Appendix I

Recommendations for Monitoring Program Design for Mid-Atlantic Estuaries

Background

The impetus of environmental management is ultimately for a sustainable environment for humans and other living organisms. The specifics of how one characterizes this sustainable environment depend on the perspective of the individual. For example, it would be different for economic concerns, recreational use, or aesthetic pleasure. It is generally accepted that observations of the environment are necessary to determine how well things are doing and if they are getting better or worse. An observation program to provide this information does not easily follow. There is a theoretical foundation in the statistical literature for the design of observation programs (Cochran 1977; Gilbert 1987; Cressie and Aldworth 1994; Stehman and Overton 1994). However, the actual design that would be employed to acquire this information depends on the specific question(s) that is (are) to be addressed with the observations.

Monitoring is a term frequently used to describe the acquisition of data in the environment. The use of the term monitoring can be put into at least the following categories (Olsen, personal communications): compliance monitoring, baseline monitoring, trend monitoring, implementation monitoring, effectiveness monitoring, project monitoring, and validation monitoring. Each of these categories requires the acquisition of data in a specific fashion to address the types of questions/issues/concerns. The reality is one monitoring design does not fit all categories. The monitoring design must be specifically matched to the information needs or the result of the effort may be unnecessarily disappointing.

Probability sampling is an efficient approach for monitoring when it is necessary to collect information at a finite number of sites and use that information to extrapolate beyond the sampled sites. A probability design is one in which every element in the population has a known (non-zero) probability of being selected from the population. There are four different types of probability sampling (simple random sample, systematic sample, stratified sample, and cluster sample), each with different characteristics. Stevens, 1997, discusses additional specific implementations of probability sampling. The important point is that one needs to be more specific than just saying a probability sampling design was used.

Large spatial applications of probability survey designs were implemented in estuaries of the northeastern and Gulf of Mexico areas of the U.S. in the early 1990s by USEPA's EMAP (Holland 1990; Paul, et al. 1992; Summers, et al. 1993; Summers, et al. 1995). The foundation of the EMAP approach to monitoring a condition is probability sampling that provides the basis for estimating resource extent and condition, for characterizing trends, and for representing spatial pattern, all with known levels of confidence. Therefore, in terms of the monitoring categories listed above, the MAIA-E design may be described as including baseline monitoring, trend monitoring, and effectiveness monitoring. These early designs were variants of the RTS implementation of probability sampling (Stevens, 1997). One drawback to the actual implementations of these designs was the inability to incorporate existing non-probability monitoring sites into the overall design.

As discussed in this report, the implementation of the MAIA-E program in 1997-98 was done in conjunction with existing benthic and water quality monitoring programs in the Chesapeake Bay. The

basic RTS design was used for the program. The U.S. EPA's National Coastal Assessment (NCA) is a current EMAP geographic initiative. This five-year program (2000-2004) focuses on surveying the condition of the Nation's coastal resources (estuaries and offshore waters) through an integrated, comprehensive coastal monitoring program among the coastal states to assess coastal ecological condition. The approach for NCA focuses on a strategic partnership with all 24 U.S. coastal states. Using compatible probability designs and a common set of survey indicators, each state is surveying and assessing the condition of their coastal resources independently. These data can then be used to develop statements of condition at multiple scales. Table I-1 summarizes the evolution and advancements in sampling designs which have resulted from the studies. The next section discusses specifics of the sampling design as implemented in the northeastern states for NCA. This is followed by the monitoring program design recommended for the mid-Atlantic estuarine waters.

Design Implemented in National Coastal Assessment in the Northeastern U.S.

The probability sampling design used in NCA for the northeastern U.S. has evolved from earlier designs to provide increased flexibility in the actual implementation. The basic design is as follows:

- Define the statistical population as estuarine waters delineated by GIS from NOAA charts.
- The grid is randomly placed (origin and orientation) using the random placement selected for EMAP-VP program.
- Tessellate grid to get approximate number of desired grid cells over target population.
- Randomly select point for each grid cell. Point is restricted to lie in the target population within the grid cell.

This last step in the design (as opposed to the random point anywhere within the grid, whether on land or water) permits possible incorporation of existing monitoring program sites into the design, increasing flexibility in implementing the program.

Before existing monitoring program sites were evaluated for incorporation into the probability design, data collected from existing sites must be confirmed to meet the quality assurance protocols specified by the program. This is to ensure that the statistical inferences made from these data are not compromised. Once this has been done, the evaluation of the site selection can proceed.

If existing monitoring program sites were selected using a probability design, then they can be incorporated directly into the design. For example, some state fish trawl programs use a stratified random design for site selection, with stratification usually based upon depth and habitat. Cox and Piegorsch (1996) discuss procedures for combining samples collected with different probability designs. However, a comparison needs to be made of the target population of the stratified random design with the target population for the program. If the existing program does not include all of the target population, then the existing program sites would need to be supplemented with additional sample sites over the remainder of the target population.

For existing monitoring program sites that were not selected using a probability design, the process to determine if the sites can be incorporated into the design is based upon the two concepts identified in Overton, et al. (1993):

1. The sites are identified with a subset of the population, and
2. The sites are similar to a probability sample of the same subset of the population.

These concepts were converted into criteria that were used for evaluating whether or not the existing monitoring sites could be considered for incorporation into the design. The criteria were:

1. The sites must have been selected initially to be representative of the area from which they were selected. For example, sites that do not satisfy this criterion would include those targeted for an outfall discharge location, the end of a dock, or a bridge overpass (for convenience in acquiring samples).

2. The distribution of individual variables from the existing monitoring sites must be equivalent to the distribution for probability samples from the same subset of the statistical population. For example, cumulative distribution functions (CDFs) of bottom dissolved oxygen concentration can be compared. A CDF displays the estimated portion of the population above and below any specified value of the variable. This criterion requires that data from probability samples be available for the population subset of interest.

These two criteria are required to be met before the existing sites are determined to be acceptable for incorporation into the design. Because of the limited availability of existing probability data for multiple parameters, a third criterion was considered as confirmatory, but was not required.

3. The correlation structure between variables from the existing monitoring sites must be equivalent to the correlation structure for probability samples from the same subset of the statistical population. This criterion requires that information on multiple variables from the probability sample be available for the population subset of interest.

Recommendation for Monitoring Program Design for Mid-Atlantic Estuarine Waters

The recommendation for the sampling design for mid-Atlantic estuarine waters is to incorporate, where possible, existing ongoing monitoring program sites using the procedure developed for NCA as discussed above and to supplement with new sites where existing sites are absent. If multiple existing sites are within a grid cell, then one of the sites would be randomly selected. As a first step in implementing this recommendation, it is necessary to determine which of the existing program sites do meet the criteria for possible incorporation. One would not have a rationale of incorporating the sites if they did not meet the criteria. It is understood that some assumptions may need to be made for judging the sites against the criteria. These assumptions need to be documented as part of the actual design that would be implemented.

With this recommended design, an hierarchical spatial sampling design can be implemented, where data collected at the various scales can be put together for an overall estimate of condition. The hierarchical design uses the basic procedure outlined above but employs differing densities of grid, depending on the allocation of sites (and associated uncertainty) for the different areas. For example, a grid of 50 cells could be overlain on the mainstem of Chesapeake Bay, with grids of densities, say 10, overlain on systems such as the Severn and South Rivers. Estimates of condition could be made on these small systems with inherent uncertainty associated with sample size of 10. These data would be combined with the remaining sites in the Bay to assess the overall condition of the Bay.

Table I-1. Progression of EMAP Probability Sampling Designs in the Northeast.

Program	Survey Design	Incorporation of Existing Program Sites
EMAP-Virginian Province 1990-93	**sampling classes (strata) based on physical extent** **separate design for each class** large systems - tessellated, center point chosen small systems - list frame, random start for 1st year, all systems over 4 yrs tidal - rib & spine	**none possible**
MAIA Estuaries 1997-1998	**RTS (random tessellated stratified) for each strata** large - tessellated, random point w/in cell small - list frame **equal weights within each strata**	**incorporate other sites for entire strata**
National Coastal Assessment 2000-2004	**RTS with variable weights for each strata** **sites restricted to estuarine resource**	**criteria for possible use of existing sites** **strata may contain mix of existing and survey sites**