

US EPA ARCHIVE DOCUMENT

Example Design and Assignment of Weights: Indiana Rotating Basin Stream Sample - 1998

Indiana Upper Wabash River Basin – Note: The results of sampling at these sites are reported in: IDEM, 2000, Indiana Water Quality Report 2000, Some project specific details have been simplified for this example and results do not exactly match those in the Indiana 200 report.

Study Background: Selecting sampling sites for probability-based sampling program.

- State-wide Monitoring based on 5 major basins, one basin sampled per year
- Upper Wabash River Basin to be sampled in 1988.
- All perennial streams (Strahler order 1-4) contained in this basin within the geographic boundaries of Indiana.
- Sample design is weighted by Strahler stream order categories to achieve an approximately equal expected sample size across stream order.
- Equal stratification between 1st, 2nd, 3rd, and 4th order streams.
- Target sample size of 50 sites with an overdraw of 50 sites, total of 100 sites for basin.

Design Details:

Target Population: All perennial streams contained in the Upper Wabash River basin within the geographic boundaries of Indiana. Upper Wabash River Basin, is comprised of USGS hydrologic units 05120101, 05120102, 05120103, 05120104, 05120105, 05120106, and 05120107.

Sample frame: Includes all Strahler order streams within these cataloging units. US EPA's Reach File Version 3 - Alpha (RF3-Alpha) 1:100,000 scale with reaches coded "R" (regular reach), "S" (start/headwater reach), "T" (terminal reach), "N" (non-networked reach), and "W" (wide river - one bank only). For double-line "wide" rivers, only one bank (one side of the channel) was used for measuring and sampling.

Survey Design: GRTS for a linear continuous network with unequal probability and over-sample.

Multi-Density Categories: Strahler Orders 1st, 2nd, 3rd, and 4th. Approximately equal number of samples in each category.

Stratification: Note that this is a continuation of Indiana's rotating basin studies: important that the stratum number is unique.

Expected Sample size: 50 sites in basin

Over-sample: 50 in base sample and 50 over-sample for a total of 100 sites.

Panels: None

Nested sub-sampling: None

Intensive areas: None

IN Rotating Basins 1998

Frame Statistics:

Length in meters by Hydrologic Unit Catalog

HUC	Length	Percent	Cumulative Length	Cumulative Percent
5120101	1541271.0	20.1	1541271	20.1
5120102	576526.3	7.5	2117797	27.6
5120103	617796.3	8.1	2735593	35.7
5120104	1076210.0	14.0	3811803	49.7
5120105	1013818.0	13.2	4825621	63.0
5120106	1783637.0	23.3	6609258	86.2
5120107	1054239.0	13.8	7663497	100.0

Length in meters by Strahler Order

Order	Length	Percent	Cumulative Length	Cumulative Percent
1	4514450.0	58.9	4514450	58.9
2	1443260.0	18.8	5957710	77.7
3	740145.6	9.7	6697856	87.4
4	660294.2	8.6	7358150	96.0
5	187620.2	2.4	7545770	98.5
7	117727.1	1.5	7663497	100.0

Length in meters by Hydrologic Unit Catalog and Strahler Order

HUC			Strahler	Order			
5120101	1	2	3	4	5	7	Total
Length	930468	264845	139357	34492	109546	62563	1541271
Percent	12.14	3.46	1.82	0.45	1.43	0.82	20.11
Row Pct	60.37	17.18	9.04	2.24	7.11	4.06	
Col Pct	20.61	18.35	18.83	5.22	58.39	53.14	
5120102							
Length	386123	73194	19841	97368	0	0	576526
Percent	5.04	0.96	0.26	1.27	0.00	0.00	7.52

Row Pct	66.97	12.70	3.44	16.89	0.00	0.00	
Col Pct	8.55	5.07	2.68	14.75	0.00	0.00	
5120103							
Length	362733	125436	129628	0	0	0	617796
Percent	4.73	1.64	1.69	0.00	0.00	0.00	8.06
Row Pct	58.71	20.30	20.98	0.00	0.00	0.00	
Col Pct	8.03	8.69	17.51	0.00	0.00	0.00	
5120104							
Length	670495	221750	63985	119979	0	0	1076210
Percent	8.75	2.89	0.83	1.57	0.00	0.00	14.04
Row Pct	62.30	20.60	5.95	11.15	0.00	0.00	
Col Pct	14.85	15.36	8.64	18.17	0.00	0.00	
5120105							
Length	618908	208225	68558	62962	0	55164	1013818
Percent	8.08	2.72	0.89	0.82	0.00	0.72	13.23
Row Pct	61.05	20.54	6.76	6.21	0.00	5.44	
Col Pct	13.71	14.43	9.26	9.54	0.00	46.86	
5120106							
Length	961692	341582	216993	192353	71017	0	1783637
Percent	12.55	4.46	2.83	2.51	0.93	0.00	23.27
Row Pct	53.92	19.15	12.17	10.78	3.98	0.00	
Col Pct	21.30	23.67	29.32	29.13	37.85	0.00	
5120107							
Length	584031	208229	101782	153140	7057.4	0	1054239
Percent	7.62	2.72	1.33	2.00	0.09	0.00	13.76
Row Pct	55.40	19.75	9.65	14.53	0.67	0.00	
Col Pct	12.94	14.43	13.75	23.19	3.76	0.00	
Total							
Length	4514450	1443260	740146	660294	187620	117727	7663497
Percent	58.91	18.83	9.66	8.62	2.45	1.54	100.00

Sample Information:

To achieve an expected sample size of 50 sites in the target population, an increased sample size (100 samples) was selected for the basin study area. The over-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.

This is to ensure the survey design meets the sample size requirements when selected sites are not useable, the over-sample is generated. Whenever a sample site is not sampled, documentation should be completed, and a site from the over-sample, the next in the over-sample sequence, selected and surveyed.

Md_caty	Ordergrp	Weight *	Number of Sites	Cumulative
21	1	368.5265	25	25
22	2	117.8171	25	50
23	3	60.42005	26	76
24	4	53.90157	24	100

* Weight – each site represents this number of kilometers of stream.

Documentation:

The Design File (a comma-delimited, ASCII file) is created having the following variable definitions:

Site_ID	Sample Identifier assigned to each site
RF3Name	EPA's River Reach File V.3 Primary Name
Long-DD	Longitude, Decimal Degrees
Lat-DD	Latitude, Decimal Degrees
Stratum	Stratum (stratum # defined in this design)
MD_Caty	Multi-Density weight category - defined above.
Weight	Initial Design weight for the site, (kilometers stream/expected sample size).
RF3RCHID	EPA's River Reach File V.3 Reach ID
Strahler	Strahler Order - generated from RF3
County	County Name
Map24k	1:24,000 scale map name
Map100k	1:100,000 scale map name
Map250k	1:250,000 scale map name
Eco	Omernik Ecoregion Level III number (2000 version)

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

Projection = Albers
Datum = NAD27
Spheroid = Clarke 1866
Units = Meters
1st projection parallel = 29 degrees 30' 00"
2nd projection parallel = 45 degrees 30' 00"
central meridian = -96 degrees 00' 00"
projection origin = 23 degrees 00' 00"
false easting = 0
false northing = 0

Description of Statistical Analysis:

The statistical analysis of the data requires the weighting and stratification variables be used. Otherwise, incorrect estimates for the target population will occur. See references for estimation procedures, or contact ORD_EMAP_DESIGN@epamail.epa.gov

Bibliography:

Christensen C. 1999. 1998 Watershed Monitoring Program Study of the Upper Wabash River Basin. Indiana Department of Environmental Management, Office of Water Management, Assessment Branch, Surveys Section, Indianapolis, Indiana. IDEM 032/02/021/1999

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.

IDEM, 2000, Indiana Water Quality Report 2000, Indiana Department of Environmental Management, Office of Water Management, Planning and Restoration Branch, Indianapolis, Indiana. IDEM 34/02/001/2000

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.

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