sizes. For example, using SRS with normally distributed data of mean 1 and variance of .25, it takes 43 costly measured samples (nm) to achieve a relative error of 15%. RSS with set size n=3 only requires 24 costly measurements; with n=10 only 10 costly measurements. However, there is a disadvantage to large set sizes. Large set sizes can lead to increased sampling costs because they require a minimum sample size. With a set size of 10 (n), researchers will always need to rank a minimum of 10 sets of 10 (n^2m=100 items) and measure 10 items (nm) even when SRS only requires 4 or 8 items to be measured for the same relative error (see right-hand most column of Table 1). These hidden costs of ranked set sampling are included in the total cost of sampling investigated in the next section.

Table 1. Sample size necessary using ranked set sampling for the relative error of the estimate of the population mean to be less than or equal to the stated value, 95% of the time. Normally distributed data with mean 1, no ranking error. SRS refers to the sample size necessary under the same conditions using simple random sampling, equivalent to RSS with a set size of 1.

Measured Sample Size (nm)

		Relative Error			
	set size (n)	**0.10**	**0.15**	**0.25**	**0.50**
Variance=.25	**SRS**	97	43	16	4
	2	66	30	12	4
	3	51	24	9	3
	4	44	20	8	4
	5	35	20	10	5
	10	30	10	10	10
Variance=.50	**SRS**	193	86	31	8
	2	132	60	22	6
	3	102	45	18	6
	4	84	40	16	4
	5	70	35	15	5
	10	50	20	10	10
Variance=1	**SRS**	385	171	62	16
	2	262	118	42	12
	3	201	90	33	9
	4	164	76	28	8
	5	140	65	25	10
	10	90	40	20	10
Variance=2	**SRS**	769	342	123	31
	2	524	234	84	22
	3	402	180	66	18
	4	328	148	56	16
	5	280	125	45	15
	10	170	80	30	10

3. COST ANALYSIS

3.1. Introduction

When designing a study, several issues are important in deciding to use a particular sampling method. The data must be appropriate for the method, and a balance between sampling costs and the final precision of the estimate should be achieved. For ecological sampling, the balance of costs and precision is particularly important. Many field studies involve both monetary and societal costs (e.g., destructive sampling or harvesting), making the desire for as few samples as possible even more pressing.

Previous assessments of ranked set sampling have focused upon precision and on costs associated with the (*nm*) costly measurements. Those assessments which have briefly addressed the total cost of RSS have done so by methods which are not generally transparent to nonstatistical scientists (Dell & Clutter 1972; Bohn & Wolfe 1994). In ecological sampling, the total cost of RSS will include the cost of ranking the n^2m items since ranking is often not trivial. Ranking will often involve either a preliminary trip to the study site or at least more hours in the field, which may be a remote location. The decision to use RSS will rest on whether the precision gained by ranking is enough to compensate for the extra work of ranking.

To assess the applicability of RSS on ecological field work, we present a method for cost comparisons. We test this method on prior results from simulations, and on actual field data. Costs will be unitless quantities that can be easily adapted to a variety of conditions. Regarding all costs in terms of time is one useful way to integrate costs from varied activities (Sheldon 1984).

3.2. Method

Relative precision has been used to compare RSS to other sampling methods for a given final sample size. This method inherently ignores the effort and cost associated with ranking. A 'vertical' comparison may be made between RSS and SRS for a common final sample size, or number of costly measured items (Figure 2, Line A). RSS and SRS may also be compared for a common level of precision at various sample sizes (Figure 2, Line B). The second comparison, examining the ratio of sample sizes necessary for a common level of precision, allows for the inclusion of costs associated with each level of sampling. From equation (1),

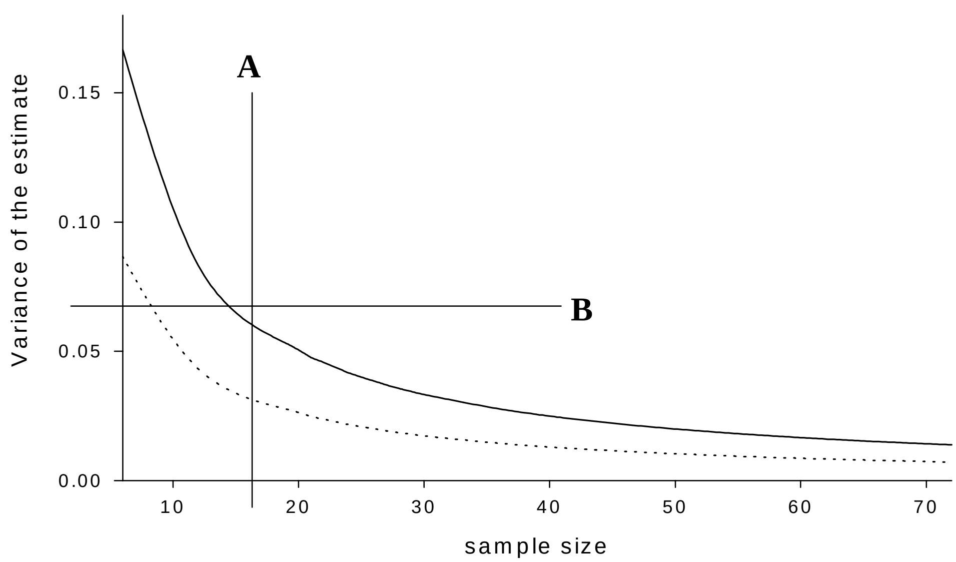
$$\frac{\mathrm{var}(\overline{X}_{RSS})}{\mathrm{var}(\overline{X}_{SRS})} = \frac{k}{nm}\left(1 - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu)^2}{n\sigma^2}\right)$$

and if the variances are equal, then

(2) $$\frac{nm}{k} = (1 - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu)^2}{n\sigma^2}).$$

Thus, the ratio of variances for a common sample size (RP, equation 1), and the ratio of sample sizes for a common precision (equation 2), both result in the same theoretical equation, the right hand side of equation (2). In practice, *n*, *m*, and *k* all need to be integers. Values for the theoretical equation have been previously simulated for several distributions (e.g., Dell & Clutter 1972; Stokes 1977).

Figure 2. Variance of the mean for standard normal data (*N(0,1)*) using both ranked set sampling (---) and simple random sampling (—) . (A) Comparison of sampling methods for a given sample size. (B) Comparison of sampling methods for a given level of precision. Comparisons using (B) allow for inclusion of per unit sampling costs.



The total cost of ranked set sampling includes per unit costs; from both ranking and measuring. All common or fixed costs are assumed to be equal under both sampling methodologies, e.g., travel costs, boundary determination of units. Simple linear cost functions are used, but these may be adapted for specific applications. The total cost for both ranked set sampling and simple random sampling is

$$
\begin{aligned}
\text{Cost}_{\text{RSS}} &= n^2 m(\text{Cost}_{ranking}) + nm(\text{Cost}_{measuring}) \\
\text{Cost}_{\text{SRS}} &= k(\text{Cost}_{measuring}).
\end{aligned} \tag{3}
$$

Combining the equations of cost (3) and precision (2) results in a final function which permits the complete evaluation of RSS to SRS comparisons. For a common level of precision, the total cost under RSS will be less than or equal to the total cost under SRS when,

$$
\begin{aligned}
& n^2 m(\text{Cost}_{ranking}) + nm(\text{Cost}_{measuring}) \le k(\text{Cost}_{measuring}) \\
& n^2 m(\text{Cost}_{ranking}) + nm(\text{Cost}_{measuring}) \le \frac{nm}{(1 - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu.)^2}{n\sigma^2})}(\text{Cost}_{measuring}) \\
& \frac{\text{Cost}_{measuring}}{\text{Cost}_{ranking}} \ge \frac{n}{\left\{(1 - \frac{\sum_{i=1}^{n}(\mu_{(i:n)} - \mu.)^2}{n\sigma^2})^{-1} - 1\right\}}.
\end{aligned} \tag{4}
$$

Thus the decision to use RSS over SRS will depend upon the amount of information gained by ranking and the set size, relative to the ratio of measuring to ranking costs. In practice, the amount of information gained by ranking will be determined by prior historical data or pilot studies. Expected mean values of order statistics ($\mu_{(i:n)}$) for normal and exponentially distributed data are available for most set sizes (Sarhan & Greenberg 1962). Set size is often restricted by physical or accuracy limitations; researchers will need to use prior information or pilot data to determine what is feasible, and what results in low ranking error.

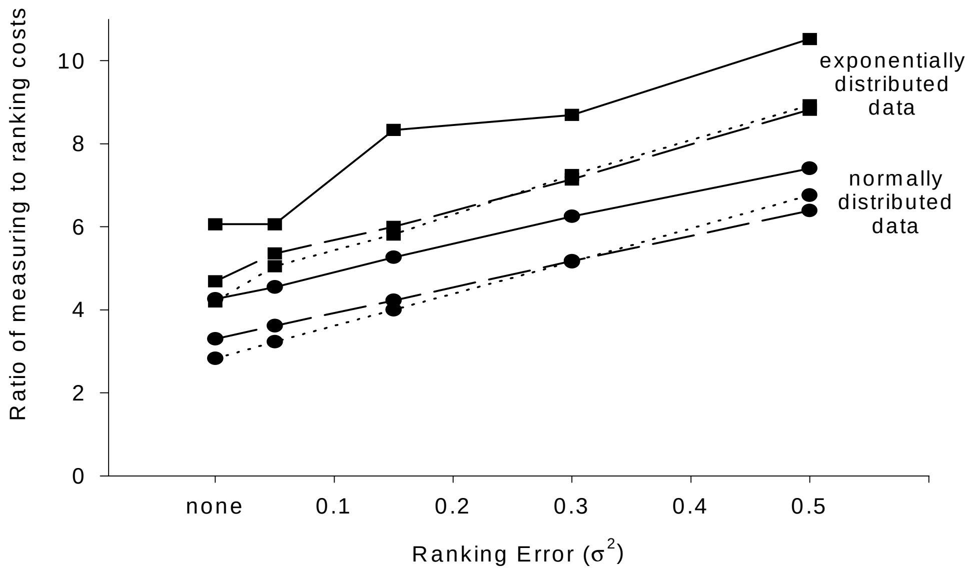
3.3. Application to prior results

Comparing the ratio of sample sizes, instead of variances, allows for the incorporation of previous work on relative precision into current cost functions. Using equation (4) we can derive a ratio of measuring costs to ranking costs such that RSS is as cost efficient as SRS for a given precision.

Two-phase sampling is often used to decrease the overall cost of sampling by using a less costly and less precise measure for the first-phase. Cost differences between the costly and frugal measures can vary substantially. Two methods used to screen for crude oil in contaminated sediment samples had a cost ratio of 5.3 ($450/$85 per sample, Skalski & Word 1994) while two methods for detecting radiation had a cost ratio of about 50 (example in Yu & Lam 1997). The two methods of estimating fish abundance used by Hankin & Reeves (1988) had a cost ratio of 20 (10 person-hours/.5 person-hours per unit).

Dell & Clutter (1972) generated relative precision values for several distributions in the presence of ranking errors. We extended their idea by developing ratios of measuring to ranking costs for different degrees of ranking error (Figure 3). These ratios illustrate the cost ratio necessary for RSS to be just as cost effective as SRS for the same level of precision. Values are presented for both normally and exponentially distributed data with increasing levels of normal (unbiased) ranking error. For example, to sample normally distributed data, ranking sets of three items at a time without ranking error, measuring needs to be about three times the cost of ranking for RSS to be as cost effective as SRS. If measuring is more than three times the cost of ranking, then RSS is less expensive than SRS for a common level of precision. The cost ratio necessary for equal total sampling costs is higher for exponentially distributed data since the distribution is skewed (for known skewed distributions see Kaur, Patil, & Taillie 1997, for RSS with unequal allocation).

Figure 3. Ratio of measuring to ranking costs for different degrees of ranking error. Ratio based upon relative precision values from Dell and Clutter (1972, Table 2) and David and Levine (1972, Table A1). Ranking error modeled as normally distributed (μ=0, σ^2) for set sizes of 2 (——), 3 (- - - -), and 5 (⋯⋯). Cost ratio necessary for RSS to be cost efficient is larger in the presence of ranking error, and is larger for exponentially distributed data (■) than normally distributed data (●).

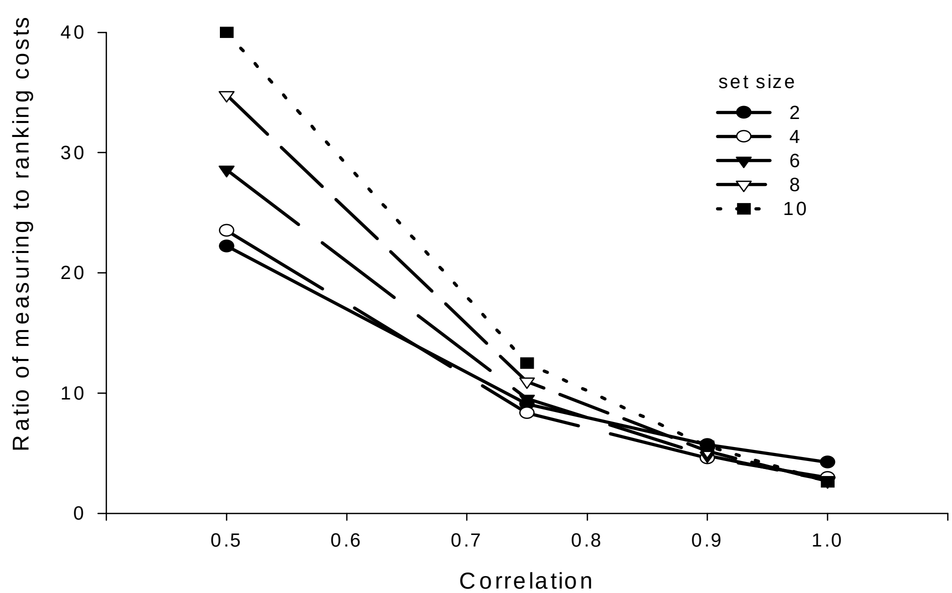


RSS increases in relative precision with increasing set size (*n*) as shown in equations (1) and (2). Some authors have suggested that set size should be chosen as large as is practical due to the beneficial effect on precision (Patil et al. 1994). When ranking is not perfect and has a cost, there is a point of diminishing returns with increasing set size. At high levels of ranking error (Figure 3, e.g., σ^2=.50) a set size of 5 has a higher necessary cost ratio than a set size of 3 for equal total costs among SRS and RSS, even though a set size of 5 results in greater precision (for a given final sample size, *nm*). The diminishing return results when the added precision of increased set size is lowered due to ranking error.

An earlier paper by Stokes (1977) adapted ranked set sampling for use with a concomitant, or correlated ranking variable which differs from the measured variable. We extended her results and developed ratios of measuring to ranking costs for different degrees of correlation between the measured and concomitant variables (Figure 4). These ratios are the cost ratios necessary for

RSS, ranking on a correlated variable, to be as cost effective as SRS on the measured variable. Stokes used a model where the relationship between the concomitant and measured variable is linear, and the variables are distributed bivariate normal. As the two variables become less related, the information gained by ranking on the concomitant variable also lessens. Ultimately at low correlations, a higher cost ratio is necessary for RSS to be as cost effective as SRS. The effect of diminishing returns with increasing set size and ranking error, is even more evident under these conditions. When the correlation between ranking and measuring variables is as low as, say, .50, a set size of two requires a cost ratio of 22, a set size of six requires a cost ratio of 29, and a set size of ten requires a cost ratio of 40. The added cost of ranking more items exceeds the benefits of increased precision with increased set size.

Figure 4. Ratio of measuring to ranking costs for different degrees of correlation between measured and concomitant variables. The variables are linearly related and distributed bivariate normal. Ratios are based upon relative precision values from Stokes (1977) and David & Levine (1972). When there is low correlation between the measuring and ranking (concomitant) variables, the increased cost of ranking more items outweighs the increased precision with larger set sizes.



These results can be used by researchers to determine if RSS is appropriate for their ecological study. If the person-hours required for a costly measurement is about six times of that required for a frugal measurement, and past data sets have been fairly normally distributed, then RSS will be more cost effective than SRS unless the chosen ranking method will result in substantial ranking error (Figure 3), or is based on a concomitant variable that is not very highly correlated (Figure 4). When ranking 8 or fewer things at a time on a concomitant variable highly correlated with the variable of interest ($\rho \geq .75$), with normally distributed data, measuring needs only be 11 times the cost of ranking. Ranked set sampling has been suggested when measurement is costly and when ranking a small set can be done frugally. These graphs give actual definitions to costly and frugal for researchers to use.

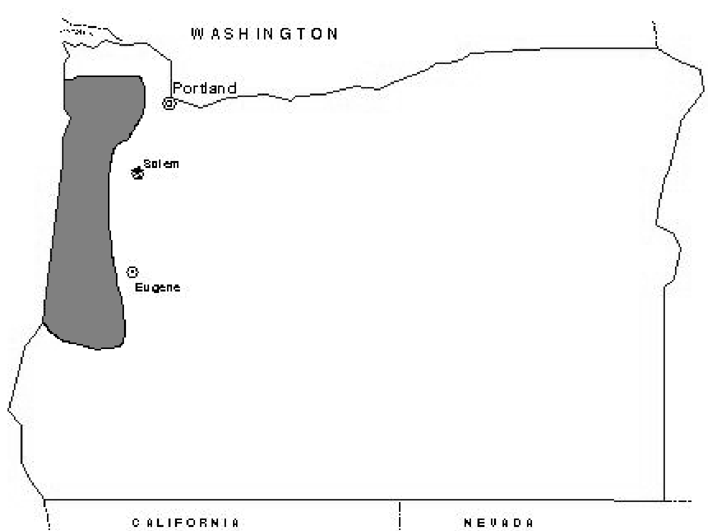
3.4. Application to stream habitat data

Results for known distributions provide useful guidelines in comparing methods and planning experiments. Ecological data may not follow standard distributions, and depending on what variables are measured, standard distributions may be modified by spatial or temporal trends. To be properly assessed for ecological data, RSS needs to be tested on naturally occurring distributions.

The USDA Forest Service collects data on Pacific Northwest streams as part of a large scale monitoring project (U.S.D.A. Forest Service 1997). A variable of interest is habitat size, particularly pool area, which has been linked to salmon production (e.g., Sharma 1998). Measuring habitat areas accurately and precisely is time consuming and labor intensive (see Poole, Frissell, & Ralph 1997 for discussion of issues). In an effort to make both habitat assessment and fish estimation more cost effective, Hankin & Reeves (1988) outlined a sampling methodology based on ratio estimation (see Cochran 1977, pp. 150-186). Many stream segments are measured visually (frugal measurement), and then a subset is also physically measured by a team of field assistants (costly measurement). The subset, with both measurements, is used to determine a correction or bias factor with which to scale the other only visually measured observations. This method results in a best linear unbiased estimator (BLUE) when the two variables (visual and physical measurements) are linearly related and when the variance of the visual measurement increases proportionally with the value (Cochran 1977). The method has the greatest utility in practice when visual measurements are consistently over or consistently under the physical measurements made by the team of field assistants (Conquest, Cardoso, Seidel, & Ralph 1991).

Visual and actual habitat length and width measurements from 21 coastal Oregon streams were obtained from the USDA Forest Service. The streams are mostly in forested areas which drain into the Pacific Ocean from and including the Umpqua River Basin north to the Columbia River Basin Boundary (Figure 5). Each stream has between 36 and 108 habitat areas (median 50) which were both visually estimated and physically measured. Data consists of visually estimated and physically measured length and width of each habitat which can be used to calculate areas. Resampling (with replacement) was used to assess the relative precision of RSS to SRS on the distributions of stream habitat measurements. Each stream was treated as an empirical distribution and 4000 random, independent samples for each were drawn for both RSS and SRS sampling methods. Relative precision values were calculated using Equation (1). Since RP depends only on set size (n) and the distribution of the data, results are only given for nm=12. RP values for other final sample sizes (nm), for a given set size, are the same. Estimated RP values were used with Equation (4) to calculate cost ratios.

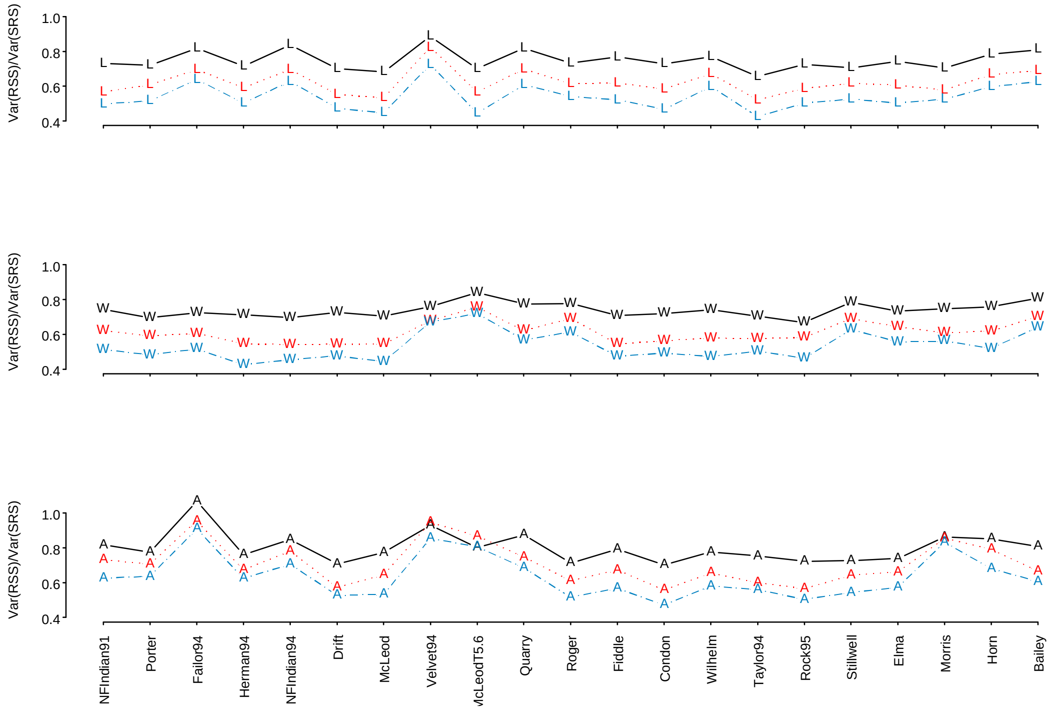
Figure 5. Map of Oregon showing river basins included in stream data set (shaded area).



Mean length and width habitat estimates using RSS were more precise than estimates using SRS for all 21 streams (RP range(0.42,0.88), mean=0.63, Figure 6). For a given final sample size, samples with larger set sizes and fewer cycles (e.g., n=4 & m=3 versus n=2 & m=6) had greater precision relative to SRS. The difference in RP values across the streams was due to differences in the distributions, mainly the skewness, of the 21 data sets. Data for the McLeod tributary and Velvet stream reach (McLeodT5.6 and Velvet94, Figure 6) were heavily skewed due to extreme data points. These values caused a multimodal distribution of the sample mean after 4000 simulations for both SRS and RSS, and resulted in the largest set size (n=4) having approximately the same RP value as the medium set size (n=3) due to random variation. Under equal allocation, RSS had greater RP for symmetric distributions versus skewed distributions but in both conditions was superior to SRS.

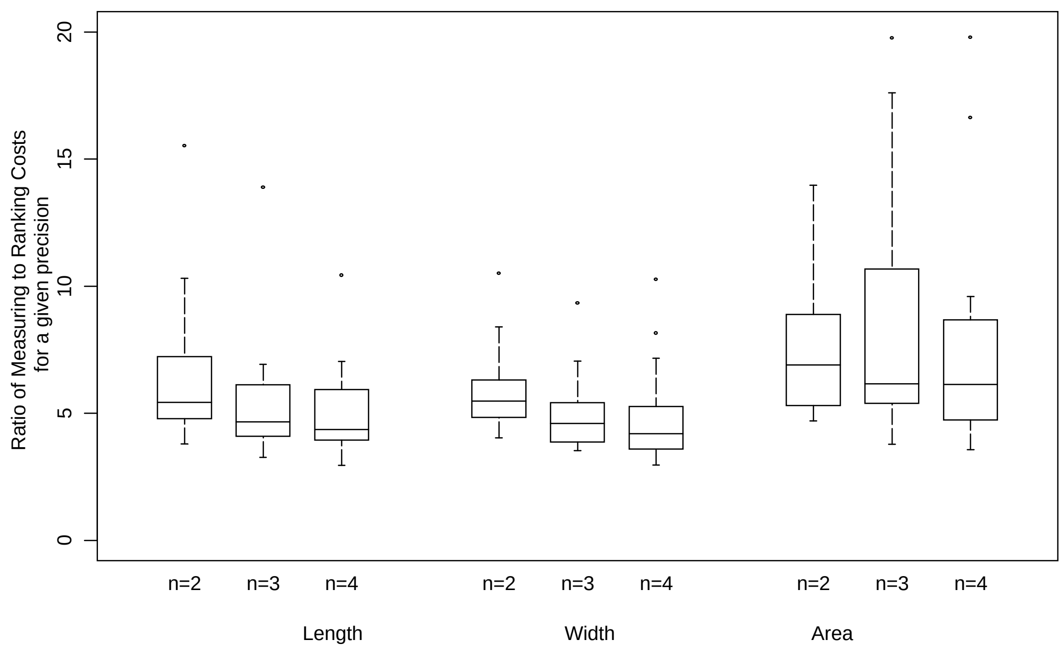
Mean area estimates varied across the streams, and showed different patterns than the length and width variables (Figure 6). Visual estimates and physical measurements of area were calculated as the product of length and width variables. Both visually and physically measured areas showed extreme points for Failor94, Velvet94, Morris, and McLeodT5.6. Failor94 had an outlying point in length, which when converted to area was 6000 units above the next largest value for that stream. In the presence of such extreme values, generally all sampling schemes do poorly and in this case the RP value was 1. Since area was the product of two variables, multimodal distributions of the estimates due to extreme values were more prevalent for both SRS and RSS. Some of these problems might be remedied with more simulations, but such results would not be readily applied to real situations with only one sample.

Figure 6. Ratio of variances for estimated mean habitat length (L), width (W), and area (A=L*W) for each Oregon stream based on 4000 resamplings. Values are presented for set sizes of 2 (——), 3 (·····), and 4 (- - - -). Lower values indicate that ranked set sampling (RSS) resulted in a more precise estimate than simple random sampling (SRS).



For ecological sampling, a more appropriate comparison is one including costs of ranking and measuring. Using equation (4), cost ratios were calculated for each of the 21 streams under each sampling condition (set size) such that RSS was as cost efficient as SRS and resulted in the same precision (Figure 7). The benefit of increased set size was evident by the lower measuring to ranking cost ratio needed for RSS to be cost effective. Due to low ranking error associated with these data (correlation(visual, physical) length=.96, width=.93, area=.97), the precision gained by increased set size was greater than the cost associated with ranking more items. Cost ratios for estimating area were generally larger than those for estimating either width or length, and did not show a dramatic decrease with larger set sizes. These results are consistent with previous work showing that increased, or in this case compounded, error requires a higher measuring to ranking cost ratio to be cost effective relative to SRS (Figure 3). Although for most individual streams an increase in set size increased precision for estimating area (Figure 6), streams differed substantially on how much RSS actually improved precision. The difference in precision levels across the streams resulted in a larger range of cost ratios. For over 75% of the streams, under all conditions, RSS was more cost effective than SRS when measuring costs were at least 11 times that of ranking costs. A preliminary study on Taylor River in Washington found actual cost ratios for estimating mean habitat size to be between 5.8 and 8.2 (mean=7.2, 4 replications, N.Mode unpublished data).

Figure 7. Boxplot of measuring to ranking cost ratios for estimating mean habitat length, width, and area under three set size (*n*) conditions. Boxplots display the median (line) and interquartile range (box) of costs ratios calculated for each of the 21 streams. Two outlying points are not shown for each set size for area (Failor94, Velvet94). For over 75% of the streams, under all conditions, RSS is more cost effective than SRS when measuring costs are at least 11 times that of ranking costs.



4. SUMMARY

The balance of cost and precision is a complex problem for any research study. Ranked set sampling provides a methodology for incorporating additional information into the sampling framework. Unlike other two-phase sampling methods which use “extra” information, ranked set sampling can incorporate non-quantitative information such as expert opinion, and is robust to departures from non-linear relationships among the variables. When ranking is not trivial, as in many types of field work, the precision gained from the “extra information” must be balanced against the costs. The cost analyses we presented allow RSS to be assessed using the total costs of sampling.

Cost ratios of measuring to ranking costs present comparisons of ranked set sampling to simple random sampling which are readily accessible to researchers and managers. The effects of ranking error on known distributions, either by a less precise measure of the same variable or by a concomitant variable, can be assessed in terms of costs. Ranking error reduces the RSS precision advantage over SRS, and thus a larger cost ratio is needed for the two methods to be equally cost effective. Although increased set size always increases precision, even in the presence of ranking error, there is a point of diminishing returns when ranking costs are taken into account. The diminishing return results when the added precision of increased set size is lowered due to ranking error.

Actual cost ratios necessary for RSS to be as cost effective as SRS for the same precision of an estimated mean are consistent with many ratios currently found in two-phase sampling. For known exponentially or normally distributed data without ranking error, measuring needs to be at most 6 times the cost of ranking. Even with substantial ranking error and an exponential (skewed) distribution, measuring need only be 11 times the cost of ranking. This value also held true with the skewed stream habitat data. When ranking 8 or fewer things at a time on a concomitant variable highly correlated with the variable of interest ($\rho \geq .75$), with normally distributed data, measuring also need only be 11 times the cost of ranking.

ACKNOWLEDGEMENTS

The authors would like to thank the National Research Center for Statistics and the Environment at the University of Washington for supporting this work, Barry Nussbaum of the United States Environmental Protection Agency for his contributions in shaping the focus of our work, Shaun McKinney of USDA Forest Service, Pacific Northwest Region for providing the stream data, and Owen Hamel and Marianne Turley for improving the quality of the report. Although the research described in this article has been funded in part by the U. S. Environmental Protection Agency through agreement CR825173-01-0 to the University of Washington, it has not been subjected to the Agency’s required peer and policy review and therefore does not necessarily reflect the views of the Agency and no official endorsement should be inferred. L. Conquest received additional support from the U.S. Environmental Protection Agency's National Health and Environmental Effects Research Laboratory, Western Ecology Division, through cooperative agreement CR824682.

REFERENCE LIST

Bohn, L. L., & Wolfe, D. A. (1994). The effect of imperfect judgment rankings on properties of procedures based on the ranked-set samples analog of the Mann-Whitney-Wilcoxon statistic. Journal of the American Statistical Association, 89(425), 168-176.

Cobby, J. M., Ridout, M. S., Bassett, P. J., & Large, R. V. (1985). An investigation into the use of ranked set sampling on grass and grass-clover swards. Grass and Forage Science, 40, 257-63.

Cochran, W. G. (1977). Sampling Techniques. New York: John Wiley and Sons.

Conquest, L. L., Cardoso, T. P., Seidel, K. D., & Ralph, S. C. (1991) Using visual estimation in a watershed monitoring study. Appendix C. Status and Trends of Instream Habitat in Forested Lands of Washington: The Timber-Fish-Wildlife Ambient Monitoring Program, Biennial Progress Report, Center for Streamside Studies, Seattle, Washington.

David, H. A., & Levine, D. N. (1972). Ranked set sampling in the presence of judgment error. Biometrics, 28, 553-5.

Dell, T. R., & Clutter, J. L. (1972). Ranked set sampling theory with order statistics background. Biometrics, 28, 545-53.

Gilbert, R. O. (1995) Ranked set sampling. DQO Statistics Bulletin: Statistical Methods for Data Quality Objective Process, PNL-SA-26377, The Pacific Northwest Laboratory.

Halls, L. K., & Dell, T. R. (1966). Trial of ranked set sampling for forage yields. Forest Science, 12, 22-6.

Hankin, D. G., & Reeves, G. H. (1988). Estimating total fish abundance and total habitat area in small streams based on visual estimation methods. Canadian Journal of Fisheries and Aquatic Sciences, 45, 834-844.

Johnson, G. D., Nussbaum, B. D., Patil, G. P., & Ross, P. N. (1996). Designing cost-effective environmental sampling using concomitant information. Chance, 9(1), 4-11.

Kaur, A., Patil, G. P., Sinha, A. K., & Taillie, C. (1995). Ranked set sampling: An annotated bibliography. Environmental and Ecological Statistics, 2, 25-54.

Kaur, A., Patil, G. P., & Taillie, C. (1997). Unequal allocation models for ranked set sampling with skew distributions. Biometrics, 53, 123-130.

Kish, L. (1965). Survey Sampling. New York: John Wiley and Sons.

Martin, W. L., Sharik, T. L., Oderwald, R. G., & Smith, D. W. (1980) Evaluation of Ranked Set Sampling for Estimating Shrub Phytomass in Appalachian Oak Forests, Publication Number FWS-4-80, Blacksburg, Virginia: School of Forestry and Wildlife Resources, Virginia Polytechnic Institute and State University.

McIntyre, G. A. (1952). A method for unbiased selective sampling, using ranked sets. Australian Journal of Agricultural Research, 3, 385-90.

McIntyre, G. A. (1978). Statistical aspects of vegetation sampling. Measurement of Grassland Vegetation and Animal Production (pp. 8-21). Hurley, Berkshire, UK: Commonwealth Bureau of Pastures and Field Crops.

Muttlak, H. A., & McDonald, L. L. (1992). Ranked set sampling and the line intercept method: A more efficient procedure. Biometrical Journal, 34(3), 329-346.

Nussbaum, B. D., & Sinha, B. K. (1997). Cost effective gasoline sampling using ranked set sampling. ASA Proceedings of the Section on Statistics and the Environment (pp. 83-87).

Patil, G. P., Sinha, A. K., & Taillie, C. (1994). Ranked set sampling. G. P. Patil, & C. R. Rao (eds), Handbook of Statistics (pp. 167-200). New York: North-Holland.

Patil, G. P., Sinha, A. K., & Taillie, C. (1993a). Ranked set sampling from a finite population in the presence of a trend on a site. Journal of Applied Statistical Science, 1(1), 51-65.

Patil, G. P., Sinha, A. K., & Taillie, C. (1993b). Relative precision of ranked set sampling: a comparison with the regression estimator. Environmetrics, 4(4), 399-412.

Poole, G. C., Frissell, C. A., & Ralph, S. C. (1997). In-stream habitat unit classification:

Inadequacies for monitoring and some consequences for management. Journal of the American Water Resources Association, 33(4), 879-896.

Ridout, M. S., & Cobby, J. M. (1987). Ranked set sampling with non-random selection of sets and errors in ranking. Applied Statistics, 36, 145-52.

Sarhan, A. E., & Greenberg, B. G. (1962). Contributions to Order Statistics. New York: Wiley.

Sharma, R. (1998). The Influence of Habitat on Smolt Production of Coho Salmon in 14 Western Washington Streams. M.S. thesis, Quantitative Ecology and Resource Management, University of Washington, Seattle, Washington.

Sheldon, A. L. (1984). Cost and precision in a stream sampling program. Hydrobiologia, 111, 147-152.

Skalski, J. R., & Word, J. Q. (1994). Detection of contaminated samples using a two-way compositing scheme. Environmental Toxicology and Chemistry, 13, 15-19.

Stokes, S. L. (1977). Ranked set sampling with concomitant variables. Communications in Statistics - Theory and Methods, A6, 1207-11.

Takahasi, K., & Wakimoto, K. (1968). On unbiased estimates of the population mean based on the sample stratified by means of ordering. Annals of the Institute of Statistical Mathematics, 20, 1-31.

U.S.D.A. Forest Service. (1997) Stream Inventory Handbook: Level I & II, Version 9.7(.

Yu, P. L. H., & Lam, K. (1997). Regression estimator in ranked set sampling. Biometrics, 53, 1070-1080.