# Idaho Arsenic Survey Design 2018

## Contact:

Jason Pappani
Water Quality Standards Lead
Idaho Department of Environmental Quality
1410 N Hilton
Boise, Idaho 83706
(208) 373-0515
Jason.Pappani@deq.idaho.gov

## Description of Survey Design

**Target population:** All streams and rivers within Idaho excluding streams on Native American Reservations and Wilderness areas.

**Sample Frame:** DEQ provided the shapefile that was used as basis for the sample frame (305bPerennielStreamsC). All perennial streams were included. Streams that occur on Native American Reservations and Wilderness areas were excluded.

**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:** Categorized by stream size based on Strahler Order (Small = $1^{st} - 5^{th}$, Large = $6^{th} - 8^{th}$), with unequal probability selection.

**Stratification:** Stratified by six Idaho administrative basins (Bear, Clearwater, Panhandle, Salmon, Southwest, and Upper Snake).

**Panels:** None.

**Expected sample size:** Total of 24 sites.

**Over sample:** 500% (120 sites) for a total of 144 sites.

**Site Use:** Sites are listed in siteID order and must be used in that order *within each stratum (administrative basin)*. 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 24 sites are to be sampled and it required that 50 sites be evaluated in order to locate 24 sites able to be sampled, then the first 50 sites in siteID order would be used.

# Sample Frame Summary

Stream and River length in kilometers

|              | 01   | 02   | 03   | 04   | 05  | 06  | 07  | 08 |
|--------------|------|------|------|------|-----|-----|-----|----|
| Bear River   | 292  | 332  | 189  | 82   | 85  | 124 | 0   | 0  |
| Clearwater   | 1901 | 1351 | 806  | 509  | 278 | 16  | 0   | 0  |
| Panhandle    | 2046 | 1376 | 782  | 330  | 230 | 1   | 0   | 2  |
| Salmon       | 3628 | 2037 | 1221 | 499  | 284 | 185 | 121 | 0  |
| Southwest    | 3035 | 1843 | 1386 | 751  | 386 | 643 | 10  | 21 |
| Upper Snake  | 4068 | 2367 | 1654 | 1254 | 747 | 653 | 26  | 0  |

|         | 1st_5th  | 6th_8th | Sum      |
|---------|----------|---------|----------|
| BEAR    | 1863.26  | 142.86  | 2006.11  |
| CLEAR   | 10515.32 | 25.35   | 10540.67 |
| PANHAND | 9701.06  | 4.27    | 9705.33  |
| SALMON  | 17047.26 | 296.37  | 17343.64 |
| SWEST   | 14786.02 | 530.00  | 15316.02 |
| USNAKE  | 15547.21 | 574.10  | 16121.31 |
| Sum     | 69460.13 | 1572.95 | 71033.07 |

# Site Selection Summary

Number of sites by Basin and Strahler order

Number of Base sites

|         | 1st_5th | 6th_8th | Sum |
|---------|---------|---------|-----|
| BEAR    | 3       | 1       | 5   |
| CLEAR   | 3       | 1       | 5   |
| PANHAND | 3       | 1       | 5   |
| SALMON  | 3       | 1       | 5   |
| SWEST   | 3       | 1       | 5   |
| USNAKE  | 3       | 1       | 5   |
| Sum     | 18      | 6       | 5   |

# 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:

| SiteID     | Unique site identification (character)                                              |
|------------|-------------------------------------------------------------------------------------|
| xcoord     | x-coordinate from map projection                                                    |
| ycoord     | y-coordinate from map projection                                                    |
| mdcaty     | Multi-density categories used for unequal probability selection                     |
| wgt        | Weight (in 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 provided |

## 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 |
| IA | Inaccessible | physical barrier or distance prevented access to the site |
| NT | Non-Target | site is not a member of the target population, i.e., not a stream or river (may be marsh, dry channel, etc.) |
| 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.