Appendix A

Sample Site Selection Design for MAIA Estuaries 1997-98

Two basic sample site selection designs were implemented for the sampling in 1997-98 for MAIA Estuaries: for estuarine waters treated as a continuous resource, sampling stations were selected randomly within hexagons of a randomly overlaid grid; and for estuaries treated as discrete resources, a random selection from a list frame of individual estuarine systems was employed. The approach for continuous resource is the randomized-tessellation stratified (RTS) design (Stevens 1997). Multiple Density-RTS (MD-RTS) (Stevens 1997) approach was used for the Delaware Estuary. The basic estimation procedures are from the report by Heimbuch, et al. (1995), and also appear in an appendix in Strobel, et al. (1995). The estimation procedures are based upon ratio-type estimators, which constrain the estimates of the proportion of the area to be no greater than 1. This eliminates the possibility of estimates that are outside reasonable bounds and reduces the variance of the estimates.

Discrete resource estimation procedure:

For discrete resources, estimates of the proportion of the area and associated variances are computed based on a random selection from a list of the estuarine systems, with replicate samples taken from a subset of the selected systems (Cochran 1977). The resulting estimate of the proportion of the area is:

$$\hat{P}_z = \frac{\sum_{i=1}^{n} A_i \bar{y}_i}{\sum_{i=1}^{n} A_i} ,$$

where

$\hat{P}_z =$ estimated proportion of the area at or below response value of z,

$n =$ number of estuarine systems sampled.

$A_i =$ area of estuarine system i,

$\bar{y}_i = \frac{1}{m_i} \sum_{j=1}^{m_i} y_{ij}$, and

$y_{ij} = \begin{cases} 1 \text{ if response value is less than or equal to } z, \\ 0 \text{ otherwise} \end{cases}$

$m_i =$ number of samples in estuarine system i,

Since replicate samples may be obtained only at a subset of the sampled systems (only a subset of total number of systems is sampled), the formula for the estimated variance taken from Cochran (1977 eq. 11.30) was modified to produce the following estimate of the approximate mean squared error (MSE) of the estimate for the proportion of the area:

$$MSE(\hat{P}_z) = \frac{\frac{N^2}{n}(1-f_1)\frac{\sum_{i=1}^{n} A_i^2(\bar{y}_i - \hat{P}_z)^2}{n-1} + \frac{N}{n*}\sum_{i=1}^{n*}\frac{A_i^2 S_{2i}^2}{m_i}}{A^2}$$

where,

$f_1 = n/N$,
n^* = number of estuarine systems with replicate samples,

$$S_{2i}^2 = \frac{\sum_{j=1}^{m_i} (y_{ij} - \bar{y}_i)^2}{m_i - 1} ,$$

A = the total area of estuarine systems in the list frame, and
N = number of estuarine systems in the list frame.

Continuous resource estimation procedure:

For continuous resources, the estuarine waters are treated as spatially extensive and the Horvitz-Thompson estimator (Cochran 1977) is used. The proportion of the area estimate is:

$$\hat{P}_z = \frac{1}{A} \sum_{i=1}^{n} \frac{y_i}{\Pi_i} ,$$

where

$\hat{P}_z$ = estimated proportion of the area at or below response value of z,

$$y_i = \begin{cases} 1 & \text{if response value is less than or equal to } z, \\ 0 & \text{otherwise} \end{cases}$$

Π_i = inclusion probability for a station (1/area),
A = the total area of estuarine waters in strata, and
n = number of estuarine systems sampled.

To produce unbiased estimates of variance, joint inclusion probabilities, Π_{ij} , must be non-zero. The variance estimates were obtained by applying the Yates-Grundy estimate of variance (Cochran 1977) and using an approximation for the joint inclusion probability:

$$v\hat{a}r(\hat{P}_z) = \frac{1}{A^2} \sum_{i=1}^{n} \sum_{j>1}^{n} \left(\frac{\Pi_i \Pi_j - \Pi_{ij}}{\Pi_{ij}} \right) * \left(\frac{y_i}{\Pi_i} - \frac{y_i}{\Pi_j} \right)^2 ,$$

where

Π_{ij} = joint inclusion probability, probability that stations i and j are selected for sampling,

$$\frac{(n-1)}{n} \Pi_i \Pi_j \quad .$$

The approximation for joint inclusion probability assumes that sites are a simple random sample over the continuous population.

Procedure for combining estimates for strata

The estimate for the proportion of the area for multiple strata is the weighted average of the proportion of the area estimates of the individual strata:

$$\hat{U} = \sum_i w_i \hat{P}_i \quad ,$$

where,

$\hat{U}$ = the estimated proportion area for the combined strata,

$\hat{P}_i$ = the estimated proportion area for the stratum i, and

w_i = relative area of stratum i, (area of stratum i/total area for combined strata).

The estimated variance for the proportion of the area in the combination of strata is the sum of the component variances:

$$v\hat{a}r(\hat{U}) = \sum_i w_i \, v\hat{a}r(\hat{P}_i)$$

where,

$v\hat{a}r(\hat{U})$ = the estimated variance of the proportion of area in the geographic area of interest, and

$v\hat{a}r(\hat{P}_i)$ = the estimated variance of the proportion of area in stratum i.

These methods for estimation of the combined strata are based on the assumption that the strata are independent.

Procedure for combination intervals

To produce confidence intervals (CIs), the proportion of the area estimates are assumed to be normally distributed. The 95% CIs are based on:

$$\text{CI} = \pm 1.96 \sqrt{MSE\ (\hat{P}_z)} \quad .$$

For small sample sizes, the normal approximation may not be acceptable. A different form has to be used: exact binomial CIs. However, the exact binomial CI can be used only for constant inclusion probabilities. For variable inclusion probabilities, the distribution of the percent area is not binomially distributed, and normal approximate CIs are the more appropriate intervals to use regardless of the sample size.

Table A-1. Initial Strata Used in Selection of Sampling Sites for MAIA Estuaries 1997-98.

Major tributaries (mainstem portion only)

Delaware River
Choptank River
Patuxent River
Potomac River (benthic/sediment quality)
Lower Potomac River (water quality)
Upper Potomac River (water quality)
Rappahannock River
York River
James River
Chowan River
Neuse River

Chesapeake Bay subsystems

eastern tributaries
Pocomoke Sound (water quality)
Tangier Sound (water quality)
Tangier/Pocomoke Sound (benthic/sediment quality)

Small estuarine systems in MAIA

Intensively-sampled small estuarine systems

Schuylkill River
Salem River
Severn River
South River
Eastern Bay
Pocomoke River
St. Jerome Creek
Pamunkey River
Mobjack Bay
Cherrystone Inlet
Sinepuxent Bay
Virginia coastal bays

Chesapeake Bay mainstem

water quality - single stratum
benthic/sediment quality - separate strata for Maryland and Virginia waters

Delaware Bay

Chincoteague Bay

Albemarle-Pamlico estuarine system open-water area

Table A-2. Combinations of Strata Used to Produce Estimates for Geographic Areas.

Delaware Estuary
small estuarine systems Delaware River Delaware Bay
Chesapeake Bay
mainstem (one stratum for water quality, two strata for benthic/sediment quality) tributaries subsystems small estuarine systems
Potomac River (water quality)
Lower Potomac River Upper Potomac River
Coastal Bays
small estuarine systems Chincoteague Bay
Albemarle-Pamlico Estuarine system
open-water area (large system class) tidal rivers small estuarine systems
Tributaries
mainstem portion of tributary small estuarine systems that are part of tributary

Table A-3. Design Specifics for MAIA Intensively-Sampled Small Estuarine Systems.

design	randomized tessellation stratified (RTS) design uniform hexagonal grid overlain on each system one random site selected within each grid
inclusion probabilities	equal within each system, but different values for each system
estimation procedure	continuous resource equations with uniform inclusion probabilities for each system
information needs	area for each system hex grid overlay specifics for each system

Table A-4. Design Specifics for MAIA Randomly-Selected Small Estuarine Systems.

design	from list frame, random selection of systems to sample one random site within each system
inclusion probabilities	equal to area of each small estuarine system
estimation procedure	discrete resource equations
information needs	list of all small estuarine systems area of each small estuarine system total area of all small estuarine systems

Table A-5. Design Specifics for Delaware Estuary (River and Bay).

design	multiple-density, randomized tessellation stratified (MD-RTS) design (Stevens, 1997) define strata within estuary for sampling purposes uniform hexagonal grid is overlain on each strata, three random site selections are made from each hexagon with sampling at the first site accessible and meeting criteria
inclusion probabilities	equal inclusion probabilities within each strata with weighting based upon which site in order of selection actually sampled (1 for first site, 2/3 for second site, and 1/3 for third site) each stratum has inclusion probability associated with area of stratum
estimation procedure	continuous resource equations with unequal inclusion probabilities
information needs	area for each stratum strata that are within Delaware estuary hex grid overlay specifics for each strata selection order for each site actually sampled area for Delaware estuary

Table A-6. Design Specifics for Chesapeake Bay Mainstem.

design	basically treated the CBP sites as random sites, one group for water quality and one group for benthic/sediment quality
inclusion probabilities	for benthic/sediment quality, total area/number of sites for water quality, total area/number of sites
estimation procedure	continuous resource equations
information needs	hex grid overlay specifics total area for mainstem used for grid overlay (large system class for Chesapeake Bay as determined for Virginian Province effort) inclusion probabilities

Table A-7. Design Specifics for Tributaries and Subsystems.

design	basically treated the Chesapeake Bay program sites as random sites within each tributary or subsystem, one group for water quality and one group for sediment quality; can treat Potomac River as 2 strata (upper and lower) or could treat as one with different inclusion probabilities; Neuse River was overlain with three different grid densities
inclusion probabilities	area of hexagonal grid for each tributary or area of each tributary divided by number of sites
estimation procedure	continuous resource equations
information needs	total area for each tributary used for grid overlay hex grid overlay specifics for each tributary inclusion probabilities

Table A-8. Design Specifics for Albemarle-Pamlico Estuarine System.

design	list frame for small estuarine systems randomized tessellation stratified (RTS) design for tidal rivers and large systems (open water)
inclusion probabilities	small systems - area of each small estuarine system tidal rivers/large systems - area of hexagon
estimation procedure	discrete resource equations for small systems continuous resource equations for large systems and tidal rivers
information needs	inclusion probabilities hex grid overlay specifics area of systems and classes