National Lake Assessment
Survey Design: 2007

Contact:

Susan Holdsworth
Office of Water
OWOW Monitoring Branch
USEPA (4503T)
1200 Pennsylvania Ave. NW
Washington, DC 20460
(202) 566-1187
Email: Holdsworth.Susan@epa.gov

Description of Sample Design

Target population: All lakes, reservoirs, and ponds within the 48 contiguous United States greater than 1 hectare in surface area that are permanent waterbodies. Note that the minimum surface area actually implemented may be increased to either 4 hectares or 10 hectares. The word "lake" in the remainder of this document includes lakes, reservoirs and ponds. Lakes that are saline are excluded as are those used for aquaculture, disposal-tailings, sewage treatment, evaporation, or other unspecified disposal use.

Sample Frame: The sample frame was derived from the National Hydrography Dataset (NHD).

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

Multi-density categories: A complex unequal probability selection process was used to allocate number of lakes to be sampled to a number of lake categories. First, 791 lakes were identified as lakes sampled during the 1970s National Lake Eutrophication study (NES). Two lake area categories, (1,4] and (4,10] hectare ranges, were defined (excluding any NES lakes). All remaining lakes were categorized (1) by four lake area size categories, (10, 20], (20, 50], (50, 100], and >100 hectares; (2) by 48 states, and (3) by nine (9) aggregated Omernik Level 3 ecoregions. This resulted in 427 unique lake categories for the unequal probability design

Stratification: No stratification

Panels: Design uses a single panel with an over sample.

Expected sample size: Expected sample size 1000 lakes including all lakes/reservoirs greater than 1 hectare in surface area in the sample frame. Of the 1000

lakes, the expected sample size for NES lakes is 100; for lakes (1, 4] is 90, for lakes (4, 10] is 90, and 720 for the remaining 424 categories. The 720 lakes were allocated so that initially (1) 180 would occur in each of the four lake surface area categories of (10, 20], (20, 50], (50, 100], and >100 hectares and (2) 15 would occur in each of the 48 states. Assuming independence of lake area and states resulted in an expected number of lakes for each state and lake area category combination. These were then adjusted to have an expected sample size of 80 lakes within each of the nine aggregated ecoregions.

Over sample: 4000 lakes were selected as potential replacement lakes. The large over sample size was done to accommodate those states who may want to increase the number of lakes sampled within their state for a state-level design.

Site Use: Each lake selected to be sampled is given unique site identification (siteID) that consists of two parts, NLA06608, and a number between 1 and 5000. It critical this siteID be used in its entirety to make sure that the lakes are correctly identified. Within each state, lakes evaluated for potential sampling must have all siteIDs from the largest to the lowest number evaluated. For example, if NLA06608-0235 is the largest siteID evaluated within a state, then all siteID that are lower than "0235" within the state must be evaluated. Even more critical is that if NLA06608-0235 is the largest siteID that is actually sampled in the field, then all lower siteIDs within the state that are evaluated to be a target lake and is accessible must be sampled in the field.

If lakes (1, 4] or (4, 10] hectares in size are included in the target population, then only those lakes in the base sample of 1000 lakes (panel=Panel_1 lakes) will be evaluated. If one of these lakes can not be sampled, then it will not be replaced by another lake.

Sample Frame Summary

Number of NES lakes by lake surface area category

	# Lakes
(0,1]	0
(1,4]	2
(4,10]	8
(10,20]	5
(20,50]	44
(50,100]	62
>100	670
Total	791

Number of lakes by lake surface area and Three aggregated Ecoregions

	WMTNS	EHIGH	PLNLOW	OTHER	Sum
(0,1]	0	0	0	0	0
(1,4]	14272	28156	112773	0	155201
(4,10]	6951	9591	52017	0	68559
(10,20]	2644	3696	18562	0	24902
(20,50]	1684	2595	12209	0	16488
(50,100]	605	1085	4444	0	6134
>100	1058	1412	4886	0	7356
Sum	27214	46535	204891	0	278640

Number of lakes by lake surface area and nine aggregated ecoregions

```
        (0,1]   (1,4]  (4,10] (10,20] (20,50] (50,100] >100     Sum
  CPL       0   45228   20781    7018    4199     1334  1475   80035
  NAP       0   10847    5335    2421    1923      838  1047   22411
  NPL       0   11363    5137    1438     764      286   235   19223
  SAP       0   17309    4256    1275     672      247   365   24124
  SPL       0   15536    6006    1863     881      252   351   24889
  TPL       0   23213    9114    3041    2026      775   809   38978
  UMW       0   17433   10979    5202    4339     1797  2016   41766
  WMT       0    9537    4804    1763    1039      351   598   18092
  XER       0    4735    2147     881     645      254   460    9122
  OTH       0       0       0       0       0        0     0       0
  Sum       0  155201   68559   24902   16488     6134  7356  278640
```

Number of lakes by state and lake surface area

```
        (0,1]   (1,4]  (4,10] (10,20] (20,50] (50,100] >100     Sum
  AL        0    4781    1440     359     179       67    55    6881
  AR        0    2768    1445     570     362      118   143    5406
  AZ        0     276     124      56      45       22    39     562
  CA        0    3051    1479     586     346      112   240    5814
  CO        0    2305    1073     332     202       72   117    4101
  CT        0    1071     376     152     128       51    38    1816
  DE        0     209      62      45      32       10     3     361
  FL        0    7691    5213    1926    1296      451   476   17053
  GA        0    7290    2974     709     325       73    62   11433
  HI        0       0       0       0       0        0     0       0
  IA        0    3014     711     202     152       48    58    4185
  ID        0    1094     472     177     105       46    80    1974
  IL        0    4077    1635     524     285       95   102    6718
  IN        0    3224    1184     371     200       72    63    5114
  KS        0    6428    1239     272     115       33    50    8137
  KY        0    1433     434     172      73       25    27    2164
  LA        0    5881    3188    1180     720      280   386   11635
  MA        0    1427     742     372     313      130    86    3070
  MD        0     597     242      93      50        9    12    1003
  ME        0    1519     921     547     468      256   482    4193
  MI        0    5164    3132    1380    1133      429   450   11688
  MN        0    9312    5831    2876    2552     1114  1311   22996
  MO        0    3976    1514     378     165       51    47    6131
  MS        0    6389    1792     535     308       64    79    9167
  MT        0    5685    2284     599     301      127   162    9158
  NC        0    4544     846     266     145       62    94    5957
  ND        0    3925    3157    1314     882      355   258    9891
  NE        0    2278    1261     478     320      103    73    4513
  NH        0     622     434     234     161       98    77    1626
  NJ        0    1077     526     230     130       26    35    2024
  NM        0     521     167      68      50       25    39     870
  NV        0     204     127      70      67       20    72     560
  NY        0    3576    1860     783     546      201   255    7221
  OH        0    3148     931     217     123       53    76    4548
  OK        0    4550    1749     472     212       61    89    7133
  OR        0     999     554     237     171       59   138    2158
  PA        0    2182     835     324     239       80    71    3731
  PR        0       0       0       0       0        0     0       0
  RI        0     226     115      53      49       25    20     488
  SC        0    3755    1137     322     153       32    44    5443
  SD        0    5093    2152     600     382      135   152    8514
```

```
TN       0    1030     467     175     114       33    49    1868
TX       0   14937    5880    1943     929      245   346   24280
UT       0     857     560     206     115       65   110    1913
VA       0    2912     801     263     146       39    45    4206
VT       0     389     205      76      75       38    39     822
WA       0    2237    1086     395     289       93   124    4224
WI       0    4200    2929    1312    1070      450   487   10448
WV       0     339     111      27       8        4    12     501
WY       0    2938    1162     424     257       77    83    4941
Sum      0  155201   68559   24902   16488     6134  7356  278640
```

Site Selection Summary

Number of NES lakes by Lake surface area in sample

```
 Area Ca       Base        Sum
 (0,1]           0    0      0
 (1,4]           0    2      2
 (4,10]          1    1      2
 (10,20]         1    2      3
 (20,50]        11   16     27
 (50,100]        9   31     40
 >100           91  354    445
 Total         113  406    519
```

Number of lakes by Aggregated Ecoregion and lake surface area

Base Sample:

| | Lake surface area category (hectares) | | | | | | | |
	(0,1]	(1,4]	(4,10]	(10,20]	(20,50]	(50,100]	>100	Sum
CPL	0	25	39	19	23	19	19	144
NAP	0	4	7	15	21	26	25	98
NPL	0	9	7	30	22	22	11	101
SAP	0	16	5	22	16	12	29	100
SPL	0	5	12	28	19	13	30	107
TPL	0	13	10	31	28	16	31	129
UMW	0	12	16	11	19	20	35	113
WMT	0	4	8	18	19	25	44	118
XER	0	3	0	11	17	19	40	90
OTH	0	0	0	0	0	0	0	0
Sum	0	91	104	185	184	172	264	1000

Over Sample:

| | Lake surface area category (hectares) | | | | | | | |
	(0,1]	(1,4]	(4,10]	(10,20]	(20,50]	(50,100]	>100	Sum
CPL	0	126	100	96	102	61	101	586
NAP	0	28	23	65	86	100	111	413
NPL	0	33	34	104	85	74	59	389
SAP	0	23	27	100	64	73	126	413
SPL	0	32	36	107	62	66	105	408
TPL	0	61	38	94	91	83	137	504
UMW	0	41	50	64	101	94	137	487
WMT	0	28	28	82	76	60	115	389
XER	0	14	12	62	68	71	135	362

```
OTH        0      0       0        0        0        0     0     0
Sum        0    386     348      774      735      682  1026  3951
```

Total Sample:

```
              Lake surface area category (hectares)
        (0,1] (1,4] (4,10] (10,20] (20,50] (50,100] >100   Sum
  CPL      0   151    139     115     125       80   120    730
  NAP      0    32     30      80     107      126   136    511
  NPL      0    42     41     134     107       96    70    490
  SAP      0    39     32     122      80       85   155    513
  SPL      0    37     48     135      81       79   135    515
  TPL      0    74     48     125     119       99   168    633
  UMW      0    53     66      75     120      114   172    600
  WMT      0    32     36     100      95       85   159    507
  XER      0    17     12      73      85       90   175    452
  OTH      0     0      0       0       0        0     0      0
  Sum      0   477    452     959     919      854  1290   4951
```

Number of lakes by state and lake surface area

```
        OverSamp    Base   Sum
  AL         62      18    80
  AR         55      11    66
  AZ         74      21    95
  CA         90      22   112
  CO         96      30   126
  CT         49      14    63
  DE         29       9    38
  FL        107      23   130
  GA         63      22    85
  IA         72      19    91
  ID         69      21    90
  IL         87      20   107
  IN         96      23   119
  KS         97      20   117
  KY         60      13    73
  LA         75      17    92
  MA         40      11    51
  MD         40       5    45
  ME         61      11    72
  MI        139      31   170
  MN        200      49   249
  MO         79      20    99
  MS         48      15    63
  MT        180      44   224
  NC         65      21    86
  ND        179      43   222
  NE        112      32   144
  NH         48      13    61
  NJ         48      10    58
  NM         91      15   106
  NV         82      19   101
  NY         69      12    81
  OH         67      20    87
  OK        103      36   139
  OR         73      30   103
  PA         67      20    87
```

```
    RI         44      12    56
    SC         60       9    69
    SD        139      43   182
    TN         61      12    73
    TX        143      31   174
    UT         82      23   105
    VA         60      16    76
    VT         49       9    58
    WA         78      20    98
    WI        144      32   176
    WV         36       8    44
    WY         83      25   108
    Sum      3951    1000  4951
```

Number of lakes by state and lake area category for base sample only

	(0,1]	(1,4]	(4,10]	(10,20]	(20,50]	(50,100]	>100	Sum
AL	0	4	3	2	1	1	7	18
AR	0	1	3	0	3	2	2	11
AZ	0	0	0	2	7	4	8	21
CA	0	1	1	4	5	3	8	22
CO	0	0	2	7	4	6	11	30
CT	0	0	0	2	6	4	2	14
DE	0	1	0	1	1	5	1	9
FL	0	3	10	2	3	2	3	23
GA	0	8	6	1	3	0	4	22
HI	0	0	0	0	0	0	0	0
IA	0	1	1	8	4	2	3	19
ID	0	2	1	1	6	3	8	21
IL	0	3	1	4	5	2	5	20
IN	0	2	3	4	6	3	5	23
KS	0	2	1	7	4	3	3	20
KY	0	3	0	3	3	3	1	13
LA	0	3	4	0	1	4	5	17
MA	0	0	2	2	2	3	2	11
MD	0	1	0	3	1	0	0	5
ME	0	1	1	0	1	4	4	11
MI	0	2	6	1	6	7	9	31
MN	0	8	7	4	4	8	18	49
MO	0	1	6	6	2	3	2	20
MS	0	4	3	2	3	2	1	15
MT	0	4	4	10	6	10	10	44
NC	0	3	1	3	3	6	5	21
ND	0	5	3	8	13	9	5	43
NE	0	0	2	10	10	5	5	32
NH	0	0	1	3	0	5	4	13
NJ	0	2	0	1	4	1	2	10
NM	0	0	0	5	1	3	6	15
NV	0	0	0	0	2	4	13	19
NY	0	0	2	1	2	2	5	12
OH	0	1	1	4	3	4	7	20
OK	0	1	2	11	7	4	11	36
OR	0	0	2	4	5	6	13	30
PA	0	2	2	6	4	3	3	20
PR	0	0	0	0	0	0	0	0
RI	0	1	0	4	3	2	2	12

```
SC      0    3    3    1     0      0    2    9
SD      0    4    0   16    10      7    6   43
TN      0    0    0    3     5      1    3   12
TX      0    6   14    2     0      0    9   31
UT      0    0    1    2     1      6   13   23
VA      0    1    1    6     6      0    2   16
VT      0    0    0    0     3      3    3    9
WA      0    2    0    5     4      5    4   20
WI      0    3    2    4     6      5   12   32
WV      0    0    0    4     2      1    1    8
WY      0    2    2    6     3      6    6   25
Sum     0   91  104  185   184    172  264 1000
```

Description of Sample Design Output:

The dbf file for the shapefile ("NLA Lake Sites Final") has the following variable definitions:

Variable Name	Description
siteID	Unique identification label for each lake in the sample
xcoord	x-coordinate of lake centroid (see Albers projection information below)
ycoord	y-coordinate of lake centroid (see Albers projection information below)
mdcaty	inclusion probability used for the unequal probability selection of the sample lake
wgt	Weight (number of lakes) to be used in the statistical analyses. It is the inverse of the inclusion probability
stratum	Strata used in design. All equal to NLALake
panel	Panel_1 identifies the 1000 lakes in the base design. Oversamp identifies lakes to be used as replacements as necessary.
EvalStatus	Placeholder to record the results of the lake recon evaluation (see below)
EvalReason	Placeholder to record reason for the evalution result.
oid_	ID from original lake frame shapefile
rch_com__1	From original NHD lake frame shapefile
com_id	From original NHD lake frame shapefile
ftype	NHD ftype field
fcode	NHD fcode field
gnis_id	GNIS identification
name	Lake name
statecty	FIPS state and county code
st	State two-letter abbreviation
cntyname	County name
eco	Omernik level 3 ecoregion number (DO NOT USE)
level3	Omernik level 3 ecoregion number (USE THIS for ecoregion)
level3_nam	Omernik level 3 ecoregion name (DO NOT USE; not all are present)

huc_12__14	USGS hydrologic catalog code for 12-14 digit units
cu_name	USGS hydrologic catalog unit name
region	USGS hydrologic region number
storetassi	Storet number for NES lakes
area_ha	Area in hectares calculated from polygon lake frame in Albers projection
epa_reg	EPA region number
wsa_eco9	Name of nine aggregated Omernik level 3 ecoregion
wsa_eco3	Name of three aggregated Omernik level 3 ecoregion
area_cat	Detailed lake surface area categories in hectares
area_cat3	Three category lake surface area categories in hectares
area_cat5	Five category lake surface area categories in hectares
area_cat7	Seven category lake surface area categories in hectares
dsgn_cat	Multi-density categories used in unequal probabilty survey design
inc_nw	Inclusion probability based on nationwide sample of (1, 10] hectare lakes
inc_st	Inclusion probability based on states
inc_cat7	Inclusion probability based on seven area categories
inc_st7	Inclusion probability based on combined state and seven area categories
inc_9st7	Inclusion probability based on combined nine aggregated ecoregions,
inc_nes	Inclusion probability based on NES lakes
neslake	Identifies NES lake

Projection Information

PROJCS["USA_Contiguous_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",-96.0],
PARAMETER["Standard_Parallel_1",29.5],
PARAMETER["Standard_Parallel_2",45.5],
PARAMETER["Latitude_Of_Origin",37.5],
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.

Horn, C.R. and Grayman, W.M. (1993) Water-quality modeling with EPA reach file system. Journal of Water Resources Planning and Management, 119, 262-74.

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

Web Page: http://www.epa.gov/nheerl/arm