

Columbia Basin Master Sample

Contact:

Phil Larsen
US EPA
NHEERL-WED
541-754-4362
Larsen.phil@epa.gov

Description of Survey Design

Target population: All stream and river channels within USGS Hydro Region 17 and portions of Hydro Regions 16 and 18 that are within boundaries of Oregon and Idaho. This covers all of Idaho, Oregon and Washington as well as small sections of Montana, Wyoming, Utah, Nevada and California that are part of the Columbia River Watershed.

Sample Frame: Sample frame was constructed in two parts: (1) NHD Plus (downloaded August, 2005) where all segments within NHD Plus are included and (2) NHD Plus (downloaded March 2010) where all segments are included. Strahler order and other NHD Plus attributes were added from the October 2008 release. This newest version had directional corrections to the original dataset in places, as well as adjustments to other NHDPlus features not used directly for sample frame design (see release notes for NHD Plus – Pacific Northwest).

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

Multi-density categories: None.

Stratification: The master sample is stratified by Hydro Region 17, Hydro Region 16 & 18 within Oregon, and Hydro Region 16 within Idaho. Stratification was required when a decision was made to include the portions of Hydro Regions 16 and 18 that are within boundaries of Oregon and Idaho.

Panels: None.

Expected sample size: The sample size was selected to have a sample site approximately every 1 km. This results in a weight approximately equal to 1 for each sample site. The number of sites in Hydro Region 17 is 532,268, Oregon supplemental stratum (OR_S) is 11,793 and Idaho supplemental stratum (ID_S) is 6,985.

Over sample: None.

Site Use: Sites are listed in siteID order within each stratum and must be used in that order within each stratum. All sites that occur prior to the last site used within a stratum

must have been evaluated for use and then either sampled or reason documented why that site was not used. As an example, if 50 sites are to be sampled within a stratum and it required that 61 sites be evaluated in order to locate 50 stream sites able to be sampled, then the first 61 sites in siteID order would be used.

Sample Frame Summary

The Hydro Region 17 sample frame has 532,267.94 km of stream channel length. The Idaho supplemental sample frame stratum has 6,984.832 km of stream channel length. The Oregon supplemental sample frame stratum has 11,792.76 km of stream channel length. Note that in order to select all sites for a state, make sure to query for border river sites which are denoted by ST:ST in the State attribute field (i.e. sites on the Columbia River would be denoted as OR:WA). Sites on border rivers are treated the same way in the County field.

Description of Sample Design Output:

The output is provided as a shapefile for the sites. Note that the “.dbf” file may be read in Excel. The attributes are as follows:

Variable Name	Description
SITEID	Unique site identification (character)
ALBERS_X	x-coordinate from map projection (see below)
ALBERS_Y	y-coordinate from map projection (see below)
LAT_DD	Latitude in decimal degrees
LON_DD	Longitude in decimal degrees
LAT_DMS	Latitude in degree minutes seconds
LON_DMS	Longitude in degree minutes seconds
WEIGHT	Weight (in km), inverse of inclusion probability, to be used in statistical analyses
STRATUM	Strata used in the survey design
COMID	Common Identifier of NHD feature
REACHCODE	NHD reach code assigned to feature
MEASURE	Distance along reach of sample point (as percent along reach beginning at bottom of reach)
FTYPE	NHD feature type
FCODE	Numeric code for feature attributes in NHD
DES_FTYPE	Modified Ftype for additional design information specifying intermittent, perennial, and artificial line types (lake or stream). Additional fields are: 'ArtificialPathLake', 'ArtificialPathStreamRiver', 'ArtificialPathStreamRiver Perennial', 'StreamRiverPerennial', and 'StreamRiver Intermittent'.
SO	Strahler stream order derived from NHDPlus using improved Strahler order algorithm
SC	Strahler calculator value from NHDPlus used in improved Strahler order algorithm
ELEV_M	Elevation at point in meters derived from National Elevation Data DEM (NED)
GNIS_ID	Geographic Names Information System ID for the value in GNIS_NAME

GNIS_NAME	Feature name from the Geographic Names Information System
STATE	State point lands in – NOTE: if point on a border river, state includes both states, i.e. OR:WA
LENGTHKM	Stream reach length in kilometers
COUNTY	County point lands in – NOTE: if point on a border river, COUNTY includes both counties, i.e. Multnomah:Clark
HUC_2	1st Level USGS Hydrologic Unit Code
HUC_4	2nd Level USGS Hydrologic Unit Code
HUC_6	3rd Level USGS Hydrologic Unit Code
HUC_8	4th Level USGS Hydrologic Unit Code
HUC_10	5th Level USGS Hydrologic Unit Code
HUC_12	6 th Level USGS Hydrologic Unit Code
LEVEL4	Level 4 ecoregion codes for the US
LEVEL4_NM	Level 4 ecoregion names
LEVEL3	Level 3 ecoregion codes for the US
LEVEL3_NM	Level 3 ecoregion names
CEC_L2	Level 2 ecoregion codes for North America
CECL2_NM	Level 2 ecoregion names
CEC_L1	Level 1 ecoregion codes for North America
CECL1_NM	Level 1 ecoregion names
COL_BASIN	'Yes' or 'No' indicating if site is within the Columbia basin
WRIA_NM	Washington WRIA regions
SA_REC_REG	Washington Salmon recovery regions

Projection Information

```

PROJCS["USA_Contiguous_Albers_Equal_Area_Conic_USGS_version",
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",-96.0],
PARAMETER["Standard_Parallel_1",29.5],
PARAMETER["Standard_Parallel_2",45.5],
PARAMETER["Latitude_Of_Origin",23.0],
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.

Web Pages:

Aquatic Resource Monitoring <http://www.epa.gov/nheerl/arm>

NHD web page <http://nhd.usgs.gov>

NHD Plus web page (July 2006) <http://www.horizon-systems.com/nhdplus>