US ENVIRONMENTAL PROTECTION AGENCY

Aquatic Survey Design and Analysis

A guide for practitioners

Anthony R. Olsen

8/13/2013

Organizations that monitor aquatic resources are incorporating probability survey designs as one component in their approach to monitoring. This document is a practitioner's guide to topics that arise when designing a probability survey and the subsequent statistical analyses of data from the survey.

Table of Contents

List of Tables

1 Introduction

The objective of this document is to serve as a practitioner's guide to probability survey design and analysis for aquatic resource monitoring. We assume that readers will have a rudimentary understanding of probability survey design and analysis. We also do not present basic finite population survey design and analysis statistical theory, such as that covered by Lohr (1999), Thompson (1992), or Cochran (1987). Readers may have read books, e.g., Gitzen, Millspaugh, Cooper, et al. (2012) or Gilbert (1987), that provide general information on the design and analysis of environmental monitoring studies. Our intent is to cover topics that typically arise during the design stages of an aquatic resource survey design and that arise in completing statistical analyses of data from those surveys.

Design topics include understanding the concept of a target population applied to aquatic resources, potential options for sampling frames that represent the target population and are the basis for selecting sample units, the concept of spatial-balance and why it is important to consider in the choice of a survey design, a monitoring framework that can be used to ensure the survey design addresses the objectives of the aquatic resource monitoring study, and decisions concerning sample size.

Analysis topics include weight adjustment to reflect the actual implementation of the survey design and to account for sample units that could not be sampled.

2 Design weights and their adjustment

Survey weights are essential in producing estimates of aquatic resource characteristics based on data from an aquatic probability survey design. For example, an estimate for a population total has the form $\hat{t} = \sum_{i=1}^{n} w_i y_i$ where y_i is a response for the ith sample unit and w_i is the corresponding analysis weight. Without the weight the total only reflects the sum of the response for the particular sample and does not reflect the population total. Similarly, an estimate for a population mean has the form $\hat{t} = \sum_{i=1}^{n} w_i y_i / \sum_{i=1}^{n} w_i$. For a simple random sample, or any equal probability sample, of a finite population the weight for each sample unit will be the same, w_i = N/n, so that the estimate for the population mean is the simple unweighted sample mean and the population total is N times the sample mean. In this case the analysis weights can be thought of as the number of population units that a sample unit represents. For example, if an equal probability survey design is used to select 100 lakes from a population of 1,000 lakes, then each sample lake "represents" 10 lakes. If a probability sample is not an equal probability sample, then the weights must be taken into account if the objective is to estimate the population total or mean.

This chapter discusses the calculation of weights appropriate for use in the analysis of data from a probability survey design. The approach follows a similar discussion in the book by Valliant, Dever and Kreuter (2013). Determining the analysis weight may involve a series of steps in most aquatic surveys. The survey design includes the computation of the design weights. If a survey design is implemented exactly as planned and no other adjustments are made to match known population characteristics, the design weights are the analysis weights. Otherwise the calculation of the analysis weights not only includes calculating the design weights, but may also include adjusting the design weights to reflect how the design was actually implemented (e.g., for use of over-sample units), for the unknown eligibility of a sample unit, for nonresponse of sample units, and to correct for sampling frame deficiencies. The sections that follow describe the calculations required to obtain the analysis weights for each of these situations.

Sampling frames likely contain errors (e.g., sample units that might not be part of the target population), or contain sample units whose eligibility status cannot be determined (in or out of the target population?). Some eligible sample units cannot be sampled (e.g., landowners deny access, or the unit is too dangerous to visit). If a monitoring program requires a specific number of sample units be sampled, survey designs are planned to select a sufficient number of sample units so that the number of sample units that actually are sampled (i.e., respond) meets the requirement, i.e., an over-sample is selected. Either all the sample units are evaluated and sampled if possible or the evaluation is completed in a way that ensures the evaluated sample units result in a probability sample. In the latter case, the sample units are evaluated and sampled until the specific number of sample units is eligible and sampled. That is, design weights must be adjusted to reflect the design as-implemented.

These idiosyncrasies or practicalities of aquatic resource monitoring generally require adjusting weights so that the units sampled have weights that sum up to the full eligible population. The following sets are defined to ease the weight adjustment process:

s = set of all sample units and n is number of sample units in set,
s_{IN} = set of sample units in s that are known to be ineligible and n_{IN} is the number of sample units in s_{IN},
s_{ER} = set of sample units in s that are known to be eligible that have a response (sampled) and n_{ER} is the number of sample units in s_{ER},
s_{ENR} = set of sample units in s that are known to be eligible that have no response (not sampled) and n_{ENR} is the number of sample units in s_{ENR},
s_{UNK} = set of sample units for which eligibility is unknown and n_{UNK} is the number of sample units in s_{UNK},
s_{NN} = set of sample units not evaluated, i.e. not needed, and n_{NN} is the number of sample units in s_{NN},
s_E = set of all sample units known to be eligible (combination of s_{ER} , s_{ENR}, i.e., $s_E = s_{ER} \cup s_{ENR}$) and $n_E = n_{ER} + n_{ENR}$,
s_{KN} = set of of sample units where eligibility is known (combination of s_{IN}, s_{ER} , s_{ENR}, i.e., $s_{KN} = s_{IN} \cup s_{ER} \cup s_{ENR}$) and $n_{KN} = n_{IN} + n_{ER} + n_{ENR}$

These sets of sample units are used in the adjustment of weights for as-implemented designs, unknown eligibility of sample units, and nonresponse of sample units. Note the total sample size $n = n_{IN} + n_{ER} + n_{ENR} + n_{UNK} + n_{NN}$.

Some weight adjustment approaches require making assumptions about why a sample unit may be ineligible or why a sample unit may not be sampleable. Little and Rubin (2002) introduced the concepts of missing completely at random (MCAR), missing at random (MAR) and nonignorable nonresponse (NINR). MCAR means that the sample unit's unknown eligibility or reason for not responding is independent of any information known about the sample unit or its response. When sample units are MCAR, the eligible sampleable units are representative of the selected sample units. MAR means that the sample unit's ineligibility or nonresponse does not depend on the response but depends only on information that is known about the sample unit. In this case, it is possible to model the dependence. For unknown eligibility, the information must be known for all units in the sample so that only information available in the sampling frame can be used. For nonresponse, the information must be known for eligible sample units that are sampleable and that are not sampleable. NINR means that the probability of a sample unit responding depends on one or more of the response variables and that this dependence cannot be removed based on information that is known for both sampleable and nonsampleable sample units.

2.1 Description of examples

The major sections are organized by the type of weight adjustment: the calculation of design weights (section 2.2), adjustments for as-implemented designs, i.e., over-sample units (section 2.3), adjustments for unknown eligibility of sample units (section 2.4), for nonresponse of sample units (section 2.5), and for correction of sampling frame deficiencies (Section 3). Within each major section a general overview of the weight adjustment process is given. This is followed by examples (1) that sample lakes as points (e.g., the National Lake Assessment), (2) that sample streams as linear networks (e.g., the National Rivers and Streams Assessment), and (3) that sample coastal waters as areas (e.g., the National Coastal Condition Assessment). Sections on design weights initially give an example for an equal probability survey design which is then followed by an unequal probability design and a stratified design. The design weight section includes a discussion on the calculation of weights when a master sample is the basis for the design. The final section discusses the calculation of analysis weights for the situation where two or more independent aquatic survey designs will be combined. First, we provide an overview of the examples that are used in this section.

2.1.1 Example Lake Finite Population

Assume a survey design is required for a monitoring program of a population of lakes within a geographic region, e.g. a state. Also assume that an explicit, written definition of the target population is available that enables a decision to be made whether a water body is eligible to be a member of the target population. Further assume that a GIS layer is available that can be used as a sampling frame where it is known that the GIS layer may include potential sample units that do not meet the definition of a lake for the monitoring program. Note that the GIS layer is a point layer which is likely derived from the polygon GIS layer of the lakes. The example finite population designs in this chapter are based on a monitoring program of lakes within Oregon. The monitoring program's general objective is to characterize the population of lakes and reservoirs (hereafter called lakes) within Oregon. The sampling frame is constructed from the National Hydrologic Database (NHD), specifically NHDPlus (NHDPlus 2013). It is known that the sampling frame contains sample units that do not meet the definition of a lake for the program. The sampling frame includes 1831 sample units (Table 2.1.1) that are categorized by five lake surface area categories and nine Omernik Level III ecoregions (Omernik 1987) which are organized by two aggregated ecoregions (Mountains and Lowlands).

Table 2.1.1-1 Sampling frame summary of Oregon lakes by lake area categories, Omernik level III ecoregions and aggregated ecoregions

	Lake Area Categories					
Omernik Level III ecoregions	< 1 ha	1 to 10 ha	10 to 100 ha	100 to 1000 ha	>1000 ha	Total
Mountains						
Blue Mountains	6	151	36	3	2	198
Cascades	15	415	75	17	7	529
Coast Range	8	81	35	12	1	137
Eastern Cascades	25	147	64	17	9	262
Klamath Mountains	4	25	3	0	0	32
Total	58	819	213	49	19	1158

Lowlands						
Columbia Plateau	4	17	1	1	0	23
Northern Basin and Range	50	204	81	21	7	363
Snake River Plain	0	3	1	0	0	4
Willamette Valley	15	220	41	6	1	283
Total	69	444	124	28	8	673
Total	127	1263	337	77	27	1831

2.1.2 Example Stream Linear Network Population

Assume a survey design is required for a monitoring program of streams and rivers within the state of Oregon. The monitoring program has an explicit, written definition that enables a decision to be made whether a stream or river channel is eligible to be a member of the target population. For simplicity, stream will refer to stream and river in the future. A critical decision by the monitoring program is to define the target population as a linear network and not as a collection of stream segments. That is, a sample element is any point on the linear network with any response measured assumed to vary continuously along the network. The monitoring program has a sampling frame that is constructed from the National Hydrologic Database (NHD), specifically NHDPlus (NHDPlus 2013). It is known that the sampling frame includes points on the stream network that do not meet the target population definition, i.e., is an over-coverage. It is also known that the sampling frame excludes streams that do meet the target population definition. The latter results in an under-coverage of the target population by the sampling frame. No inferences about streams in the under-coverage can be made. While the sampling frame consists of a collection of stream segments, a sample unit is any point on the linear network. Consequently, the survey design is for a continuous population within a bounded area and is not a finite population sample.

The sampling frame includes 88,115 km of stream length (Table 2.1.2-1). Streams are categorized by mountain and lowland ecoregions, small and large streams, and type of land ownership.

2.1.2-1 Sampling frame stream and river length (km) by ecoregion, stream size and land ownership

Ecoregion	Land ownership					
Stream size	Non-federal	USFS	BLM	Other federal	Tribal	Total
Mountain						
Small	26413	21842	5163	353	505	54276
Large	10983	4220	1404	81	344	17033
Total	37396	26063	6566	435	849	71309
Lowland						
Small	7472	7	1517	414	9	9420
Large	5628	3	1444	251	60	7386
Total	13100	10	2961	666	69	16806
Total						

	Small	33885	21849	6680	768	514	63696
	Large	16611	4223	2847	333	405	24419
	Total	50496	26073	9527	1101	919	88115

2.1.3 Example Coastal Water Area Population

2.2 Design Weights

Probability survey designs specify the probability that a sample unit will be selected. This probability, π_i , is called the sample unit's inclusion probability for finite populations or inclusion density for linear network or area populations. The inverse of the inclusion probability is the design weight $w_{di} = 1/\pi_i$. For survey designs that have a fixed sample size, n, the sum of the inclusion probabilities over the entire population equals the sample size. The population extent for a finite population, e.g. lakes, is the number of population units, N. The population extent for a linear network is the total length of the linear network and for an area the total area of the sampling frame. The sum of the weights over the units in a sample equals the population size or extent. This identity provides a check on whether the weights have been calculated correctly. If the sum of the weights does not equal the population extent, then you know that a mistake was made in calculating the weights. It is important to recognize that weights have units. For example, a finite population of lakes has weights where the units are number of lakes, a linear network of streams may have units of km and an area population of coastal waters may have units of hectares.

2.2.1 Equal Probability Design

For a finite population suppose that a sample of size n is selected with equal probability without replacement from a sampling frame that has N sample units. In this case the design weight is equal to $w_i = N/n$. For the lake example, if a sample size of 300 is selected, then the design weight would be w_i = 1831/300 = 6.10333 for all units in the sample. That is, each lake in the sample represents 6.1033 lakes. For a linear network the extent of the sampling frame is the network's total length, L, resulting in design weights w_i = L/n. A sample of size 300 for the Oregon stream example results in design weights w_i = 88115/300 = 440.5735 km. For a sampling frame consisting of polygonal areas, the extent of the sampling frame is the total area, A, of the polygons resulting in design weights w_i = A/n.

2.2.2 Unequal Probability Design

Many surveys select sample units with unequal probability, e.g., small lakes may have a smaller probability of being selected that large lakes. Determining the design weights depends on knowing the specific algorithm used to complete the selection and a number of potential algorithms are available. We describe a method suggested by Madow (1949) that combines systematic and random sampling for sampling without replacement with unequal probability. The probability of selecting a unit is proportional to an arbitrary weight associated with each unit. The method requires that the units be arranged in some linear order, calculating the total weight of the units and then drawing a systematic sample with a random start using a fixed length sampling interval along the cumulative weight totals. An implementation of this method for finite, linear network and areal populations using either independent random sampling or spatially-balanced sampling is available in the R package spsurvey by Kincaid and Olsen (2012) and is described by Stevens and Olsen (1999), Stevens and Olsen (2004) and Olsen, Kincaid and Payton (2012).

The implementation may be either class-based or proportional to an auxiliary variable. A brief description of the class-based approach is

- Create k=1, . . . , K classes based on information that is known for all sample units in the sampling frame. The classes are groups of sample units that have different unequal probability of selection.
- Define s_k as the set of sample units in class *k* where information in the sampling frame is used to assign the sample unit to class *k*. All units must be assigned to one and only one class.
- Define $\pi_{ki} = n_k / M_k$ and $u_{ki} = 1/\pi_{ki} = M_k / n_k$ as the inclusion probability and weight for the *i*th sample unit in the *k*th unequal probability class, respectively, and where n_k is the desired sample size and M_k is the extent for the *k*th unequal probability class. $M = \sum_{k=1}^{K} M_k$ is the total extent of the sampling frame. Extent is the number of sample units in the class for a finite sampling frame, the total length of all sample units in the class for a linear network sampling frame, and the total area of all sample units in the class for an areal sampling frame.
- Create a line whose length is equal to the sum of all the weights in the sampling frame and consisting of segments whose lengths are equal to the weight of the sample units. The sample units are either randomly ordered for an independent random sample or ordered by a spatially-balanced algorithm.
- Select a systematic sample of size n using a random start on the line. A sample unit is selected if a member of the systematic sample occurs on the sample unit's segment.

The selection process guarantees a fixed sample size but it does not guarantee that the weights will sum to the extent of the sampling frame. The latter is due to the number of sample units in each class being random, i.e., the actual number of sample units in class *k* may differ from the desired sample size n_k. The weights are based on n_k, consequently the sum of the weights will only be equal to the extent of the sampling frame when the actual sample size is the same as the desired sample size.

2.2.2.1 Example Lakes

Assume an unequal probability survey design is used to select 300 lakes where the unequal probability classes are a combination of ecoregion and lake area and that the expected sample size for each class is 75. In this case the weight for each class is calculated as $u_{ki} = 1/\pi_{ki} = M_k / 75$ and the results are given in Table 2.2.2.1. Each lake in the sample that is ≤ 10 ha in the mountain ecoregion represents 11.69333 lakes while lakes > 10 ha in the lowland ecoregion represents 2.13333 lakes. The weights are the same as those for the lake stratified design. The sum of the weights for each class is not equal to the class extent and the sum across all classes does not equal the sampling frame extent. For example, the mountain ecoregion with lakes ≤ 10 ha class weight sum equals 935.4664 (80*11.69333). The sum of all the weights is 1862.9333 which is close to the sampling frame extent of 1831. This is due to the unequal selection algorithm that selects a fixed sample size overall but does not guarantee the desired sample size in each class, i.e. 75 in this example.

Table 2.2.2-1 Number of lakes and unequal probability class weights by ecoregion and lake area

Ecoregion	Class Extent			Class Weight		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha	Total	≤ 10 ha	>10 ha	
Mountain	877	281	1158	11.6933	3.7467	
Lowland	513	160	673	6.8400	2.1333	
Total	1390	441	1831			
Ecoregion	Class Realized Sample Size			Sum of Design Weights		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha	Total	≤ 10 ha	>10 ha	Total
Mountain	80	71	151	935.467	266.012	1201.48
Lowland	73	76	149	499.320	162.133	661.45
Total	153	147	300	1434.79	428.14	1862.93

2.2.2.2 Example Streams

Assume an unequal probability survey design is used to select 300 sample units (i.e., sites) where the unequal probability classes are a combination of ecoregion (Mountain and Lowland) and stream size (Large and Small) and that the expected sample size for each class is 75. In this case the weight for each class is calculated as $u_{ki} = 1/\pi_{ki} = M_k/75$ and the results are given in Table 2.2.2-1. Each stream sample unit that is a small stream in the mountain ecoregion represents 723.68 km of stream length while large stream sample units in the lowland ecoregion represent 98.48 km of stream length. The weights are the same as those for the stream stratified design. The sum of the weights for each class is not equal to the class extent and the sum across all classes does not equal the sampling frame extent. For example, the mountain ecoregion, small stream class weight sum equals 50,658 km (70*723.6822). The sum of all the weights is 86,540 which is close to the sampling frame extent of 88,115. This is due to the unequal selection algorithm that selects a fixed sample size overall but the sample size in each class depends on the random selection process.

Ecoregion	Class Extent		Class Weight	
	Stream Size		Stream size	

	Small	Large	Total	Small	Large	
Mountain	54,276	17,033	71,309	723.6822	227.1034	
Lowland	9,420	7,386	16,806	125.5954	98.4817	
Total	63,696	24,419	88,115			
	Class realized sample size			Sum of design weights		
	Stream Size			Stream size		
Ecoregion	Small	Large	Total	Small	Large	Total
Mountain	70	86	156	50,658	19,531	70,189
Lowland	80	64	144	10,048	6,303	16,351
Total	150	150	300	60,706	25,834	86,540

2.2.2.3 Example Coastal Waters

2.2.3 Stratified Design

Stratified designs divide the sampling frame into H mutually exclusive and exhaustive subsets so that each sample unit is in one and only one of the H subsets. Within each of the strata sample units can be selected either with equal probability or with unequal probability. The selection in each stratum is completed independently of the selection in the other strata. First, assume equal probability of selection is used within the strata. Assume that n_h is selected from stratum h = 1, …, H where $n = \sum_{h=1}^{H} n_h$ is the total sample size and that the extent of each strata is M_h where $M = \sum_{h=1}^{H} M_h$ is the extent of the sampling frame. The design weights are $w_{hi} = M_h / n_h$ for the ith sample unit in the hth stratum. Note that the sum of the weights in each stratum is equal to the stratum extent and sum of weights across all strata is equal to the total frame extent.

2.2.3.1 Example Lakes

Define four strata for the Oregon lakes example based on the combination of whether the lake is (1) in a Mountain or Lowland ecoregion and (2) is less than or equal to 10 ha or greater than 10 ha (Table 2.2.3.1). Assume that the sample size is 75 in each stratum. Then the design weights are calculated as $w_{hi} = M_h / n_h$ and given in Table 2.2.2-1. Each lake in the sample that is $\leq$ 10 ha in the mountain ecoregion represents 11.69333 lakes while lakes $>$ 10 ha in the lowland ecoregion represents 2.13333 lakes. Note that the sum of the weights equals both the stratum extent and overall sampling frame extent.

Table 2.2.3-1 Number of lakes and stratum weights by ecoregion and lake area

Ecoregion	Stratum Extent			Stratum Weight	
	Lake Area			Lake Area	
	$\leq$ 10 ha	>10 ha	Total	$\leq$ 10 ha	>10 ha
Mountain	877	281	1158	11.69333	3.74667
Lowland	513	160	673	6.84000	2.13333
Total	1390	441	1831		

2.2.3.2 Example Streams

Define four strata for Oregon streams based on combination of a stream (a) being in a mountain or lowland ecoregion and (b) being a small or large stream. Assume that the desired sample size is 75 sample units in each stratum. Then the design weights are calculated as $w_{hi} = M_h / n_h$ where M_h is the stratum extent n_h = 75 for all strata (Table 2.2.3-1).

	Stratum extent			Stratum weight	
	Stream size			Stream size	
Ecoregion	Small	Large	Total	Small	Large
Mountain	54,276	17,033	71,309	723.6822	227.1034
Lowland	9,420	7,386	16,806	125.5954	98.4817
Total	63,696	24,419	88,115		

2.2.3.3 Example Coastal Waters

2.2.4 Design based on a Master Sample

Occasionally several monitoring organizations conduct monitoring of an aquatic resource, e.g. streams, where the geographic regions overlap but the target populations differ. It is not always possible to coordinate the planning and design of the monitoring programs but the organizations would like to be able to combine data from their studies after the fact. Where possible they would like to be able to select sample units from a common sample frame and survey design. One approach to address this issue is the concept of a master sample (King and Jessen 1945, Yates 1981, Larsen, Olsen and Stevens 2008). The actual design for a particular survey is constructed by the way that the master sample is implemented. When sample units are used in order to evaluate *n* sample units, then the design is an equal probability survey design with design weights equal to the sampling frame extent divided by *n*. When sample units are used in order within *k* classes to achieve a sample size of n_k, then the design is an unequal probability survey design with design weights equal to the sampling frame extent in class *k* divided by n_k.

A brief description of the class-based approach is

- Define s_{EV} as the set of sample units in *s* that have been evaluated for potential eligibility and sampling. The number of sample units in s_{EV} is $n_{EV} = n_{IN} + n_{ER} + n_{ENR} + n_{UNK}$
- Create k=1, … , K classes based on information that is known for all sample units in the sampling frame. The classes are formed to reflect how sample units are chosen to be evaluated for potential use.

- Define M_k as the extent in the sampling frame in the kth class and $M = \sum_{k=1}^{K} M_k$ is the total extent of the sampling frame.
- Define s_k as the set of sample units in class k where information in the sampling frame is used to assign the sample unit to class k. All units must be assigned to one and only one class.
- Define $s_{k,EV}$ as the set of sample units in class k that were evaluated for potential eligibility and sampling, $s_{k,EV} = s_k \cap s_{EV}$. Then the number of sample units in k that have been evaluated is $n_{k,EV}$ and $n_{EV} = \sum_{k=1}^{K} n_{k,EV}$
- Define the master sample weight for sample unit i as $w_{Di} = M/n_M$ where n_M is the master sample size.
- Calculate the implemented design weights for sample units in class k as $a_k = \dfrac{M_k}{\sum_{i \in s_{k,EV}} w_{Di}}$ where w_{Di} is the master sample design weight for the ith sample unit.
- The design weight for the implemented design for sample unit i in $s_{k,EV}$ is $w_{Vi} = w_{Di} a_k$.
- The adjusted weight w_{Vi} is zero for the sample units not needed in class k, i.e. in $s_{k,NN} = s_k \cap s_{NN}$.

2.2.4.1 Example Lakes

To select a master sample for a finite population, such as lakes, the sample size is selected to be equal to the extent of the finite population. For the lake example, the sampling frame contains 1831 lake objects so the sample size is set to 1831. A spatially-balanced sample selected using GRTS selects all the lake objects without replacement such that the lakes appear in reverse hierarchical order. Essentially a master sample is a very large over-sample from the sampling frame. The design weights are equal to 1.0 since each lake represents only itself in the master sample.

Assume that a spatially-balanced master sample for lakes is selected and that a survey design is implemented based on k = 4 classes: Mountains ≤ 10 ha, Lowlands ≤ 10 ha, Mountains > 10 ha, Lowlands > 10 ha. Sufficient sample units are evaluated in each class until 40 lakes are found that can be sampled. The calculation of design weights is completed using the procedures for unequal probability survey designs where the initial design weights are equal to 1.0. The R function *wgtadjOS* can be used to complete the weight calculation.

Table 2.2.4-1 Design weights based on a master sample of lakes where 40 sampled lakes are achieved for each combination of ecoregion and lake area

Ecoregion	Class Extent			Class Realized Sample Size		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha	Total	≤ 10 ha	>10 ha	

Mountain	877	281	1158	69	47	116
Lowland	513	160	673	64	50	114
Total	1390	441	1831	133	97	230
Ecoregion	Class Weight			Sum of Design Weights		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha		≤ 10 ha	>10 ha	Total
Mountain	12.7101	5.9787		877.0	281.0	1158.0
Lowland	8.0156	3.2000		513.0	160.0	673.0
Total				1390.0	441.0	1831.0

2.2.4.2 *Example Streams*

To select a master sample for a linear network, such as streams and rivers, the sample size is selected based on the density of sample units desired. For example, a density of one sample unit per km of stream length is likely to provide a sufficient number of sample units for most anticipated survey designs. Essentially a master sample is a very large over-sample from the sampling frame. Assume that a master sample of 10,000 sample units is selected from the Oregon stream sample frame. In this case, the design weights are equal to 8.8115 km since the extent of the sample frame is 88,115 km and the sample density is 1 sample unit per 8.8115 km.

From this spatially-balanced master sample assume that a survey design is implemented based on k = 4 classes: Mountain small streams, mountain large streams, lowland small streams and lowland large streams. Moreover assume that a sufficient number of sample units are evaluated in each class until 40 streams are found that can be sampled. The calculation of design weights is completed using the procedures for unequal probability survey designs where the initial design weights are equal to 8.8115 km. The R function *wgtadjOS* can be used to complete the weight calculation.

Table 2.2.4-2 Design weights based on a master sample for streams where 40 stream sites are sampled for each combination of ecoregion and stream size

Ecoregion	Class Extent			Class Realized Sample Size		
	Stream size			Stream size		
	Small	Large	Total	Small	Large	
Mountain	54,276	17,033	71,309			
Lowland	9,420	7,386	16,806			
Total	63,696	24,419	88,115			
Ecoregion	Class Weight			Sum of Design Weights		
	Stream size			Stream size		
	Small	Large		Small	Large	Total
Mountain						

Lowland						
Total						

2.2.4.3 Example Coastal Waters

2.3 As-Implemented Design Weight Adjustment

Aquatic resource monitoring surveys commonly have units in a sample that are determined not to be eligible sample units or if they are eligible sample units cannot be sampled for various reasons. The cost associated with field sampling is sufficiently high that a program's budget cannot afford to sample more than planned. Moreover, sampling less than the planned number many times is not desirable as funds allocated to fieldwork may be lost or the actual sample size may be much smaller than planned. Experience from prior studies provides some information about expected eligibility rates and nonresponse rates. That information can be used to determine a sample size necessary to come close to the desired number of responding sample units. While this may be preferred from a statistical survey design perspective, it does not guarantee the desired number of responding sample units. An alternative, termed an over-sample, is to select a sample that is known to be sufficiently large and use the sample units in that sample in such a way to preserve the integrity of the survey design, albeit with a random sample size. For example, from a simple random sample of size 300, the first 204 sample units of that sample are a simple random sample so that if only those sample units are evaluated for eligibility and sampled if possible to achieve exactly 150 sampled units, they are a simple random sample of size 204 from the sampling frame. For a spatially balanced sample produced by the GTRS algorithm with reverse hierarchical ordering, the first 204 sample units of a GRTS equal probability design of size 300 are also a GRTS equal probability design of size 204. Both statements are also true for unequal probability designs. For stratified designs, similar statements are true for each stratum. Since the original design weights are based on an assumed sample size of n (300 in the example) but the actual sample size is less than n (204 in the example), the design weights must be adjusted.

A variation of an over-sample is to select the sample as stated but instead of calculating the initial weights for the entire sample, calculate the weights based on the desired number of sample units to be sampled (e.g. 150 in the sample). For this alternative, the initial design weights are based on the desired sample size and the over-sample initial design weights are as well. If 204 sample units must be evaluated to determine their eligibility and ability to be sampled so that 150 are actually sampled, then the original design weights must be adjusted to account for the use of 204 units instead of the planned 150 units.

Adjusting the original design weights for an as-implemented design is accomplished using a class-based approach. A brief description of the class-based approach is

- Define s_{EV} as the set of sample units in s that have been evaluated for potential eligibility and sampling. The number of sample units in s_{EV} is $n_{EV} = n_{IN} + n_{ER} + n_{ENR} + n_{UNK}$
- Create k=1, … , K classes based on information that is known for all sample units in the sampling frame. The classes are formed to reflect how sample units are chosen to be evaluated for potential use.
- Define M_k as the extent in the sampling frame in the kth class and $M = \sum_{k=1}^{K} M_k$ is the total extent of the sampling frame.
- Define s_k as the set of sample units in class k where information in the sampling frame is used to assign the sample unit to class k. All units must be assigned to one and only one class.
- Define $s_{k,EV}$ as the set of sample units in class k that were evaluated for potential eligibility and sampling, $s_{k,EV} = s_k \cap s_{EV}$. Then the number of sample units in k that have been evaluated is $n_{k,EV}$ and $n_{EV} = \sum_{k=1}^{K} n_{k,EV}$
- Calculate the adjustment for sample units in class k as $a_k = \dfrac{M_k}{\sum_{i \in s_{k,EV}} w_{Di}}$ where w_{Di} is the design weight for the ith sample unit.
- The adjusted weight for sample unit i in $s_{k,EV}$ is $w_{Vi} = w_{Di} a_k$.
- The adjusted weight w_{Vi} is zero for the sample units not needed in class k, i.e. in $s_{k,NN} = s_k \cap s_{NN}$.

2.3.1 Example Lakes

Assume that a monitoring program requires 150 lakes to be sampled. To ensure that exactly 150 lakes can be sampled, the program elects to use an unequal probability survey design to select 300 lakes where the unequal probability classes are a combination of ecoregion and lake area and that the expected sample size for each class is 75. Past experience suggests that a sample size of 300 lakes will provide a sufficient number of lakes to result in 150 sampled lakes. Table 2.2.3.1 summarizes the initial design and design weights. Assume that to sample 150 lakes the lakes are evaluated in order from the total sample of 300 lakes and that n_{EV} = 213 evaluations were required. To adjust the design weights to account for the as-implemented design, a single class is used with M = 1831 sample units in the class. The adjustment factor is $a_1 = \dfrac{M}{\sum_{i \in s_{1,EV}} w_{Di}} = \dfrac{1831}{1321.787} = 1.385246$ and the as-implemented adjusted weight is then $w_{Vi} = w_{Di} * 1.385246$ where the unique weights are 16.1981 for Mountain lakes ≤ 10 ha, 9.4751for Lowland lakes ≤ 10 ha, 5.19001 for Mountain lakes >10 ha and 2.955192 for Lowland lakes >10 ha.

Suppose instead that lakes are replaced within the classes created by the combination of Mountain and Lowland lakes and of lakes ≤ 10 ha and lakes >10 ha. That is, if a lake cannot be sampled it is replaced by a lake within the same class until 40 lakes are sampled within that class. Moreover once 40 lakes in a class are sampled no additional lakes in that class are evaluated. The same initial design is used as before. The adjustment factors using the four classes are $a_1 = \frac{M_1}{\sum_{i \in s_{1,EV}} w_{Di}} = \frac{877}{689.9067} = 1.2712$ for Mountain lakes ≤ 10 ha, $a_2 = \frac{M_2}{\sum_{i \in s_{2,EV}} w_{Di}} = \frac{513}{437.7600} = 1.1719$ for Lowland lakes ≤ 10 ha, $a_3 = \frac{M_3}{\sum_{i \in s_{2,EV}} w_{Di}} = \frac{281}{217.3067} = 1.2931$ for Mountain lakes > 10 ha, and

$a_4 = \frac{M_4}{\sum_{i \in s_{2,EV}} w_{Di}} = \frac{160}{108.8000} = 1.4706$ for Lowland lakes > 10 ha. The adjusted weights and sum of the adjusted weights are given in Table 2.3.1.1. The R function *wgtadjOS* can be used to complete the weight calculation.

Table 2.3.1-1 As-implemented weight adjustment summary when sample units are replaced by ecoregion and lake area

Ecoregion	Class Extent			Number of Sample Units Evaluated		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha	Total	≤ 10 ha	>10 ha	Total
Mountain	877	281	1158	59	58	117
Lowland	513	160	673	64	51	115
Total	1390	441	1831	123	109	232
Ecoregion	Design Weights			Sum of Design Weights		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha		≤ 10 ha	>10 ha	Total
Mountain	11.6933	3.7467		982.2	264.1	1246.3
Lowland	6.8400	2.1333		441.2	172.8	614.0
				1423.4	436.9	1860.3
Ecoregion	Adjusted Weights			Sum of Adjusted Weights		
	Lake Area			Lake Area		
	≤ 10 ha	>10 ha		≤ 10 ha	>10 ha	Total
Mountain	14.8644	4.8448		877	281	1158
Lowland	8.0156	3.1273		513	160	673
Total				1390	441	1831

The two examples for as-implemented design weight adjustment illustrate the importance of ensuring that the classes used in the weight adjustment are the same classes that are used when selecting lakes to replace lakes that cannot be sampled. The adjusted weights differ because different classes are used. The weight adjustment also depends on a random process since whether a lake must be replaced depends on it being selected for evaluation and on whether it can be sampled. Consequently, the adjusted weights are not fixed but are random in the sense that they are influenced by random events. To minimize the impact that the random events have on the adjusted weights, the classes should be selected so that the number of sample units evaluated is at least greater than 10 and preferably much larger. When 10 sample units are evaluated, the adjusted weights will change by less than 10 percent if one more or one less sample unit is evaluated.

2.3.2 Example Streams

Table 2.3.2-1 As-implemented design weight adjustment summary when sample units are replaced by ecoregion and stream size

	Class Extent			Number of Sample Units Evaluated		
	Stream size			Stream size		
Ecoregion	Small	Large	Total	Small	Large	Total
Mountain	54,276	17,033	71,309			
Lowland	9,420	7,386	16,806			
Total	63,696	24,419	88,115			
	Design Weights			Sum of Design Weights		
	Stream size			Stream size		
Ecoregion	Small	Large		Small	Large	Total
Mountain						
Lowland						
Total						
	Adjusted Weights			Sum of Adjusted Weights		
	Stream size			Stream size		
Ecoregion	Small	Large		Small	Large	Total
Mountain						
Lowland						
Total						

2.3.3 Example Coastal Waters

2.4 Unknown Eligibility Weight Adjustment

Sampling frames for aquatic surveys typically include sample units that are ineligible with respect to the definition of the target population. It might not possible to determine a sample unit's eligibility, i.e., its eligibility is unknown. For example, field visits might be required to determine eligibility, and access to sites might be denied by landowners, or by physical inaccessibility. When sample units' eligibility can be determined, e.g., when field visits are possible, it is typical that some units are eligible and some are ineligible. This is evidence that some of the sample units whose eligibility is unknown might be ineligible. An adjustment of the weights to account for this unknown eligibility may be desirable. Three courses of action are possible to account for the unknown eligibility.

2.4.1 No Adjustment

One option is not to adjust the design weights and define three classes for the sample units: known eligible ("Target"), known ineligible ("Non-target") and unknown eligibility ("Unknown"). In this case any estimates based on sampled units do not reflect the characteristics of the entire population of lakes but to a subset of the population whose eligibility could have been determined and sampled. If no adjustment is made for unsampleable units, then this subset is called the "sampled population" by many authors (Lohr 1999). In this case, the design weights, w_{Di}, are used for the analyses. For the equal probability lake example, the weights are w_{Di} = 6.10333 for all $i \in s$.

2.4.2 Deterministic Decision Adjustment

Another option is to make a decision about the unknown eligibility of the sample units. For example, the evaluation process for determining the eligibility status of a sample unit may be sufficient to conclude that the sample units are almost certainly eligible. That is, the sample units are deemed eligible but not sampleable and any weight adjustment would be done when adjustments are made for nonresponse. In this case, the sample units in s_{UNK} are assigned to s_{ENR} and the design weights, w_{Di}, do not change. For the equal probability lake example, the weights do not change and are w_{Di} = 6.10333 for all $i \in s$.

2.4.3 Class-based Adjustment

Another method of handling the unknown eligibility sample units is to distribute their total sample weight among those whose eligibility status is known using a class-based approach. A class-based approach is used since little information may be available about the sample units with unknown eligibility. In general, the class-based approach is

- Create b=1, . . . , B classes based on information that is known for all sample units with known eligibility.
- Define s_b as the set of sample units in class b where information in the sampling frame is used in assigning the sample unit to class b. This is done regardless of the sample unit's eligibility or sampleability.
- Define $s_{b,KN} = s_b \cap s_{KN}$ as the set of sample units in s_b with known eligibility.
- Calculate the eligibility adjustment for sample units in class b as
  $a_{Ub} = \frac{\sum_{i \in s_b} w_{Di}}{\sum_{i \in s_{b,KN}} w_{Di}}$ where w_{Di} is the design weight for the ith sample unit.

- The adjusted weight for unknown eligibility for sample unit *i* in $s_{b,KN}$ is $w_{Ubi} = w_{Di} a_{1b}$.
- Set the weights for the remaining sample units in class b, i.e., those with unknown eligibility, to zero.

Note that the factor $1/a_{Ub}$ is an estimate of the probability of a sample unit having a known eligibility status within class b. The R function *wgtadjU* can be used to complete the weight calculation.

2.4.3.1 Example Lakes

Suppose that the 300 sample units (potential lakes) from the equal probability survey design are evaluated to determine if they are eligible for the study, i.e., meet the definition of a lake, and if they can be sampled. The initial evaluation may be completed in the office using available information. When the available information is not sufficient, then a field reconnaissance visit may be made to gather additional information about whether the lake is eligible. Finally, for sample units where the determination is that the sample unit is thought to be eligible, the sample unit is scheduled for sampling first by determining if it is accessible. A sample unit may be inaccessible for two major reasons. First, permission to access the sample unit may be required and the landowner may or may not give permission to access and sample it. Second, the sample unit may be physically inaccessible or unsafe to sample. For example, the sample unit may be in a remote wilderness area where a decision is made that the cost to access it is determined to be too great. Based on the evaluations completed, information on whether the sample unit selected is an eligible sample unit is compiled and for the eligible sample units (lakes) whether it was possible to sample them (Table 2.4.3-1).

Based on prior experience it is known that lake sample units that are ≤ 10 ha are more likely to be ineligible than lake sample units that are >10 ha. Form two classes, sample units that are less than or equal to 10 ha and sample lakes that are greater than 10 ha. Then the adjustment factors are

$$a_{U1} = \frac{\sum_{i \in s_1} w_{Di}}{\sum_{i \in s_{1,KN}} w_{Di}} = \frac{233 * 6.10333}{216 * 6.10333} = 1.0787$$

$$a_{U2} = \frac{\sum_{i \in s_2} w_{Di}}{\sum_{i \in s_{2,KN}} w_{Di}} = \frac{67 * 6.10333}{65 * 6.10333} = 1.0308$$

for sample units ≤ 10 ha (b=1) and for sample units >10 ha (b=2). The weights adjusted for unknown eligibility are

$$w_{U1i} = w_{Di} a_{U1} = 6.10333 * 1.0787 = 6.5837$$
$$w_{U2i} = w_{Di} a_{U2} = 6.10333 * 1.0308 = 6.2911$$
$$w_{U,UNK} = 0 \; .$$

To check the weight adjustment, the sum of the weights is 216*6.5837 + 65*6.2911 + 19*0 = 1831.

2.4.3-1 Unknown eligibility weight adjustment based on two lake area classes for an equal probability design

Evaluation category	Number of sample units evaluated		
	Lake area class		
	≤ 10 ha	>10 ha	Total
UNK-unknown eligibility	17	2	19
ER-eligible sampled	131	46	177
ENR – eligible not sampled	58	12	70
Access denied	29	8	37
Physically inaccessible	29	4	33
IN-ineligible	27	7	34
KN-known eligibility	216	65	281
UNK-unknown eligibility	17	2	19
Total	233	67	300
	Weight adjustment		
	Lake area class		
	≤ 10 ha	>10 ha	
Design weights	6.1033	6.1033	
Adjustment factor	1.0787	1.0308	
Weights adjusted for unknown eligibility	6.5837	6.2911	

2.4.3.2 Example Streams

Table 2.4.3-2 Unknown eligibility weight adjustment based on two stream size classes for an equal probability design

Evaluation category	Number of sample units evaluated		
	Stream size class		
	Small	Large	Total
UNK-unknown eligibility			
ER-eligible sampled			
ENR – eligible not sampled			
Access denied			
Physically inaccessible			
IN-ineligible			
KN-known eligibility			
UNK-unknown eligibility			
Total			
	Weight adjustment		
	Stream size class		
	Small	Large	

Design weights			
Adjustment factor			
Weights adjusted for unknown eligibility			

2.4.3.3 Example Coastal Waters

2.5 Nonresponse Weight Adjustment

In most aquatic surveys some of the sample units that are eligible for sampling cannot be sampled leading to nonresponse for those units. Some sample units may be physically inaccessible or unsafe to access, e.g. a stream may be located at the bottom of a steep canyon which cannot be safely accessed, a estuarine site may be located in the middle of a busy shipping channel or in an underwater ordinance disposal area. Whether a sample unit is physically inaccessible is also a function of the cost to access the unit. For example, a lake or stream sample unit may be located in a remote region that requires use of costly equipment or trips requiring multiple days. Access to sample units in many aquatic surveys requires obtaining permission of a landowner. A landowner may deny access prior to a field crew visit or even if prior approval was given may deny access when a field crew arrives at the site. The reason for nonresponse may not be independent of the response, i.e., missing completely at random (MCAR). In cases, the nonresponse may be missing at random (MAR). That is, the sample unit's nonresponse does not depend on the value of the response but depends only on information that is known about the sample unit and all other sample units in the sample that are eligible to be sampled regardless of whether they are sampled or not sampled. If the probability that a sample unit can be sampled, i.e. responds, depends on one or more of the variables that are being measured and this dependence cannot be removed based on information that is known for both sampleable and nonsampleable sample units, then nonignorable nonresponse (NINR) is present. In this situation, it is not possible to adjust the weights.

Two courses of action can be taken to address the presence of nonresponse. One action is not to adjust the weights and restrict the inference to the sampled population, as in the same option for addressing unknown eligibility. This is the only action that can be taken for NINR. It must be made clear to those who use the results that the results do not apply to the entire population of interest but only to the sampled population. If all the variables measured and reported for the survey are categorical, then another option is to explicitly create an additional category for each variable that identifies the result as not available, e.g., a category "No Data". This is not an option when reporting means or other summary statistics for continuous variables.

A second option is to adjust the weights by a class-based approach. The process is the same as the class-based approach described when adjusting weights for unknown eligibility. The class-based approach is

- Create c=1, . . . , C classes based on information that is known for all sample units in the sample.

- Define s_c as the set of sample units in class c where information in the sampling frame is used in assigning the sample unit to class c. This is done regardless of the sample unit's eligibility or sampleability.
- Define $s_{c,E} = s_c \cap \left(s_{ER} \cup s_{ENR}\right)$ as the set of sample units in class c that are eligible
- Define $s_{c,ER} = s_c \cap s_{ER}$ as the set of sample units in s_c that are sampleable, i.e., have a response.
- Calculate the nonresponse adjustment for sample units in class c as $a_{Rc} = \dfrac{\sum_{i \in s_{c,E}} w_{Ui}}{\sum_{i \in s_{c,ER}} w_{Ui}}$ where w_{Ui} is the weight for the *i*th sample unit, possibly adjusted for unknown eligibility and as-implemented design.
- The adjusted weight for nonresponse for sample unit *i* in $s_{c,ER}$ is $w_{Rci} = a_{Rc} w_{Ui}$. That is, the adjustment is only made to sample units that were sampleable.
- Set the weights for the remaining sample units in class c, i.e., those that are not sampleable, to zero.

Note that the factor $1/a_{Rc}$ is an estimate of the probability of a sample unit being sampleable within class c. Also note that if no weight adjustment for unknown eligibility is made then w_{Ui} is replaced by w_{Di} , the design weight. The R function *wgtadjNR* can be used to complete the weight calculation.

The weight after adjustment for sample unit i is

$$
\begin{aligned}
w_{Ri} &= \begin{cases} w_{Ui} a_{2c} & \text{when } i \in s_{c,ER} \\ w_{Ui} & \text{when } i \in s_{IN} \\ 0 & \text{when } i \in s_{NK} \cup s_{ENR} \end{cases} \\
&= \begin{cases} w_{Di} a_{Ub} a_{Rc} & \text{when } i \in s_{b,KN} \cap s_{c,ER} \\ w_{Di} a_{Rb} & \text{when } i \in s_{b,KN} \cap s_{IN} \\ 0 & \text{when } i \in s_{UNK} \cup s_{ENR}. \end{cases}
\end{aligned}
$$

The sample units that have known eligibility and are sampleable receive adjustments for both unknown eligibility and nonresponse. Sample units that have known ineligibility are only adjusted for unknown ineligibility. Sample units that have unknown eligibility and are not sampleable are assigned zero weights, i.e., they are not included in the estimation process. The classes used in the class-based approach do not need to be the same for the unknown eligibility adjustment and the nonsampleable (nonresponse) adjustment.

The class-based approach requires making a decision on what classes should be used. The decision involves knowing, or making assumptions, about whether the nonresponse is missing completely at random (MCAR) or missing at random (MAR). Having sample units be MCAR is not common in aquatic surveys. In the event that they are MCAR, then a single class can be used for the weight adjustment. MAR is more common occurrence in aquatic surveys and leads to weight adjustment using classes based on factors known, or expected, to be related to the response. For example, sample units that are difficult to access are more likely not to have local human disturbance which may impact the response.

2.5.1.1 Example Lakes

Continuing the lake example (Example 2.4.3.1) with the equal probability survey design, assume that 247 lakes were determined to be eligible for the study (Table 2.5.1-1). Of these 177 were sampled, 33 were physically inaccessible and 37 were denied access by the landowner. Based on prior experience we know that the response for lakes that are in mountain ecoregions is different from lakes in lowland ecoregions. For lakes where access is denied by landowners, we assume that the response does not differ from lakes were access is granted. We also know that physically inaccessible lakes likely differ in response from lakes that are physically accessible. While we can define a class based on inaccessibility, no lakes that were sampled are in that class. Consequently, forming a class based on inaccessibility cannot be used for nonresponse adjustment. One option is to assume that mountain lakes that are inaccessible are similar to mountain lakes that are accessible and sampled; similarly for lowland lakes. The MAR assumption may not be completely satisfied for this option. Another option is to find another characteristic related to inaccessibility, e.g. classes based on distance to road. For our example, we will use two classes lowland, s_1, and mountain, s_2, to adjust the weights for nonresponse. The number of lakes in s_1 is 148 and in s_2 is 99. The number of eligible lakes with a response is 119 and 61 for lowlands ($s_{1,ER}$) and mountains ($s_{2,ER}$), respectively.

Table 2.4.3-1 Number of eligible sample units and eligible sampleable units by nonresponse class

NonResponse Class		E: All Eligible Sample Units			
	Lake area	Sampled	Physically inaccessible	Denied access	Total
Lowland		**61**	**9**	**23**	**93**
	≤ 10 ha	45	7	17	69
	>10 ha	16	2	6	24
Mountain		**116**	**24**	**14**	**154**
	≤ 10 ha	86	22	12	120
	>10 ha	30	2	2	34
Total		**177**	**33**	**37**	**247**

For the example, 70 eligible lakes cannot be sampled (Table 3). Based on prior experience, we believe that the response for lakes that are in mountain ecoregions is different than for lakes in lowland ecoregions. Form two classes based on this experience. Then, using data from Table 3, the adjustment factors are

$$a_{R1} = \frac{\sum_{i \in s_{1,E}} w_{Ui}}{\sum_{i \in s_{1,ER}} w_{Ui}} = \frac{69*6.583688 + 24*6.291128}{45*6.583688 + 16*6.291128} = \frac{605.2616}{396.924} = 1.524880$$

$$a_{R2} = \frac{\sum_{i \in s_{2,E}} w_{Ui}}{\sum_{i \in s_{2,ER}} w_{Ui}} = \frac{120*6.583688 + 34*6.291128}{86*6.583688 + 30*6.291128} = \frac{1003.9410}{754.931} = 1.329845$$

for sample units in lowland ecoregions (c=1) and for sample units in mountains (c=2), respectively. The weights, w_{Rbci}, where *b* is class for Unknown adjustment, *c* is class for nonresponse adjustment and *i* is the *i*th lake with the joint classes are

$$w_{R11i} = w_{Ui} a_{R1} = w_{Di} a_{U1} a_{R1} = 6.583688*1.524880 = 6.10333*1.0787*1.524880 = 10.039$$
$$w_{R21i} = w_{Ui} a_{R1} = w_{Di} a_{U2} a_{R1} = 6.291128*1.524880 = 6.10333*1.0308*1.524880 = 9.5932$$
$$w_{R12i} = w_{Ui} a_{R2} = w_{Di} a_{U1} a_{R2} = 6.583688*1.329845 = 6.10333*1.0787*1.329845 = 8.7553$$
$$w_{R22i} = w_{Ui} a_{R2} = w_{Di} a_{U2} a_{R2} = 6.291128*1.329845 = 6.10333*1.0308*1.329845 = 8.3662$$
$$w_{RU1i} = w_{Di} a_{U1} = 6.10333*1.0787 = 6.5837$$
$$w_{RU2i} = w_{Di} a_{U2} = 6.10333*1.0308 = 6.2911$$
$$w_{RUUNK} = 0$$
$$w_{RUENR} = 0$$

To check the weight calculations we sum the weights over all 300 sample units:

$$45*10.0393 + 16*9.5932 + 86*8.7553 + 30*8.3662 +$$
$$27*6.5837 + 7*6.2911 + 19*0 + 70*0 = 1831 .$$

2.5.1.2 Example Streams

2.5.1.3 Example Coastal Waters

3 Auxiliary Data in Weighting

Post Stratification use to adjust for known under coverage of sampling frame

4 Combining Independent Survey Designs

Suppose that a state or an EPA Region knows that several probability surveys for a particular type of aquatic resource have been conducted in an area. If the surveys also measure the same indicators and were conducted in a common time period, then it is desirable to combine the probability surveys for the population estimation. Questions arise concerning when is this possible, what assumptions must be made, and operationally how is it done.

The first step is to determine that the studies of interest are in fact probability surveys. This includes knowing the characteristics of the survey design and, in particular, knowing the inclusion probabilities or weights for all sites in the studies. The survey designs must also be statistically independent. That is, the sample selection processes would have been completed independently.

Each study has set of indicators that were measured. If no indicators are common across the studies, then no reason exists to combine the studies. If the same indicators are included, then the measurement protocols and calculation procedures must be reviewed to ensure that the indicators are really the same. For example, a benthic macroinvertebrate Index of Biotic Integrity (benthic IBI) may be included in all the studies but different field protocols, laboratory counting protocols, or IBI calculation procedures may be used. Combining the indicators may not make ecological sense in this case. Assume for the remaining discussion that it does make sense to combine the indicators.

Each study also defined a specific target population. Typically, the target population not only involves a definition of the aquatic resource but also a geographic area. In some cases, the same definitions are used and the only difference among the studies is that they were independently conducted in different years. For example, two separate GRTS survey designs could have been used to select one sample to be measured in 2001 and in 2002. In other cases, the same definition of the aquatic resource is used and the studies were conducted in either entirely separate geographic regions or with one or more regions overlapping. Another possibility is that the studies use the same geographic area but use different definitions for the aquatic resource. For example, one study could focus on wadeable streams and the other on non-wadeable streams and rivers. Of course, the definitions may also overlap. A review of the studies must carefully identify all areas of overlap in geography and aquatic resource. Each site must be classified according to its membership in these groups. Sometimes a graphical presentation, such as a Venn diagram is useful to identify the groups.

5 References

Cochran, W. G. (1987). *Sampling Techniques*. 3rd edition. New York: John Wiley & Sons.

Gitzen, R. A., Millspaugh, J. J., Cooper, A. B. and Licht, D. S., editors. (2012). Design and Analysis of Long-term Ecological Monitoring Studies. New York: Cambridge University Press.

Kincaid, T. M. and Olsen, A. R. (2012). spsurvey: Spatial Survey Design and Analysis. R package version 2.5. http://www.epa.gov/nheerl/arm/

King, A. J. and Jessen, R. J. (1945). The master sample of agriculture. *Journal of the American Statistical Association* **40**, 38-56.

Larsen, D. P., Olsen, A. R. and Stevens, D. L., Jr. (2008). Using a Master Sample to integrate stream monitoring programs. *Journal of Agricultural, Biological, and Environmental Statistics* **13**, 243-254.

Little, R. J. A. and Rubin, D. B. (2002). *Statistical Analysis with Missing Data*. 2nd edition. New Jersey: John Wiley & Sons.

Lohr, S. L. (1999). *Sampling: Design and Analysis*. Pacific Grove, CA: Duxbury Press.

Madow, W. G. (1949). On the theory of systematic sampling, II. *Annals of Mathmatical Statistics* **20**, 333-354.

NHDPlus (2013). National Hydrography Dataset Plus. http://www.horizon-systems.com/NHDPlus/index.php

Olsen, A. R., Kincaid, T. M. and Payton, Q. (2012). Spatially balanced survey designs for natural resources. Pages 126-150 *in* R. A. Gitzen, J. J. Millspaugh, A. B. Cooper and D. S. Licht, editors. Design and Analysis of Long-Term Ecological Monitoring Studies. Cambridge, UK: Cambridge University Press.

Omernik, J. M. (1987). Ecoregions of the conterminous United States. *Annals of Association of American Geographers* **77**, 118-125.

Stevens, D. L., Jr. and Olsen, A. R. (1999). Spatially restricted surveys over time for aquatic resources. *Journal of Agricultural, Biological, and Environmental Statistics* **4**, 415-428.

Stevens, D. L., Jr. and Olsen, A. R. (2004). Spatially-balanced sampling of natural resources. *Journal of American Statistical Association* **99**, 262-278.

Thompson, S. (1992). *Sampling*. New York: John Wiley & Sons.

Valliant, R., Dever, J. A. and Kreuter, F. (2013). *Practical Tools for Designing and Weighting Survey Samples*. New York: Springer.

Yates, F. (1981). *Sampling Methods for Censuses and Surveys*. 4th edition. London: Charles Griffen & Co.

6 Glossary

Target populations:

The target population refers to the resource to be described. For example, a stream network in a particular watershed, all streams and rivers in a state, or the streams in a national forest. Critical in developing the design is an explicit definition of the target population. For this guidance document, we will consider a stream network as the target population and our goal is to describe attributes of that stream network, such as the number of a particular species of fish it contains, or their spatial distribution within that network, or the variation in physical habitat structure across the network (e.g., the distribution of % fines or habitat complexity across the network). The definition of the target population should contain specific information about the stream network: its spatial extent, its flow status (the perennial network? Includes the intermittent channels?); its size (all stream sizes? Just first order streams?). Should it contain only the fish bearing portion of the network, or that portion occupied by a particular species or population? The definition should be specific enough to determine unequivocally whether a location on a stream network is part of the target population. It is important to be very explicit about the target stream network that is the subject of study because the survey design will be developed to describe this target population.

Statisticians distinguish two types of target populations: discrete and continuous. Discrete populations consist of populations whose 'parts' can be identified and listed, such as the population of lakes or wetlands in a region. Alternatively, stream or road networks are often considered continuous. Continuous populations can be converted into discrete populations by the application of specific rules that break the resource into discrete elements. For example, stream networks could be converted into discrete form by identifying unique reaches defined at the network confluences. In general, we treat stream networks as continuous populations.

Elements of a population Elements of a population refer to the 'parts' that make up the target population. Elements of a discrete population are easy to describe in that they are the individuals that make up the population. Each lake or wetland in a population of lakes or population of wetlands is a population element. For continuous resources, population elements are points on the target resource, e.g., points on a stream network. Clearly, there can be an infinite number of points associated with a continuous population (more on this later). An important rule in the characterization of the population elements is an explicit definition so that members of a field crew can determine whether the site visited is a member of the target population. As indicated above, the stream network could be divided into discrete reaches or habitat units (using a consistent reach definition). In this case, each reach or habitat unit is a population element, and the collection of all the reaches or habitat units make up the target population.

Sample frame:

The frame is the representation of the target resource used in the selection of the sample. For discrete populations, the frame is often a list containing each population element, e.g., a list of lakes in the region of interest, sometimes referred to as a "list frame". For continuous resources, such as stream networks, a digital map of the stream network is the usual form of the frame. Accurate representations of stream networks therefore become critical as they become the functional target population.

Two types of frame errors occur: 1) mapped parts of the stream network that are not part of the target population, or 2) parts of the target population not represented on the maps. The first case is easier to handle than the second in that the set of sites selected as a sample can be evaluated with respect to target status prior to or at the time of the field visit, and then adjustments can be made to the final estimates/inferences by accounting for the fraction of the frame that was "non-target". Dealing with the second case is more difficult because it entails gathering information "outside" the perceived frame to evaluate how much of the actual stream network that should have been part of the frame was missed. An explicit example of this second case arose as part of the GRTS based stream survey developed for the Oregon Department of Fish and Wildlife's coastal coho monitoring program. The 1:100 K USGS digital hydrography was initially used as the frame. The ODFW knew that this was a somewhat inaccurate representation of the coho domain, but didn't know what fraction of the resource was missed. During the first 8 years of the survey, ODFW field crews gathered information about parts of the network missed by recording information on salmon-bearing streams observed in the field that were not in the frame. Approximately 10 % of the coho domain was excluded from the 1:100 K frame. During 2007, the survey design was modified for a variety of purposes; during this process, the frame was modified, partly to include streams not included in the original frame. This example illustrates the critical need to incorporate continual evaluation of the frame as part of an ongoing monitoring program.

Sample selection rules:

Sample selection rules describe the mechanics by which a sample will be selected from the frame to represent the population. Three general types of selection rules are: simple random sampling; systematic sampling; and GRTS. Each of these can include stratification, as well as a number of other refinements (e.g., nested sampling, adaptive sampling, cluster sampling; consult a survey design text such as Thompson (…) for the various ways that the simple designs can be tailored). See the variety of statistics texts for details of simple random sampling and systematic sampling and their permutations. This guidance manual sticks with GRTS as described elsewhere. With respect to the use of the master sample, selection rules indicate which part of a master sample is selected to meet the particular design requirements.

Weights, inclusion probabilities and inclusion densities:

Random sampling (simple, systematic, or GRTS) allows each element of the target population (as represented by the frame) a known chance of being included in the sample. This likelihood of being included is the inclusion probability (or inclusion density for continuous populations); its reciprocal, the sample weight, identifies how much of the population is represented by the sample point.. As a brief example, take a stream network 1000 km in length. Select 20 sites from the frame using an equi-probable GRTS selection, so that every point on the stream has the same chance of being selected. The inclusion density describes the number of sample points per unit length of the stream. In this case, the inclusion density is constant and given by 20/1000, or 0.02. Each point then represents (1000/20) = 50 km of the stream network.

Of course, inclusion probabilities and densities do not have to be constant. There are many reasons why a variable probability design might be used. For example, if there is some sub-population of particular interest, we may want to ensure a certain number of samples in the sub-population. We can do this by stratifying, or by specifying a variable inclusion probability. For example, it is often sensible to use some measure of stream size, e.g., Strahler order, to determine inclusion probabilities. An equiprobable sample will result in having most sample sites on small streams, because the preponderance of stream length is in small streams. If we want to increase the relative number of sample points on larger streams, then we need to increase the inclusion probability for larger streams. We could stratify, for example, by specifying that we want an equal number of sample sites on first, second, and third or higher order streams. Alternatively, we could say that we want the inclusion probability for second order to be twice the inclusion probability of first order, and the inclusion probability of third or higher to be four times that of first order.

The above example using stream order to define inclusion probabilities is an example of using an ancillary variable to structure the sample. The ancillary variable need not be discrete, as stream order is. For example, a continuous variable such as elevation or annual precipitation could be used. One could also develop a function that combined several ancillary variables. The only essential requirement is that the ancillary must be known for every population element (because we must be able to determine the inclusion probability or density for every element in the population). Determining inclusion probabilities and weights can become complex depending the complexity of the design.

Stratification and variable probability:

Target populations can be divided into discrete subpopulations, or strata, on which to increase/decrease sample size; or selection probabilities can vary along environmental or other gradients. A stream network's elevation could be used to divide the population into elevation strata, allocating an equal number of sites per stratum (likely yielding inclusion probabilities that vary by stratum because the amount of stream length in each stratum likely would vary). Or the elevation gradient could be used in the GRTS site selection algorithm such that an "elevation balanced" sample could be drawn, each site likely with a different inclusion probability.

Over-sample:

In most environmental surveys, it will not be possible to sample every site in the selected sample. This may be due to lack of access permission, dangerous access, bad weather, or the selected site may be non-target. Whatever the reason, there is frequently a need to add additional sites to the original sample to achieve an acceptable number of samples. The most convenient way to do this rigorously with a GRTS sample is to pre-select “over-sample” sites at the same time as the base sample is selected. These over-sample sites can then be used, in sequence, to “replace” sites that were inaccessible or non-target. Because of the spatial balance property of a GRTS sample, the resulting sample of base sites plus over-sample sites will be a spatially balanced sample of the accessible population.

Panels:

Designing surveys to estimate status and trends requires the parsimonious allocation of field sampling across space (the stream network) and time (usually years). Knowledge of spatial and temporal variation of the indicator of choice is critical for the efficient allocation of visits to new sites, or to revisits to existing sites. For example, given a fixed total sampling effort, should I sample more sites during an annual index window (e.g., the spawning season), or should I sample more sites? If I am interested in change or trend detection, should I revisit all sites annually, or would a revisit pattern across years (in which not all sites are sampled every year) be more efficient? Often, the most efficient design for one type of question is not the most efficient for another type of question. Sampling more sites, at the expense of revisits to sites, improves the precision for estimating status (i.e., abundances) or spatial distribution. In contrast, revisiting sites across years generally improves change or trend detection capability. To meet these differing needs for site-visits, statisticians have proposed designs that allow balancing the need for more sites for status estimation with the need revisits to sites for change/trend detection. These designs consist of panels of sites, each panel with a particular pattern of visits across years. A simple design consists of one panel in which all sites are visited every year. A slightly more complex design consists of an annual panel (sites are visited each year), panels that are visited on a specified cycle, and panels of randomly selected sites each year. For example, a four panel design could consist of: an annual panel, a year 1 panel of sites visited every year, starting with year 1, a year 2 panel of sites visited every year, starting in year 2, a year 3 panel of sites visited every year, starting in year 3. These are sometimes called rotating panel designs. Urquhart and Kincaid (1999) and McDonald (2003) give examples of a variety of panel designs, and McDonald (2003) proposes a nomenclature for panel designs.

7 Index