March 10, 2004

Alaska Tanana Basin Wadeable Streams Design

Contact:

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

Description of Sample Design

Target population: All wadeable perennial streams within the Tannana Basin in Alaska.

Sampling Frame: National Hydrologic Dataset (NHD) for the Tannana Basin in Alaska. NHD for this basin has not yet undergone the process of updating and review that has been completed for most of the United States. It does not have Strahler order available as an attribute. USGS in Alaska initiated the calculation of Strahler order and determined that sufficient problems were present that it was not feasible to complete within the time frame required. Other attributes in NHD will be used as surrogates for stream size and importance. The categories will be: (1)Named rivers and streams including headwater and start reaches that are named, (2) Headwater and Start reaches excluding any named reaches , (3) In network streams (0 and 1) that are not in either category 1 or 2, and (4) other reaches, non-networked and isolated reaches. In the documentation please explicitly describe the NHD codes used to create these categories. Note the fourth category will be excluded from the frame for sampling purposes.

Survey Design: GRTS survey design for a linear network with RHO. GRTS: Generalized Random Tessellation Stratified. RHO: Reverse Hierarchical Order.

Stratification: None
Panels: None

Multi-Density Categories: Use the first 3 categories described under the sample frame. For convenience refer to them as: Headwater, Named River, and Other Network reaches. The fourth category will not be included.

Sample size: 50 total base samples. Allocate to Multi-Density Categories as follows:

Headwater - 12, Named River - 25, Other Network - 13. Expectation is that more Headwater streams will be replaced due to being non-perennial (NonTarget) than streams in the other categories and that more Other network streams will be non target than named rivers. If this is the case, then more than 50% of the sites will be on Named
Rivers and less than 25% will be in the other two categories. Rationale for assigning more sites to Named Rivers is that they are most likely to be streams and rivers that are of interest.

Over sample: 300% (150 sites) for a total of 200 sites.

Sample Frame Description (USGS National Hydrography Dataset – High Resolution data for Alaska)

The Cataloging Units 19040501 – 19040509 and 19040511 (available from the NHD ftp site as of February, 2004) were appended using the "append_NHD tool 2.27". This created one coverage for the Tanana Basin. An attribute for the sample categories listed above was added to the coverage. No coverage was available for CU 19040510 in the southwest portion of the Tanana River Basin, therefore, no streamsin this CU were used for selection.

The attributes were assigned to the reach (RCH) route in the following order:

1. Named routes were selected (variable "NAME" contained any character string) and assigned sample category 1.
2. From the remaining reaches, using the related NHD.rflow table (rch2tflw relate), routes where "delta_lvl" = -9999 were assigned sample category 2.
3. From the remaining reaches, routes where "delta_lvl" is greater or equal to 0 were assigned sample category 3.
4. All remaining reaches were assigned sample category 4.

Description of Sample Design Output:

To achieve an expected sample size of sites in the target population, an appropriate sample size was selected for each separate study area. The extra/reserve samples are to provide alternate sites for sample sites not conforming to target population rules (e.g. non-wadeable, mis-mapped features) or being inaccessible due to safety concerns or denied access by landowners. The design has a base set of 50 samples spread over a single panel. The design has a 300% oversample, i.e. 150 reserve sites, for a total of 200 potential sample sites. The oversample sites should be added, as needed, in numerical order.

The following tables show the distribution of the frame information, as well as the sample sites, for each separate design (stratum).

Sums of Multipliers for Sample Design
AK Tanana Basin streams 2004

sampclass	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	55323.61	50.00	55323.61	50.00
2	26555.97	24.00	81879.57	74.00
3	28774.57	26.00	110654.1	100.00

clasmult	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1	26555.97	24.00	26555.97	24.00
1.571	28774.57	26.00	55330.54	50.00
3.262	55323.61	50.00	110654.1	100.00

MD_CATY values for AK Tanana Basin streams 2004
sums are in meters

md_caty	numbering scheme for design strata	clasmult	sampclass	sum of units	sum of weighted units
1	1	1.000	2	26555965.52	26555965.52
2	1	1.571	3	18316088.14	28774574.47
3	1	3.262	1	16960026.48	55323606.37

Selected Sites for AK Tanana Basin streams 2004

Table of nest_id by oversamp

nest_id
oversamp(If 0: routine site; Else a reserve site)

Frequency	0	1	Total
1	50	150	200
Total	50	150	200

Table of sampclass by oversamp

sampclass
oversamp(If 0: routine site; Else a reserve site)

Frequency	0	1	Total
1	24	68	92
2	11	41	52
3	15	41	56
Total	50	150	200

The attached comma-delimited, ASCII file (akw04450.csv) has the following variable definitions:

Site_ID	Sample Identifier assigned to each site 14 characters
Site Name	Name (if provided) 30 characters
Long-DD	Longitude, Decimal Degrees 12.6 numeric
Lat-DD	Latitude, Decimal Degrees 12.6 numeric
Stratum	Stratum (1 stratum defined in this design) 2 integer
Panel	Used if Multiple years/seasons/etc. sample 1 integer
Oversamp	Defines “backup” or “oversample” sites. 1 = oversample site, 0 = expected sample site 1 integer
Division	Division breaks down panels and expected/replicate sites. 1 integer
MD_Caty	Multi-Density weight category - defined above. 2 integer
Nest_ID	More than one if multiple levels/classes of samples drawn. 1 integer
Nest1	Defines sites within this nest class (1 = within, 0 = not in) 1 integer
Nest1_N	Expected number of samples for initial design categories. 2 integer
Nest1_wt	Initial Design weight for the site. 12.7 numeric
Long-DMS	Longitude, Degrees Minutes Seconds 20 characters
Lat-DMS	Latitude, Degrees Minutes Seconds 20 characters

The location information (if provided) is based on the 1927 North American Datum for projection. The Arc/INFO export files, if delivered with these data, have the following projection parameters:

Projection	Albers
Datum	NAD27
Spheroid	Clarke 1866
Units	Meters

1st projection parallel	55.0
2nd projection parallel	65.0
central meridian	-154
projection origin	50
false easting	0.0
false northing	0.0

Description of Statistical Analysis:

The statistical analysis of the data requires the weighting and stratification variables be used, even if computation of descriptive statistics (means, medians, standard errors, etc.) is all that is desired. After fieldwork and sampling, information on sampled and unsampled sites, along with reasons for non-sampling, need to be used to adjust sample weights. Otherwise, incorrect estimates for the target population will occur. See references for estimation procedures, or contact emapdesign@mail.cor.epa.gov.

For any questions about these data, please contact:

Anthony R. Olsen
US EPA National Health and Environmental Effects Research Laboratory
Western Ecology Division
200 SW 35th Street
Corvallis, Oregon 97333
Voice: 541-754-4790 Fax: 541-754-4716
email: olsen.tony@.epa.gov

Bibliography:

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

US Department of the Interior – US Geological Survey, National Mapping Division and the US Environmental Protection Agency (1999) Standards for National Hydrography Dataset – High Resolution: DRAFT.