Alaska Cook Inlet Lakes
Survey Design

Contact:

Doug Dasher
Alaska Department of Environmental Conservation
610 University Avenue
Fairbanks, AK 99709
Phone +1 (907)-451-2172
Fax +1 (907)-451-5146
doug_dasher@dec.state.ak.us

Description of Sample Design

Target population: All non-glacier fed lakes within the Cook Inlet basin, Alaska.

Sample Frame: To identify the target population streams, USGS/US EPA's National Hydrography Database (NHD) was used as the sampling frame. The polygon shapefile was used to create a point shapefile based on the lake centroids. The sample frame is restricted to lakes identified as greater than 4 hectares. An additional attribute was added that identified glacier and non-glacier lakes by Ted Moran, USGS Alaska Science Center

For more information on NHD and NHDPlus please check the websites: http://nhd.usgs.gov or http://www.horizon-systems.com/nhdplus.

Survey Design: A Generalized Random Tessellation Stratified (GRTS) survey design for a finite resource was used. The GRTS design includes reverse hierarchical ordering of the selected sites.

Multi-density categories: Two attributes are used to create continuous inclusion (selection) probabilities. Lake area in hectares was categorized as (4,10], (10, 20], (20, 50], (50, 100], and > 100 hectares. The second is proportional to the inverse lake density, estimated by a two-dimensional statistical density estimator, kdhe, from the MASS library. The two inclusion probabilities were multiplied and then adjusted to ensure an equal expected sample size for the lake area categories.

Stratification: Non-glacier fed lakes

Panels: None

Expected sample size: Expected sample size 50 sites with 10 in each lake area category.

Over sample: 200% (100 lakes).

Site Use: The base design has 5 lakes. Lakes are listed in SiteID order and must be used in that order. All sites that occur prior to the last site used must have been evaluated for use and then either sampled or reason documented why that site was not used. As an example, if 50 lakes are to be sampled and it required that 55 lakes be evaluated in order to locate 50 sampleable lakes, then the first 55 lakes in SiteID order would be used.

Sample Frame Summary

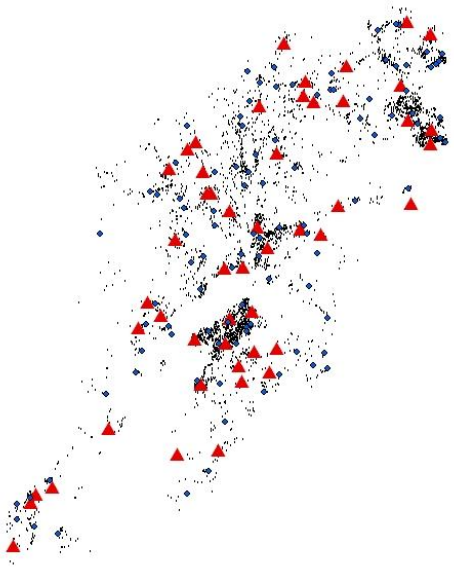
Total stream length in the sample frame is 7024.9 km Wadeable (1st to 4th order).

	stratum	
	Glacial Fed	Non-Glacial Fed
(4,10]	5	2338
(10,20]	12	931
(20,50]	12	664
(50,100]	6	202
>100	20	128
Total	55	4263

Site Selection Summary

Number of sites in sample

	BaseSites	OverSamp	Sum
[4,10]	10	22	32
(10,20]	8	21	29
(20,50]	14	22	36
(50,100]	8	21	29
(100,3e+04]	10	14	24
Sum	50	100	150



Description of Sample Design Output:

The tab-delimited ASCII file (Can be directly read into Excel) has the following variable definitions:

Variable Name	Description
SiteID	Unique site identification (character)
x	x-coordinate from map projection (see below)
y	y-coordinate from map projection (see below)
mdcaty	Multi-density categories used for unequal probability selection
weight	Weight (in square km), inverse of inclusion probability, to be used in statistical analyses
stratum	Strata used in the survey design
panel	Identifies base sample by panel name and Oversample by OverSamp
EvalStatus	Site evaluation decision for site: TS: target and sampled, LD: landowner denied access, etc (see below)
EvalReason	Site evaluation text commment
auxiliary variables	Remaining columns are from the sample frame (NHDPlus)

Projection Information

PROJCS["Alaska_Albers_Equal_Area_Conic",
GEOGCS["GCS_North_American_1983",
DATUM["D_North_American_1983",
SPHEROID["GRS_1980",6378137.0,298.257222101]],
PRIMEM["Greenwich",0.0],
UNIT["Degree",0.0174532925199433]],
PROJECTION["Albers"],
PARAMETER["False_Easting",0.0],
PARAMETER["False_Northing",0.0],
PARAMETER["Central_Meridian",-154.0],
PARAMETER["Standard_Parallel_1",55.0],
PARAMETER["Standard_Parallel_2",65.0],
PARAMETER["Latitude_Of_Origin",50.0],
UNIT["Meter",1.0]]

Evaluation Process

The survey design weights that are given in the design file assume that the survey design is implemented as designed. Typically, users prefer to replace sites that can not be sampled with other sites to achieve the sample size planned. The site replacement process is described above. When sites are replaced, the survey design weights are no

longer correct and must be adjusted. The weight adjustment requires knowing what happened to each site in the base design and the over sample sites. EvalStatus is initially set to "NotEval" to indicate that the site has yet to be evaluated for sampling. When a site is evaluated for sampling, then the EvalStatus for the site must be changed. Recommended codes are:

EvalStatus Code	Name	Meaning
TS	Target Sampled	site is a member of the target population and was sampled
LD	Landowner Denial	landowner denied access to the site
PB	Physical Barrier	physical barrier prevented access to the site
NT	Non-Target	site is not a member of the target population
NN	Not Needed	site is a member of the over sample and was not evaluated for sampling
Other codes		Many times useful to have other codes. For example, rather than use NT, may use specific codes indicating why the site was non-target.

Statistical Analysis

Any statistical analysis of data must incorporate information about the monitoring survey design. In particular, when estimates of characteristics for the entire target population are computed, the statistical analysis must account for any stratification or unequal probability selection in the design. Procedures for doing this are available from the Aquatic Resource Monitoring web page given in the bibliography. A statistical analysis library of functions is available from the web page to do common population estimates in the statistical software environment R.

For further information, contact

Anthony (Tony) R. Olsen
USEPA NHEERL
Western Ecology Division
200 S.W. 35th Street
Corvallis, OR 97333
Voice: (541) 754-4790
Fax: (541) 754-4716
email: Olsen.Tony@epa.gov

Bibliography:

Diaz-Ramos, S., D. L. Stevens, Jr, and A. R. Olsen. 1996. EMAP Statistical Methods Manual. EPA/620/R-96/002, U.S. Environmental Protection Agency, Office of Research and Development, NHEERL-Western Ecology Division, Corvallis, Oregon.

Stevens, D.L., Jr. 1997. Variable density grid-based sampling designs for continuous spatial populations. Environmetrics, 8:167-95.

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 A. R. Olsen. 2003. Variance estimation for spatially balanced samples of environmental resources. Environmetrics **14**:593-610.

Stevens, D. L., Jr., and A. R. Olsen. 2004. Spatially-balanced sampling of natural resources in the presence of frame imperfections. Journal of American Statistical Association:99:262-278.

Strahler, A.N. 1957. Quantitative Analysis of Watershed Geomorphology. Trans. Am. Geophys. Un. 38,913-920.

Web Pages: http://www.epa.gov/nheerl/arm
http://nhd.usgs.gov
http://www.horizon-systems.com/nhdplus