Appendix B

Methods and Indicators for MAIA Estuaries 1997-98

Samples and *in situ* measurements were collected for characterization of: (1) physical habitat (depth, temperature, salinity, dissolved oxygen, pH, water clarity, organic carbon content in sediments, and grain size of sediments); (2) water quality (dissolved and particulate nutrients, total suspended solids, chlorophyll *a*, and phaeophytin); (3) contamination in sediments (total metals, simultaneously-extracted metals, acid volatile sulfide, PAHs, PCBs, pesticides, butyltins, and sediment toxicity); (4) contaminants in fish and crab tissue (total metals, PAHs, PCBs, and pesticides); and (5) biotic condition (diversity and abundance of benthic invertebrates, fish and shellfish, and external pathology and spleen macrophage aggregates in fish).

Sampling was carried out by the host agency responsible for the station network in the region, i.e., NOAA in the Delaware Estuary, NOAA and the University of North Carolina in the APES; NPS in the Chincoteague Bay; CBP in the Chesapeake Bay; and the U.S. EPA in the Maryland and Virginian coastal bays and in the intensively-sampled estuaries. Sites in the Delaware Inland Bays were not sampled in the MAIA estuaries program because they were recently included in an earlier EMAP assessment program (Chaillou, et al. 1996). Samples for analyses of water, sediment, and benthic community quality were collected predominantly in 1997. The U.S. EPA and the NPS conducted the fish trawls in 1998. Generally, samples for all analyses were collected at each station, except in the Chesapeake Bay, where separate station networks are maintained by the CBP for water, sediment, and benthic analyses; and in the APES, where water analyses are not routinely performed.

A Hydrolab Datasonde was used to measure *in situ* values of physical habitat parameters at meter intervals, and water clarity was determined with a Secchi disk. Water samples were collected with a 5-L Go-Flo® bottle in the surface and bottom layers (one meter from the air and sediment interface, respectively), and filtered with 0.7-micron glass-fiber filters. The water and filters were frozen for later analysis. A wide range of nutrient parameters was measured (Table 3-3), including dissolved inorganic and organic components of nitrogen and phosphorus nutrients, as well as particulate forms. Total nitrogen concentrations are calculated as the sum of total dissolved nitrogen and particulate organic nitrogen. Total phosphorus concentrations are calculated as the sum of total dissolved phosphorus and particulate phosphorus. Replicate field samples were analyzed at 6% of the stations to evaluate the repeatability of the sampling procedure. Analytical methods are described in D'Elia, et al. (1997).

Sediments were collected with a 0.04-m^2 Young-modified Van Veen grab sampler. Surface sediments (composites of upper 2 cm) were collected from each station and used to measure physical, chemical, and toxicological characteristics of the sediments. The chemical contaminants are those measured by the NOAA NS&T program (Valette-Silver 1992). Sediments analyzed for total metals were dried and completely digested in nitric/hydrofluoric acids (acid persulfate for mercury). For measurement of acid volatile sulfide (AVS) and simultaneously extracted metals (SEM), wet sediments were treated with 1M HCl to release the AVS (USEPA 1991). For the organic analyses, sediments were extracted using the procedures of NOAA NS&T program (Lauenstein, et al. 1993). All analyses were performed on samples that were stored frozen.

Three measures of sediment toxicity were measured: (1) Static, 10-day survival tests for *Ampelisca abdita* (ASTM 1991, USEPA 1995) were performed at 20° C and 30 ppt salinity on sediments from 401 stations. Total ammonia concentrations were measured and the sediment flushed if levels were greater than 20 mg/L. Sediments were considered to be toxic if the survival rate of the amphipods relative to a corresponding negative control was < 80% at $\alpha = 0.05$. (2) A solid-phase Microtox® bioassay was performed on whole sediments (Microbics 1992) at 326 stations. A dilution series of sediment in 2% saline (0.01 to 10% sediment) was inoculated with the photoluminescent bacterium *Vibrio fischerii*, incubated for 20 minutes, filtered, and analyzed for remaining luminescence with a Microtox® Model 500 Analyzer. A log-linear regression model was used to determine the EC_{50} (the sediment concentration that reduced luminescence by 50% relative to a non-toxic control), corrected for moisture content of the sediment. Muddy sediments can diminish luminescence (Ringwood, et al. 1997); therefore, toxicity criteria depended on grain size: sediments were considered to be toxic when EC_{50} was $\leq$ 0.2% for silt fractions $\geq$ 20%, and when EC_{50} was $\leq$ 0.5% for silt fractions $\leq$ 20%. (3) An exploratory organic-extract Microtox® bioassay (Johnson and Long 1998) was performed on sediments from 75 stations. The sediments were dried with anhydrous sodium sulfate, extracted with dichloromethane, and exchanged into a mixture of dimethylsulfoxide, toluene, and isopropyl alcohol. A dilution series was prepared from the extract and analyzed as above. Sediments were considered toxic if the EC_{50} exceeded the EC_{50} for a reference sediment.

Sediment was also collected separately with a 0.04-m^2 Young-modified Van Veen grab sampler for the purpose of measuring species composition, enumeration, and biomass determination of infaunal and epifaunal benthic macroinvertebrates. One to three grab samples were taken from each station. The contents were live-sieved in the field with a 0.5 mm mesh screen, and organisms retained on the screen were fixed in a 10% buffered formalin with rose bengal for preservation and visualization. Only organisms larger than 0.5 mm were processed; therefore, groups such as turbellarian flatworms, nematodes, ostracods, harpacticoid copepods and foraminifera were excluded from the identification process. Taxa were identified to the lowest possible taxon, usually species; however, because of complexities involved with precise identification, the following groups of organisms were routinely identified to the indicated taxonomic level: anthozoa (class), chironomidae (family), hirudinea (class), nemertinea (phylum), oligochaeta (class), ostracoda (subclass), sipuncula (phylum), turbellaria (class), and copepoda (order). Biomass was calculated as the dry weight of all specimens of a taxon in a grab sample, following dehydration at 60° C and combustion in an ash oven at 500° C for 5 hours. These data were used to compute mean abundances per grab of infaunal species, epifaunal species, spionid polychaetes, and tubificid oligochaetes; the mean biomass per grab of all species; the total and mean numbers per grab of infaunal species and epifaunal species; the Shannon-Weiner Index (Shannon and Weaver 1949) and Gleason's D index (Krebs 1989), which are measures of species diversity; and a multi-metric benthic community index, developed for species in the Virginian Province, computed using expressions of species diversity and abundance of opportunistic species (Paul, et al. 1999).

A benthic community index is calculated as a weighted combination of three benthic diversity metrics: a salinity-adjusted Gleason's index, a salinity-adjusted abundance of tubificids, and the abundance of spionids. This benthic index was developed with data compiled during the 1990-1993 EMAP effort in the Virginian Province using discriminant analysis to determine a weighted combination of parameters which distinguish impacted and unimpacted sites in the EMAP-VP (Paul, et al. 1999). Indices values less than or equal to zero designate impacted conditions by definition. Stations in the APES are evaluated by a similar

index derived for the EMAP-Carolinian Province estuarine study (Hyland, et al. 1996). The threshold of impairment for this index is ≤ 3; therefore, the index scores are adjusted by subtracting 3 from the APES values and are evaluated using an impairment threshold of zero.

Fish and crab were collected in standard 10-minute trawls, using a 15 m otter trawl towed against the tide at 1-3 knots. The identity, abundance, average fork length, and frequency and location of visible pathologies (lumps, growths, ulcers, or finrot) were determined in ship-board inspections for all fish species collected in standardized fish trawls. The spleens of three target species – white perch (*Morone americana*), spot (*Leiostomus xanthurus*), and summer flounder (*Paralichthys dentatus*) – were preserved for later histological examination for macrophage aggregates, and the concentrations of metals, PAHs, PCBs, and pesticides were measured in composites samples of summer flounder or blue crab collected in standard or auxiliary trawls.

Appendix C

Criteria for Presenting Indicator Data

Surface Water Nutrients – Total Nitrogen (TN). TN was not measured directly in the MAIA-E program; rather, it was calculated as the sum of total dissolved nitrogen (TDN) and particulate organic nitrogen (PON). The units of TN are mgN/L, which is equivalent to ppm. There are no firm guidelines for classifying the nitrogen nutrient condition in estuaries; therefore, we used the 25th and 75th percentile values of all MAIA-E measurements to define the three map categories. *Low* TN concentrations are less than or equal to 0.5 mgN/L; *intermediate* concentrations are greater than 0.5 to 1.0 mgN/L; and *high* concentrations are greater than 1.0 mgN/L. Some monitoring programs measure DIN rather than TN. DIN is the sum of nitrate, nitrite, and ammonium concentrations. Plants assimilate these simple inorganic compounds directly; therefore, some would argue the DIN is a better measure of nutrient condition. However, DIN concentrations vary widely throughout the year in response to complex cycles of supply and transformation. Surface waters are often nearly depleted of DIN during summer when the MAIA-E program sampled the estuaries. Since TN is a combination of both dissolved inorganic and organic, plus particulate organic nutrient compounds, it is a more accurate measure of the overall availability of nutrients in an estuary. TN will likely be the measure used by regulatory agencies in the future to set nutrient concentration guidelines for estuaries.

Caveats regarding the interpretation. We emphasize that the map categories "low", "intermediate" and "high" are based on a simple ranking of the MAIA-E data rather than on established guidelines. The maps accurately show *relative* TN conditions in the region, but not based on any particular environmental response such as bloom activity or the survival of sea grasses. The maps indicate TN concentrations during a single, summertime sampling. It is possible that other nutrient measures are better predictors of bloom activity, e.g., nutrient availability during spring or regeneration rates of nutrients in sediments. In part, the MAIA-E program was designed to address such questions regarding interpretation. Finally, high concentrations of nitrogen nutrients are not necessarily harmful. For instance, highly turbid waters may prevent excessive plant growth even when abundant nutrients are available. This is the case in much of the Delaware Estuary and coastal bays.

Surface Water Nutrients – Total Phosphorus (TP). TP was not measured directly in the MAIA-E program; rather, it was calculated as the sum of total dissolved phosphorus (TDP) and particulate organic phosphorus. It is a measure of both organic and inorganic dissolved phosphorus species. There are no firm guidelines for classifying the phosphorus nutrient condition in estuaries; therefore, we used the 25th and 75th percentile values of all MAIA-E measurements to define the three map categories. *Low* TP concentrations are less than or equal to 0.05 mgP/L; *intermediate* concentrations are greater than 0.05 to 0.10 mgP/L; and *high* concentrations are greater than 0.10 mgP/L. As with the TN measurements described above, caution is advised when interpreting the TP maps. The TP map categories are based on simple ranking of the MAIA phosphorus measurements and may not reflect actual environmental responses. The maps are accurate displays of the *relative* availability of TP during summer in the region.

Surface Layer Chlorophyll *a*. Seawater samples were collected from about 1 meter below the surface with a 5L Go-Flo® sampling bottle. Water samples were filtered aboard ship with 0.7-micron glass-fiber filter pads. Chlorophyll *a* pigments were extracted from the filter with 90% acetone and measured with a Turner Design TD700 Fluorometer. The results were reported with units of µg/L, which is equivalent

to parts per billion. Three categories were used to characterize chlorophyll concentrations: *good* for concentrations less than or equal to 15 µg/L; *fair* for values greater than 15 to 30 µg/L; and *poor* for concentrations greater than 30 µg/L. These thresholds were used in preparing a report on the condition of the MAIA estuaries (Paul, et al. 2000). The threshold value of 15 µg/L is also equal to the restoration goals recommended for the survival of SAV in the Chesapeake Bay (Batiuk, et al. 2000). While this threshold value may not be appropriate for SAV restoration in all estuaries, it is a useful reference value.

Sediment Organic Content (%TOC). Total organic carbon contents were labeled as follows: *low* for values less than or equal to 1%; as *intermediate* if greater than 1% to 3%; and as *high* for values greater than 3%. The category thresholds are those used in the EMAP-VP study, based on findings in that study that TOC values greater than 3% were associated with impacted benthic communities (as measured by the benthic index), while values less than 1% were not (Paul, et al. 1999).

Water Clarity (Secchi Depth). The MAIA-E program characterizes water clarity by reporting the Secchi depth (SD), which is the depth (in meters) at which a white disk becomes obscured by suspended material or colored tannins present in the water. Shorter Secchi depths signify more murky water. We use three categories to describe water clarity: *clear* waters display Secchi depths greater than 1.0 meter; *intermediate* clarity designates Secchi depths greater than 0.3 to 1 meter; and *murky* conditions are indicated by Secchi depths less than or equal to 0.3 meter. There are no established criteria for water clarity; therefore, we use threshold values preciously adopted by EMAP-VP.

Relation to other measures of water clarity. Secchi depths can be compared with the light extinction coefficient K_d , another parameter used to characterize water clarity. K_d describes the exponential decrease of illumination I_z / I_0 with water depth z: $I_z / I_0 = \exp(-K_d * z)$. Although there is no firm relationship between K_d and SD, Batiuk, et al. (2000) found an expression which is useful for estuaries: $SD * K_d = 1.45$. Thus, an SD of 1.0 meter is equivalent to the transmission of 23% of ambient light at one meter depth, comparable to the restoration goals recommended for the survival of SAV in the Chesapeake Bay (Batiuk, et al. 2000). While this threshold value may not be appropriate for SAV restoration in all estuaries, it is a useful reference value for water clarity nonetheless.

Caveats regarding interpretation. The interpretation of water clarity measurements may sometimes be, pardon the expression, unclear. This is so for at least three reasons. First, measurements of Secchi depth alone (or light extinction coefficient) cannot distinguish whether the light attenuation is caused by particulate matter, by living or dead plant material, or by colored substances in the water. Such identification may be aided by consideration of other parameters such as total suspended solids and chlorophyll. Most estuarine waters are naturally turbid to some extent, especially in regions where rivers expel their loads of suspended material and nutrients into a protected water body. However, Secchi depth measurements alone cannot distinguish between natural and anthropogenic causes of loss of water clarity. Finally, there is as yet no consensus among researchers regarding criteria for adequate vs. harmful degrees of water clarity.

Bottom Layer Dissolved Oxygen (DO). Condition categories were considered *good* for levels greater than 5 mg/L; *fair* for concentrations greater than 2 to 5 mg/L; and *poor* for concentrations less than or equal to 2 mg/L. EPA's proposed saltwater quality criteria cite DO thresholds of 2.3 and 4.8 mg/L (USEPA 2000b). Most states have set their water quality standard for DO at 5 mg/L.

Caveat: Short-term hypoxic conditions may also occur in shallower water during the night, when oxygen demand from respiration exceeds supply. This indicator does not measure this effect.

Sediment Quality – Metal Contamination. This summary uses the approach of Long and Morgan (1990) to characterize contamination in sediments. These researchers identified a list of nine metals which induced impairment to biological organisms in estuaries. They specified ERL and ERM for each metal. The ERL value specifies the concentration of metal that would likely produce adverse effects in 10% of a population, while the ERM value indicates the concentration that would have an effect rate of 50%. The metals and respective ERL and ERM values (mg/g dry wt or ppm) are: arsenic: 8.2, 70; cadmium: 1.2, 9.6; chromium: 81, 370; copper: 34, 270; lead: 46.7, 218; mercury: 0.15, 0.71; nickel: 20.9, 51.6; silver: 1.0, 3.7; and zinc: 150, 410. These values were taken from Long, et al. (1995). Three categories are presented on the maps. *Good* signifies no ERL exceedances; *intermediate* represents any ERL (but no ERM) exceedance, and *poor* indicates an ERM exceedance.

Sediment Quality – Organic Contamination. The Long and Morgan (1990) approach is also used to characterize contamination levels of organic compounds in sediments (see description above). The following nineteen compounds and respective ERL and ERM values were used to prepare the maps in this summary report: acenaphthene: 16, 500; acenaphthylene: 44, 640; anthracene: 85, 1100; fluorene: 19, 540; 2-methyl naphthalene: 70, 670; naphthalene: 160, 2100; phenanthrene: 240, 1500; benz(a)anthracene: 261, 1600; benzo(a)pyrene: 430, 1600; chrysene: 384, 2800; dibenzo(a,h)anthracene: 63, 260; fluoranthene: 600, 5100; pyrene: 670, 2600; 4,4'-DDE: 2.2, 27; low MW PAH (sum of 2- and 3-ring PAHs): 550, 3160; high MW PAH (sum of 4- and 5-ring PAHs): 1700, 9600; total PAH (sum of all measured PAHs): 4000, 45000; total DDT (sum of 2,4' and 4,4' congeners of DDD, DDE and DDT): 1.6, 46; and total PCBs: 23, 180. The units are ng/g (ppb). These values were taken from Long, et al. (1995). Three categories are presented on the maps. *Good* signifies no ERL exceedances, *intermediate* represents any ERL (but no ERM) exceedance, and *poor* indicates any ERM exceedance.

Sediment Toxicity (Amphipod Survival). The toxicity of sediments was evaluated using a static ten-day assay conducted using the amphipod *Ampelisca abdita* following EMAP procedures (EPA 1994, 1995). The test is simple in concept – amphipods are added to sediment, and their survival rate is used as an indicator of sediment toxicity. Results are reported as the average number of amphipods surviving in the sample tests divided by the number of amphipods surviving in a control sediment, expressed as a percent. Lower values of this result indicate higher toxicity. The three categories used on the maps are: *good* when survival is greater than 80%; *fair* for values between 60 to 80%; and *poor* for survival rates less than or equal to 60%. The same threshold values were employed in the EMAP-VP program.

Sediment Quality – Benthic Index. The EMAP-VP benthic index is a combination of three metrics into a single index (the metrics are: salinity-adjusted Gleason's index, the salinity-adjusted abundance of tubificids, and the abundance of spionids). This benthic index was developed with data compiled during the 1990-1993 EMAP effort in the Virginian Province (Paul, et al. 1999). The majority of values range from -5 to +5, with positive values signifying healthy conditions and negative values indicating degraded conditions. On the maps in this summary, the following classifications hold: *good* when the index is greater than zero, and *poor* for values less than zero. The threshold value of zero reflects a defining feature of the index scale; values greater than zero were similar to pristine reference sites, while negative values were associated with impaired reference sites.

Number of Fish Species. The MAIA-E program conducted regular fish surveys during the summer of 1998 to characterize the structure and health of the fish communities. The stations sampled were selected according to the probabilistic design but were not identical with the stations sampled for water and sediment quality analyses conducted primarily in 1997. Therefore, it is not possible to directly compare these different analyses station by station. However, it is statistically valid to compare results among

classes of estuaries, e.g., large versus small estuaries, Delaware Estuary versus Chesapeake Bay, etc. Three categories were used to classify species richness: *low* if the number was two or less; *intermediate* for 3 to 5 species; and *high* for 6 or more species. These categories are the same as those used in the EMAP-VP.

Fish and Shellfish Tissue Contamination. Representative samples of edible tissue from summer flounder and blue crabs were analyzed for metallic and organic toxicants. The fish and shellfish were collected in 1998. Table 6-1 lists the analytes and concentration limits considered by USEPA to present risks to human consumers (USEPA 2000). The limits are based on human health risk assessment and are considered to be protective of recreational, tribal, ethnic, and subsistence fishers who are likely to consume more fish than the general population. A site was classified as poor if a limit for any analyte in Table 6-1 is exceeded.