EPA/620/R-95/007
September 1995

Quality Assurance Report
EMAP-Virginian Province
1990 - 1993

by

Charles J. Strobel
United States Environmental Protection Agency
Narragansett, RI

and

Raymond M. Valente
Science Applications International Corporation
Narragansett, RI

Darryl Keith
EMAP-Virginian Province Manager
United States Environmental Protection Agency
Narragansett, RI

Kevin Summers
EMAP-Estuaries Technical Director
United States Environmental Protection Agency
Gulf Breeze, FL

Joseph LiVolsi
Quality Assurance Officer
United States Environmental Protection Agency
Narragansett, RI

United States Environmental Protection Agency
National Health and Environmental Effects Research Laboratory
Atlantic Ecology Division
Narragansett, RI

ABSTRACT

This report documents the results of Quality Assurance activities conducted in conjunction with sampling performed by EPA's Environmental Monitoring and Assessment Program's Estuaries study (EMAP-Estuaries) in the Virginian Province from 1990 through 1993. As part of the planning stage for each years activities, a QA Plan was developed. All sampling and analytical activities were required to be conducted in accordance with the prescribed methods, and following the standards stated in the QA Plan. This report discusses the results of Quality Assurance activities by indicator, data qualifier flags, data quality, and, where appropriate, discusses lessons learned and proposes changes or solutions to improve data quality.

Data collected in the Virginian Province from 1990 to 1993 were generally of high quality. A total of 446 Base Sampling Sites were scheduled for sampling over this period. Twenty one stations were eliminated due to inadequate water depth or logistical concerns. With the exception of total suspended solids (samples for this indicator were not collected in 1990), the success rate for all indicators exceeded 80% (percent of stations with data passing QC), with most exceeding 85%.

Some significant problems were encountered in the chemical analysis of sediment samples resulting in the deletion of some data from the database. The specific problems, and a discussion of the data deleted or qualified are included in this report.

DISCLAIMER

Mention of trade names or commercial products does not constitute endorsement by the Environmental Protection Agency or recommendation for use.

PREFACE

Contractor support for the preparation of this document was supplied via contract number 68-C1-0005 to Science Applications International Corporation.

The appropriate citation for this report is:

Strobel, C.J. and R.M. Valente. 1995. Quality Assurance Report: EMAP-Virginian Province, 1990-1993. United States Environmental Protection Agency, National Health and Environmental Effects Research Laboratory, Atlantic Ecology Division, Narragansett, RI. September 1995. EPA/620/R-95/007.

This report is AED contribution # 1639.

CONTENTS

Contents (continued)

TABLES

Tables (continued)

Tables (continued)

ABBREVIATIONS

AED	Atlantic Ecology Division of NHEERL (formerly ERL-N)
AVS	Acid Volatile Sulfide
BSS	Base Sampling Site
CDF	Cumulative Distribution Function
CTD	Conductivity, Temperature, Depth datalogger
DBT	Dibutyltin
DO	Dissolved Oxygen
dry wt	Dry weight
DS3	Hydrolab DataSonde3 datalogger
EMAP	Environmental Monitoring and Assessment Program
EMAP-E	EMAP-Estuaries
ERL-N	Environmental Research Laboratory, Narragansett (renamed AED)
MBT	Monobutyltin
mg/L	milligrams per liter = parts per million (ppm)
mg/kg	milligrams per kilogram = parts per million (ppm)
kg/m^3	kilograms per cubic meter
NHEERL	National Health and Environmental Effects Research Laboratory (U.S. EPA)
ND	Not Detected
ng/g	nanograms per gram = parts per billion (ppb)
PAH	Polycyclic Aromatic Hydrocarbon
PCB	Polychlorinated Biphenyl
QA	Quality Assurance
QC	Quality Control
SEM	Simultaneously Extracted Metals
SQC	Sediment Quality Criteria
TBT	Tributyltin
μg/g	micrograms per gram = parts per million (ppm)
μ	Micron
‰	parts per thousand (ppt)

Section 1
Introduction

The Estuaries component of EPA's Environmental Monitoring and Assessment Program (EMAP-E)commenced in 1990 with a Demonstration Project in the estuaries of the Virginian Biogeographic Province (mid-Atlantic coast from Cape Cod, Massachusetts to Cape Henry, Virginia). Following the successful completion of this Demonstration Project, EMAP-E monitoring in the Virginian Province (EMAP-VP) has continued on an annual basis through 1993. Complete descriptions of the EMAP-E monitoring approach and rationale, sampling design, indicator strategy, logistics, and data assessment plan are provided in the Near Coastal Program Plan for 1990: Estuaries (Holland 1990).

The EPA mandatory Quality Assurance (QA) Program requires that every environmental monitoring and measurement project have a written and approved quality assurance project plan (QAPP). As such, a QAPP was prepared for the 1990 Virginian Province Demonstration Project (Valente *et al.* 1990), and this plan has since been revised in each subsequent year of monitoring in the Province (Valente and Schoenherr 1991; Valente *et al.* 1992; Valente and Strobel 1993). The QAPP prepared each year describes the quality assurance and quality control activities and measures that are implemented in the Province to ensure that the data meet certain established criteria (*i.e.*, measurement quality objectives).

The purpose of this report is to present and interpret the results of the various quality assurance activities and quality control checks which have been performed over the first four years of monitoring in response to the requirements of the Virginian Province QAPPs. As the various QA results are presented and discussed, an attempt is made throughout the report to describe changes and "lessons learned" as the EMAP-VP QA Program has evolved over the past four years. In addition to this document, Quality Assurance Annual Report and Work Plans (QAARWPs) have been prepared for each year of EMAP-E monitoring since 1990 (Valente 1991a; Valente 1991b; Latimer 1992; Summers 1993).

All field work was conducted by Science Applications International Corporation (SAIC), Versar Inc., or a consortium of universities under the leadership of the University of Rhode Island (URI).

A table summarizing the percent of all data collected passing QC (*i.e.*, data completeness) is presented in Section 11.

Section 2
Field Crew Training and Audits

Monitoring for the EMAP-VP Program consists of intensive annual sampling by multiple field crews operating from small boats during a two month summer index period. EMAP-E has developed Standard Operating Procedures (SOPs) for its field activities to insure the comparability of data collected by different teams operating across wide geographic distances (*i.e.*, both within and among provinces). EMAP-VP has instituted an annual cycle involving rigorous field crew training and subsequent field performance reviews to insure uniform adherence to Standard Operating Procedures.

Training sessions lasting from four to eight weeks typically occur immediately prior to the summer sampling interval. Training involves a combination of both formal classroom instruction and "hands-on" practical experience to impart necessary skills in everything from first-aid and seamanship to sample shipping and computer use. As an essential aspect of the QA program, all field crews must pass a final proficiency exam (*i.e.*, "certification") at the end of the training session before they are permitted to begin actual sampling. In addition, at least once during the sampling interval, a formal field QA audit is conducted to ascertain that SOPs continue to be followed. In addition to the audit conducted by the QA Officer, a performance review of each crew is performed by senior Program personnel. Written examinations and results of performance reviews are maintained as permanent record by the Program.

The training certification exam and the subsequent field performance reviews typically are conducted by the Province QA Coordinator. Formal procedures, involving checklists and grading systems, have been developed to facilitate the certification/auditing process. Whenever deficiencies are noted, the field personnel are re-trained immediately prior to resuming sampling activities.

Records documenting the results of the annual field crew certifications are maintained by the Province QA Coordinator. In addition, following each field review, the QA Coordinator files a written report describing his/her findings and any corrective actions undertaken. QA personnel also have performed periodic on-site evaluations of laboratories responsible for processing samples. The purpose of these evaluations is to document that each contract laboratory has adequate equipment, personnel and facilities to analyze samples in accordance with prescribed methods and QA requirements. Laboratory evaluation results also are documented in reports filed by the Province QA Coordinator.

2.1 1990 Results

Formal training was held at the University of Rhode Island's (URI) Fisheries Center in Wickford, RI from May 29 to June 15, 1990. All crew members were required to attend the entire course (however, crew chiefs were periodically pulled from training for other activities). The development and conduct of the course was sub-contracted by SAIC to the URI Marine Advisory Service and Fisheries Department. Instructors for the course were provided by the URI Graduate School of Oceanography, URI Fisheries Department, NOAA (Milford, CT; Narragansett, RI; and Woods Hole, MA), American Heart Association, Computer Sciences Corporation (CSC), and SAIC.

The class was generally divided into two groups; one classroom and one practical (on-the-water). Most classroom lessons were followed by practical training. Topics included boating safety, trailering, operation of sampling equipment, navigation (including operation of the electronic instruments), data transfer, Quality Assurance/Quality Control, fish and mollusc taxonomy, fish pathology, and CPR.

All participants were required to complete a Skills Evaluation Form on the first day of training. This information was used to assign personnel to crews based on their skills, thereby assuring that each crew possessed the necessary skill mix (computer and electronic instrument operation, fish taxonomy, bivalve taxonomy, sediment sampling, etc.) for all aspects of sampling; and to help select those who would undergo additional training in a specialty area (*e.g.*, computer operation). Throughout training crews worked together as a team during all hands-on activities. At the end of training the composition of the crews was reviewed to assure that each crew had appropriate personnel to complete all aspects of sampling. Based on the personal knowledge of crew members gained by the Crew Chiefs, and information from the contract personnel managers, no changes to the crews were deemed necessary.

Trial runs, encompassing all components of sampling activities, were originally planned to be an important component of training. This was not fully realized during formal training; no training in integrated sample processing, packaging and shipping was provided. In addition, an evaluation of the crews by experienced EMAP-VP personnel at the end of training revealed that some data collection methods were still not well understood or being followed properly. As a result, it was decided at the end of the formal training period that the crews were not adequately prepared for the Data Collection Phase. Therefore, the start of Interval 1 was delayed by 10 days. This reduction left insufficient time for all stations to be sampled in that interval; therefore, the focus of Interval 1 activities was changed from the collection of data to an extension of training. This was deemed necessary to ensure crews were fully competent in all aspects of sample collection.

This extended training consisted of the actual collection of data and samples in the field at a limited number of stations under the close supervision of senior EMAP-VP personnel familiar with the methods. This activity served as more than just "dry runs", with some of the data collected during this exercise being used in the characterization of the Virginian Province.

During field operations each crew was visited by a senior EMAP staff member (Field Coordinator or QA Coordinator). All aspects of sampling, from boat operations to shipping, were observed by the reviewer. Some of the activities included confirming the presence/ absence of external pathologies, re-measuring fish and apparent RPD (redox potential discontinuity) depth, assuring that all precautions were taken to avoid contamination of the chemistry samples, assuring proper processing of benthic infauna samples, observing data entry, and assuring that all necessary safety precautions were observed. In 1990, no "field review check-off sheet" was utilized in this review; however, a memo to the Province Manager was generated summarizing the review. Both reviewers concluded that the crews demonstrated positive attitudes to QA issues, and that all sources of field-generated error were in reasonable control.

Evaluation of Training and Lessons Learned

An evaluation of training, based on the overall results of the Data Collection Phase, indicated that the success of training was mixed. URI provided an excellent facility and staff. Their contribution was mainly geared towards boat operations and safety, areas in which they have extensive experience in providing classroom and hands-on training to marine-related groups, such as commercial fishermen. The success can be measured by the absence of any injuries during over 13,000 person-hours of field operations. Extramural instructors for fish and mollusc taxonomy, and fish pathology provided excellent instruction; however, they were not expert in the goals of EMAP, and had limited time to present their material. The material they presented was often too broad in scope, resulting in inadequate instruction in the detailed areas pertinent to the Demonstration Project. It was suggested that, in future years, such instruction should be more focused on Virginian Provinces issues, species, and conditions. In-house instructors adequately presented instructions for the operation of gear; however, the science behind the methods was not explained (*e.g.*, the characteristics of a good dissolved oxygen profile). Several areas were identified that required more attention in subsequent courses, including packaging and shipping, and general maintenance (lubricating trailer hubs, etc.).

It was suggested that crew chiefs should be much more involved in training in future years. The proposal for 1991 included extensive training for all crew chiefs prior to crew training. This training should provide them with sufficient information to perform all sampling tasks. Important components should include the operation of the field computer, understanding all Quality Assurance issues, any theory necessary for them to evaluate whether or not sample or data collection must be repeated, and trouble shooting electronic sampling equipment. Crew chiefs would then play an active role in training their crews.

2.2 1991 Results

Suggestions for improving training (detailed above) were incorporated into planning activities for the 1991 season. Crew chiefs underwent detailed training during the first two weeks of June, 1991. Training was limited to two weeks because all but one of six crew chiefs were returnees from the previous year. Training was conducted at the U.S. EPA Environmental Research Laboratory-Narragansett, RI (ERL-N) and focused mainly on the sampling methods, with emphasis placed on the electronic measurements and the computer system. Crew chief training was conducted by SAIC and CSC personnel with oversight by EPA ERL-N staff.

Crew training was held from 17 June to 19 July 1991. Both safety and sampling methods were important components of training. Crew training was broken into two phases: formal training which lasted for approximately 2½ weeks, and one week (per crew) of trial runs.

Trial runs consisted of four days in the field during which crews operated as they would during the sampling season. They were assigned four stations to monitor for all parameters, including DataSonde deployment and retrieval. Crews members stayed in motels, prepared samples for shipment, entered data into the field computer, and electronically transmitted all data to the Field Operations Center (FOC) just as they would during actual field operations. In addition, the Field Coordinator or the QA Coordinator visited each crew during trial runs, completing a performance review sheet to determine the crew's overall grasp of the Program. All crews were deemed properly prepared to begin sampling activities on 22 July, 1991.

Certification examinations for crew chiefs and field crew members were administered at the end of each course and proved to be very useful. As a result of testing, two crew chiefs were identified as needing additional training. Remedial coaching was provided and they were fully competent by the start of crew training. The examination administered at the end of crew training suggested some areas, such as contingencies for moving stations, were not adequately covered, so additional time was spent discussing these topics prior to trial runs.

In addition to the crew certification visits performed during dry runs, each crew was visited by a senior EMAP staff member (Field Coordinator or QA Coordinator) during field operations. All aspects of sampling, from boat operations to shipping, were observed by the reviewer. Some of the activities included confirming the presence/absence of external pathologies, re-measuring fish and apparent RPD depth, assuring that all precautions were taken to avoid contamination of the chemistry samples, assuring proper processing of benthic infauna samples, observing data entry, and assuring that all necessary safety precautions were observed. The reviewer used a "field review check-off sheet" to provide guidance during the review, and to document the crew's performance. Both reviewers concluded that the crews were sufficiently concerned with all QA issues, and that all sources of field-generated error were in reasonable control.

The only problem noted was the determination of the depth of the apparent RPD. This measurement was determined to be too subjective, variable, and difficult to accurately measure based on a visual inspection of a clear plexiglass core taken from a grab sample. Although reasonable measurements could be made in muddy sands, the majority of the sediments encountered by field crews were fine grained muds where adhesion to the plexiglass core creates too much smearing to allow for an accurate measurement. As a result of this observation,

RPD measurements were dropped from the sampling program, and all existing RPD data deleted from the database.

2.3 1992 Results

Crew chiefs, who were all returnees from previous years, underwent a refresher training course during the last week of May, 1992. This training was conducted at ERL-N and focused mainly on the sampling methods, with emphasis placed on the electronic measurements and the computer system. Crew chief training was conducted by SAIC and CSC personnel with oversight by EPA ERL-N staff.

Crew training was held from 15 June to 17 July 1992. Both safety and sampling methods were important components of training. Crew training was broken into two phases: formal training which lasted for approximately 3 weeks, and one week (per crew) of trial runs.

Trial runs consisted of five days in the field during which crews operated as they would during the sampling season, monitoring practice stations for all parameters. Crew members stayed in motels, prepared samples for shipment, entered data into the field computer, and electronically transmitted all data to the Field Operations Center (FOC) just as they would during actual field operations. In addition, the Field Coordinator or the QA Coordinator visited each crew during trial runs, completing a performance review sheet to determine the crew's readiness. All crews were deemed properly prepared to begin sampling activities on 27 July, 1992.

In addition to the crew certification visits performed during trial runs, each crew was visited by a senior EMAP staff member (Field Coordinator or QA Coordinator) during field operations. All aspects of sampling, from boat operations to shipping, were observed by the reviewer. Some of the activities included confirming the presence/absence of external pathologies, re-measuring fish, assuring that all precautions were taken to avoid contamination of the chemistry samples, assuring proper processing of benthic infauna samples, observing data entry, and assuring that all necessary safety precautions were observed. The reviewer used a "field review check-off sheet" to provide guidance during the review, and to document the crew's performance. Both reviewers concluded that the crews were sufficiently concerned with all QA issues, and that all sources of field-generated error were in reasonable control.

The EMAP-VP QA Officer participated in audits during both crew certification and field operations. During these audits he evaluated both the crew and the QA Coordinator's ability to conduct a performance review. His findings were then summarized in a memo to the ERL-N laboratory director and the Province Manager. Although he disagreed with some of the methods employed, he was fully satisfied that crews were adhering to EMAP SOPs, and that the QA Coordinator was competent at evaluating the remaining crews.

2.4 1993 Results

Crew chief training for 1993 was separated into pilot training and chief scientist training. Pilot training was held at the University of Rhode Island's Graduate School of Oceanography from 17 May to 21 May, 1993. This training consisted of instruction in navigation, safety and boat handling. Chief scientist training was conducted at URI from 14 June to 18 June, 1993. This training focused mainly on the sampling methods, with emphasis placed on the electronic measurements and the computer system. Crew chief training was conducted by SAIC, URI and ROW Sciences personnel, with oversight by EPA ERL-N staff. All chief scientists were returnees from previous years.

Formal crew training was held at URI from 21 June to 9 July, 1993. Both safety and sampling methods were important components of training. Crew training was followed by one week (per crew) of trial runs.

Trial runs consisted of five days in the field during which crews operated as they would during the sampling season, monitoring practice stations for all parameters. Crew members stayed in motels, prepared samples for shipment, entered data into the field computer, and electronically transmitted all data to the Field Operations Center (FOC) just as they would during actual field operations. In addition, the Field Coordinator or the QA Coordinator visited each crew during trial runs, completing a performance review sheet to determine the crew's readiness. All crews were deemed properly prepared to begin sampling activities on 26 July, 1993.

In addition to the crew certification visits performed during trial runs, each crew was visited by a senior EMAP staff member (Field Coordinator or QA Coordinator) during field operations. All aspects of sampling, from boat operations to shipping, were observed by the reviewer. Some of the activities included confirming the presence/absence of external pathologies, re-measuring fish, assuring that all precautions were taken to avoid contamination of the chemistry samples, assuring proper processing of benthic infauna samples, observing data entry, and assuring that all necessary safety precautions were observed. The reviewer used a "field review check-off sheet" to provide guidance during the review, and to document the crew's performance. Both reviewers concluded that the crews were sufficiently concerned with all QA issues, and that all sources of field-generated error were in reasonable control.

The EMAP-VP QA Officer participated in audits during both crew certification and field operations. During these audits he evaluated both the crew and the QA Coordinator's ability to conduct a performance review. His findings were then summarized in a memo to the ERL-N laboratory director and the Province Manager. Although he disagreed with some of the methods employed, he was fully satisfied that crews were adhering to EMAP SOPs, and that the QA Coordinator was competent at evaluating the remaining crews.

Section 3
QA Results for Chemical Contaminant Analyses of Sediments

3.1 Background

Measurement Quality Objectives (MQOs) for the analysis of chemical contaminants in EMAP-E sediment samples are specified in the annual Province Quality Assurance Project Plans. These plans variously require each EMAP-E laboratory to analyze the following types of quality control (QC) samples along with every batch or "set" of field chemistry samples: laboratory reagent blanks, calibration check standards, laboratory fortified sample matrix (matrix spike), laboratory fortified sample matrix duplicate (matrix spike duplicate), laboratory duplicate, and Laboratory Control Material (LCM). Results for these QC samples must fall within certain pre-established control limits for the analysis of a batch of samples to be considered acceptable.

Standard or Certified Reference Materials (SRMs or CRMs) typically are used by EMAP-E laboratories as their Laboratory Control Material (LCM). SRMs and CRMs have known or "certified" concentrations of the analytes being measured and therefore are useful for assessing both accuracy and precision. The QA Project Plan requires the laboratory's percent recovery (relative to the certified concentration in the reference material) to fall within certain pre-established control limits to be considered acceptable. If the laboratory consistently fails to meet these acceptability criteria for the CRM or SRM analysis, the values reported for the failed analytes are considered to be suspect (biased) and are flagged in the database, as described in the following section.

In addition to the above QA requirements, each laboratory analyzing EMAP sediment chemistry samples must participate in an intercomparison exercise conducted through NOAA's National Status and Trends (NS&T) Program and coordinated by the National Institute of Standards and Technology (NIST).

Many of the goals and objectives of the EMAP-E program coincide with those of NOAA's NS&T program. By interagency agreement, personnel from the two agencies have continued to coordinate their activities to ensure that data produced by the two coastal monitoring programs are compatible. This applies in particular to measurements of chemical contaminant concentrations in tissue and sediment samples. To achieve this goal, all EMAP-E laboratories participated yearly in the NIST/NOAA intercomparison exercises. A brief description of this exercise follows.

The NIST/NOAA intercomparison exercise is a key element of the National Status and Trends Program and the EMAP-E "performance-based" QA philosophy. In this continuing series of exercises, various materials are distributed in common to all laboratories for blind analysis. These exercises are coordinated for EPA and NOAA by NIST, which typically distributes a variety of materials including gravimetrically-prepared solutions, extracts of environmental samples (tissue or sediment), or actual marine samples (tissue or sediment). All EMAP-E laboratories are required to participate in the NOAA/NIST intercomparison exercises in order to become "certified" prior to analyzing actual samples, and as a means of assessing comparability on an on-going basis.

Each year the EMAP-E QA Coordinator joined laboratory personnel from the EPA's Environmental Monitoring Systems Laboratory (EMSL), Cincinnati, OH, in attending the NIST/NOAA intercomparison exercise annual meeting, where the results of the intercomparison exercises were presented and discussed. The annual meetings serve as an excellent forum for representatives of the various labs to identify common analytical problems and discuss potential solutions.

3.2 Data Qualifier Codes for Chemistry

Four data qualifier codes or "flags" are used in EMAP-E's sediment chemistry datasets:

The "SC-A" code indicates that an analyte was not detected. When the "SC-A" code is used, the concentration field is left blank and the detection limit for the analyte in that particular sample is reported under the variable "MDL" (method detection limit).

It is sometimes possible for a laboratory to detect an analyte and report its concentration at a level which is below the calculated method detection limit for the sample. In these situations, the analyst is confident that the analyte was present in the sample, but there is a high degree of uncertainty in the reported concentration. The "SC-B" code is used to flag reported values which are below the calculated method detection limit for the sample. Such values are considered estimates only and should be used with discretion.

The "SC-C" code is applied in situations where the laboratory failed to meet required control limits for one or more of the quality control samples analyzed along with each sample batch. In such situations, there is reason to believe that the concentrations reported for an analyte or group of analytes may not accurately reflect the actual concentrations present in the samples. The "SC-C" code usually is applied when the Certified Reference Material results indicate that a laboratory experienced a consistent bias in the analysis of a particular analyte or group of analytes. The "SC-C" code is also applied whenever other QC sample results suggest a possible bias in the reported values (*e.g.*, sample contamination detected in the laboratory reagent blank). Values flagged with the "SC-C" code therefore are considered estimates only and should be used with discretion.

Results of QC sample analyses are stored in the EMAP-E database and are available upon request. The "SC-C" code used to flag suspect values is applied following a thorough QA review of the entire data package submitted by the laboratory for a given year. In many instances, best professional judgement must be used to decide which values should be qualified as estimates only. In the following sections, explanations are provided for the "SC-C" codes which appear in the EMAP-E sediment chemistry datasets. Persons using these data may wish to perform their own review of the QC sample results to determine the acceptability of these data for their purposes.

For the years 1991-1993 in the Virginian Province, the laboratory used gas chromatography/electronic capture detection (GC/ECD) with dual column confirmation for the analysis of PCB congeners and chlorinated pesticides in sediments. All values reported in the database for the PCBs and pesticides represent "confirmed" results (*i.e.*, the analyte was detected and could be quantified on both the primary and secondary columns). In situations where an analyte was detected on one column, but was not confirmed on the second column, the result was treated as a "not detect" *(i.e.*, the SC-A code is used to flag the result in the database).

Close inspection of the "confirmed" results for certain pesticides revealed a number of instances where there was a significant discrepancy in the amount detected on the two GC/ECD columns (*i.e.*, greater than a factor of three difference). In these instances, it is difficult to ascertain which amount is more accurate (*i.e.*, which is the "right" answer). A decision was made to take a "conservative" approach and report the lower of the two values in the database, and to flag these values using the "SC-D" code. The SC-D code has the following meaning: "Analyses were conducted using GC/ECD with dual column confirmation. Quantitation on the two columns differed by more than a factor of three, and the lower of the two results is reported."

Although this approach was deemed necessary, the user must be cautioned that the application of the "SC-D" code may invalidate investigations of the ratios of compounds. For example, if the concentrations of p,p'-DDT from the two columns were 6.1 and 2.0 ng/g respectively, the SC-D code would be applied and the lower value of 2.0 ng/g reported. However, if the values for p,p'-DDE were 6.0 and 2.1 ng/g, the SC-D code would **NOT** be applied and the original value of 6.0 ng/g would be reported. Most likely the ratio of these two compounds

is approximately 1, but the results as reported would indicate a ratio of about 3. Therefore, ratios of compounds should only be used when either all or none of the compounds are flagged with the SC-D code.

Values which are not flagged with the SC-B, SC-C or SC-D codes are considered valid and useful for anticipated assessment purposes.

3.3 Quality Assessment Results

In the following sections, results for chemistry QC samples are summarized, and data flags associated with the 1990 to 1993 EMAP-VP chemistry datasets are explained.

3.3.1 Laboratory Audit

A technical systems audit was conducted on 29 and 30 April 1991 at EMSL in Cincinnati, OH. The audit team was led by Mr. Raymond Valente, the QAO for the EMAP-E program. Mr. Valente was assisted by two senior organic chemists from ERL-N: Dr. Richard Pruell and Mr. Don Cobb. Mr. Robert Graves, the acting EMAP QA Coordinator based at EMSL-Cincinnati, accompanied the audit team as an observer.

Major problems were uncovered in EMSL-Analytical's (the production laboratory arm of EMSL-Cincinnati) adherence to QA protocols for the analysis of PCB and pesticides in sediments. These problems were uncovered prior to and during the audit of the laboratory (see following paragraph). As a result of the audit, the 1990 sediment organic analyses were halted and a series of corrective actions were implemented to bring the process back into control. Analyses resumed in late FY '91.

The primary purpose of the audit was to review the methodology being employed at EMSL-Analytical for analysis of low-level organic compounds in estuarine sediment samples from the EMAP-E 1990 Virginian Province Demonstration Project. This review was deemed necessary partly in response to delays in sample processing and subsequent phone conversations with laboratory personnel which suggested technical difficulties with the organic analyses had been encountered. The audit had to be scheduled with a minimum of advance notice (ca. 1 week) to include one of the principal EMSL-Analytical participants prior to her departure from the laboratory on April 30. The audit team's specific goal was to determine the exact nature of any technical difficulties being encountered and provide constructive assistance as appropriate. At the same time, the audit team evaluated the adequacy of EMSL-Analytical's adherence to QA requirements in relation to the EMAP-E analyses.

The main deficiency noted was failure to adhere to QA specifications in performing sediment organic analyses (PCBs and pesticides). The audit findings were documented in a report submitted to the EMAP-E Acting Technical Director and appropriate EMSL-Analytical personnel. A series of corrective actions were implemented over the course of the spring and summer 1991, and no further audits were conducted.

3.3.2 1990 QA Results

Major and trace element analyses (except mercury)

Two methodologies, inductively-coupled plasma atomic emission spectrophotometry (ICP-AES) and graphite furnace atomic absorption (GFAA) spectrophotometry, were utilized for the analyses of major and trace elements (metals) in sediment samples collected by EMAP. The results of QC samples (e.g., calibration standards, laboratory reagent blanks, matrix spikes, and LCMs) run with each of the 18 batches of 1990 VP sediment samples generally met the pre-established EMAP criteria for acceptability.

For the ICP-AES analyses, which included the metals Ag, Al, Cr, Cu, Fe, Mn, Ni, Pb, and Zn, a total of 18 analytical sets or "batches" of samples were analyzed. SRM 2704 (Buffalo River Sediment, issued by NIST) was analyzed along with every batch as the Laboratory Control Material. The analysis of a LCM is a particularly important component of EMAP's performance-based approach to QA/QC that provides assessments of accuracy as well as precision. The 1990 QAPP required the laboratory's percent recovery (relative to the certified concentration in the reference material) to fall within a range of 85% to 115% for each metal. Except for silver, the average percent recovery of each metal (relative to the certified concentration in SRM 2704) was within the acceptability range of 85% to 115% (Table 3-1), and no "SC-C" codes were applied.

Table 3-1. Summary results for SRM 2704 (Buffalo River Sediment) used as a set control for the 1990 Virginian Province sediment inorganic analyses.

ICP-AES METALS (n = 18 analysis sets or "batches"):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Ag	na	na	na	na	na
Al	96	1.8	1.9	92	99
Cr	87	2.7	3.1	80	91
Cu	95	2.4	2.5	90	99
Fe	88	1.6	1.8	83	90
Mn	96	2.2	2.3	92	99
Ni	90	5.5	6.2	84	110
Pb	93	4.5	4.8	85	99
Zn	96	1.6	1.7	93	99

GFAA METALS (n = 18 analysis sets):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
As	78	4.1	5.3	70	89
Cd	100	7.0	7.0	87	111
Sb	79	11.9	15.1	51	99
Se	97	12.4	12.8	70	119
Sn	80	30.0	37.5	29	144

[1] Average percent recovery relative to the SRM certified value.
[2] Standard deviation of the percent recovery values.
[3] Coefficient of variation of the percent recovery values.
[4] Minimum percent recovery for 18 analysis sets
[5] Maximum percent recovery for 18 analysis sets

Silver was not detected in most of the 1990 Virginian Province samples; however, the laboratory's detection limit of 1 ppm was well above the target detection limit of 0.01 ppm specified in the QA Plan. If the target detection

limit had been achieved, silver probably would have been detected and quantified in a much higher number of samples. Therefore, the 1990 results are not reliable for assessing silver concentrations in Virginian Province sediments. This problem was corrected in 1991 by analyzing for silver using GFAA rather than ICP.

The GFAA analyses included the metals As, Cd, Sb, Se, and Sn; a total of 18 analytical sets or "batches" of samples were analyzed. SRM 2704 was analyzed along with every sample batch as the Laboratory Control Material. Average SRM percent recoveries fell outside the acceptability range of 85% to 115% for the following metals: As (78%), Sb (79%) and Sn (80%) (Table 3-1). In addition, matrix spike recoveries for these metals were highly variable. These low and variable recoveries are attributed to both the low concentrations of these metals in SRM 2704 (*i.e.*, close to the detection limit) and the less rigorous digestion procedure used (*i.e.*, hydrofluoric acid was not employed). Therefore, data users are cautioned that the reported concentrations for As, Sb, and Sn may underestimate the true amount present in each sample, but this bias is not considered severe given that the recoveries of these metals from SRM 2704 ranged between 78% and 80%. Given this slight low bias in the SRM results, all reported concentrations for As, Sb and Sn in the 1990 Virginian Province dataset are qualified with the "SC-C" code.

It should be noted that in 1991 the control limits were changed from 85-115% recovery to 80-120% recovery. The only 1990 data this would potentially affect is Sn, which showed a recovery rate of 80%. Sn results for 1990 remain flagged in the EMAP database.

Mercury analyses

For the 1990 Virginian Province sediment mercury analyses, the Certified Reference Material BEST-1 (issued by the National Research Council of Canada) was analyzed along with every sample batch as the Laboratory Control Material (n = 18 sample batches). The average percent recovery of 82% for mercury in this reference material fell just outside the accuracy control limit range of 85% to 115%, suggesting that mercury may have been slightly under-recovered in some sample batches. However, an average percent recovery of 96% was achieved for the matrix spike samples analyzed in each batch. Overall, these results indicate acceptable accuracy for the mercury analyses, and no "SC-C" codes were used to qualify the data. The 1990 mercury results were deemed acceptable for use without qualification.

Organic analyses

Data users are cautioned that there are several major deficiencies in the 1990 Virginian Province sediment organics dataset that might limit or preclude the use of these data. These deficiencies, described below, were the result of numerous methodological and QA/QC problems experienced by the laboratory responsible for the analyses.

As stated earlier, all EMAP-E analytical laboratories were required to participate in interlaboratory comparison exercises as ongoing demonstrations of capability. EMSL-Analytical took part in these exercises in 1990. The 1990 Round 1 exercise required the laboratory to identify and quantify a mixture of PAHs, PCBs and pesticides (six each) in hexane. EMSL-Analytical failed to identify seven of the 18 compounds and had relatively high variability between sample replicates. The second round requirements were similar to the first and EMSL-Analytical improved their performance by lowering the variability between samples. In the final round the results showed difficulties with all three groups (PAHs, PCBs, and pesticides) having varying degrees of error in each, especially the pesticides with 4,4'-DDT and dieldrin having exceedingly high values for mean absolute percent error (2,142 and 3,113 percent, respectively).

In general, results for reagent blanks and calibration check samples analyzed with each batch of field samples fell within control limits and serve to verify that sample contamination did not occur and that all instruments were calibrated properly throughout the analytical runs. However, the matrix spike results are of limited use in assessing overall data quality because the spiking solutions used by the laboratory for the PAH and PCB/pesticide analyses contained only a small subset of the analytes of interest and not the full suite as originally specified in the QA Plan. Furthermore, it is difficult to evaluate laboratory performance solely on the basis of matrix spike results because it is often equivocal whether low recoveries are due to flawed methodology, poor technique, or a true matrix interference.

Results for laboratory duplicate samples, intended to serve as a check on precision, also are of limited value in assessing the quality of the 1990 Virginian Province organics data because the laboratory usually failed to detect the analytes of interest in the sample chosen at random for duplicate analysis (*i.e.*, most of the analytes in laboratory duplicate samples were reported as "not detected").

Given the above limitations on using the matrix spike and laboratory duplicate results to assess the overall quality of the 1990 Virginian Province organics data, great emphasis was placed on the LCM results. For both the PAH and PCB/pesticide analyses, SRM 1941 (Organics in Marine Sediment, issued by NIST) was analyzed as the LCM along with each batch of samples. The QAPP required the laboratory's percent recovery (relative to the certified concentration in the reference material) to fall between 70% and 130% for each organic analyte.

For most of the individual PAH compounds and PCB congeners with "known" concentrations in SRM 1941, the average percent recovery achieved by the laboratory (based on n=20 batches for PAHs and n=22 batches for PCB/pesticides) consistently fell within the control limit range of 70% to 130% (Tables 3-2 and 3-3). Very high and variable SRM 1941 recovery rates were experienced for the pesticides heptachlor epoxide (231%), cis-chlordane (322%), trans-nonachlor (412%), and 4,4'-DDT (186%) (Table 3-3). In general, significant problems were experienced by EMSL in their analyses of samples for PCBs and pesticides. In addition to their poor performance for pesticides in the intercomparison exercise, problems existed with their extraction and analysis of samples.

Due to problems with integrating the internal standard peak consistently between standards and samples, EMSL-Analytical switched to an external standard quantitation. The EMAP QA Team was concerned about this since the external calibration does not allow any accounting for errors introduced by different extract volumes and injection volumes. The EMAP QA Team also felt that the internal standard chosen was not the best choice since its chemical structure and properties are not the same as the analytes of interest. EMSL-Analytical was strongly urged to begin using PCB 198 as the internal standard for quantification. This would eliminate the external standard usage and allow analytes to be quantified directly from a similar compound.

Tetrachloro-m-xylene (TCMX) was used as an internal standard as well in the 1990 analyses. In most cases, peak area was used to quantify results; however, at times peak height was used due to interferences. The peak height from the internal standard was used along with the ratio of the peak area/height from the 5 ppb standard to calculate the peak areas in samples with interferences. It is not clear whether or not the ratio is constant with increasing or decreasing concentrations. TCMX was recovered in excess of 200% in some samples and greater than 100% in others.

EMSL-Analytical used two GC columns (RTX-50 and OV-5) in order to quantify and confirm PCBs and pesticides. It was noted in a review of the 1990 raw data that use of a particular column for quantifying results was dependent upon the TCMX recovery. It has become clear that EMSL-Analytical chose to report the best result from each column. The EMAP Audit Team allowed the use of both columns to quantify results of SRMs as long as the use is consistent. Weeks later it was found that EMSL-Analytical was still picking and choosing the analytes instead of consistently measuring them on the same column. Given the possible differences in results (>3X) this practice is questionable at best.

These analytical problems result in significant doubts regarding the quality of the 1990 PCB/pesticide data. These findings, along with EMSL's acknowledgement of the problems, have resulted in the deletion of all 1990 PCB and pesticide results from the EMAP database.

A major deficiency in the 1990 Virginian Province organics dataset is related to the laboratory's failure to achieve the target detection limits originally specified in the QA Plan. These target detection limits were 10 ng/g (dry weight) for each PAH compound and 0.25 ng/g for each PCB congener and pesticide. In general, the detection limits achieved by the laboratory ranged from 1.5 to 30 times higher than the target value for PAH compounds and up to 15 times higher than the target value for PCB congeners and pesticides (Table 3-4). In addition, the detection limits varied widely because the laboratory analyzed a different amount (*i.e.*, dry weight) of sediment from each sample. As a result, the analytes of interest were not detected in a large number of samples, and the "calculated" detection limit (*i.e.*, the theoretical concentration of each analyte necessary for detection) differed significantly from sample to sample (Table 3-4).

If the target detection limits had been achieved and consistent sample sizes had been used, the organic analytes of interest probably would have been detected and quantified in most of the 1990 Virginian Province samples. In reality, analytes of interest present in the samples at low concentrations were not detected and therefore not reported. This limits the comparability of the 1990 Virginian Province organics data with other data sets for which lower detection limits were achieved and limits data users' ability to make quantitative evaluations of sediment contamination for these organic compounds in the Virginian Province. As a result of this problem, EMSL's poor performance in the intercomparison exercise, and the results presented in Table 3-2, the "SC-C" code has been applied to all 1990 PAH data. This QA code informs the user that problems were identified which question the quality of the results. Therefore these results should be treated as estimates and used with caution.

Table 3-2. Results for SRM 1941 (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1990 Virginian Province sediment PAH analyses (n = 20 analysis sets or "batches").

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
Phenanthrene	98.8	22.0	22.3	62	138
Anthracene	71.6	17.9	25.0	37	101
Fluoranthene	99.2	22.4	22.6	65	149
Pyrene	87.6	18.7	21.3	65	121
Benz[a]anthracene	93.9	20.8	22.1	57	141
Benzo[b+k]fluoranthene	104.6	18.9	18.1	67	142
Benzo[a]pyrene	64.9	15.4	23.7	40	90
Perylene	64.4	16.2	25.2	35	93
Benzo[ghi]perylene	86.2	23.3	27.0	48	145
Indeno[1,2,3-cd]pyrene	118.9	29.5	24.8	65	182

[1] SRM 1941 has certified concentrations for only a subset of the PAH compounds analyzed by the laboratory in 1990.
[2] Average percent recovery relative to the SRM certified value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 20 analysis sets
[6] Maximum percent recovery for 20 analysis sets

Table 3-3. Results for SRM 1941 (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1990 Virginian Province sediment PCB/pesticide analyses (n = 22 analysis sets or "batches").

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
PCB 18	79.4	17.1	21.5	23	101
PCB 28	54.8	9.2	16.8	34	76
PCB 52	101.5	23.5	23.1	60	146
PCB 66	67.7	9.7	14.3	47	80
PCB 101	73.9	17.1	23.1	48	105
PCB 118	99.2	14.4	14.5	65	116
PCB 153	94.5	15.1	16.0	60	121
PCB 105	96.3	17.9	18.6	67	130
PCB 138	77.1	16.3	21.1	53	105
PCB 187	82.7	18.6	22.5	58	122
PCB 180	97.0	19.5	20.1	66	132
PCB 170	82.3	20.5	24.9	57	143
PCB 195*	147.0	39.0	26.5	80	213
PCB 206*	100.3	27.9	27.8	61	176
PCB 209	93.9	21.5	23.0	61	134
Heptachlor epoxide*	231.0	91.7	39.7	109	448
cis-Chlordane*	322.0	81.5	25.3	87	450
trans-Nonachlor*	411.9	710.7	172.5	86	2770
4,4'-DDE	104.8	32.0	30.5	65	212
4,4'-DDD	92.3	21.4	23.2	33	123
4,4'-DDT*	185.8	135.4	72.9	63	660

[1] SRM 1941 only lists "non-certified" or informational values for this group of PCB congeners and pesticides (* = concentration in the SRM is less than 10 times the target detection limit).
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 22 analysis sets
[6] Maximum percent recovery for 22 analysis sets

Table 3-4. Range in detection limits (in ng/g dry weight) reported for organic compounds in 1990 Virginian Province sediment samples. The target detection limits were 10 ng/g for each PAH compound and 0.5 ng/g for each PCB congener and pesticide.

Polycyclic Aromatic Hydrocarbons (PAHs)

	Minimum	Maximum	Median
Acenaphthene	21	207	34
Anthracene	17	121	28
Benz(a)anthracene	17	72	28
Benzo(a)pyrene	23	151	38
Benzo(e)pyrene	23	153	37
Biphenyl	23	150	36
Chrysene	22	72	35
Dibenz(a,h)anthracene	24	252	43
2,6-dimethylnaphthalene	24	156	38
Fluoranthene	16	114	24
Fluorene	25	176	43
2-methylnaphthalene	25	162	39
1-methylnaphthalene	23	150	34
1-methylphenanthrene	13	86	21
Naphthalene	30	54	39
Perylene	27	189	46
Phenanthrene	16	44	26
Pyrene	15	39	22
Benzo(b+k)fluoranthene	22	145	33
Acenaphthlylene	22	212	38
Benzo(g,h,i)perylene	31	325	55
Ideno(1,2,3-c,d)pyrene	26	249	43
2,3,5-trimethylnaphthalene	23	219	38

DDT and its metabolites

	Minimum	Maximum	Median
2,4'-DDD	0.13	1.93	0.24
4,4'-DDD	0.12	6.10	0.20
2,4'-DDE	0.10	1.11	0.18
4,4'-DDE	0.04	0.45	0.07
2,4'-DDT	0.12	1.26	0.22
4,4'-DDT	0.18	3.22	0.58

continued

Table 3-4, continued.

Chlorinated pesticides other than DDT

	Minimum	Maximum	Median
Aldrin	0.10	1.78	0.27
Alpha-Chlordane	0.09	1.16	0.19
Trans-Nonachlor	0.04	0.87	0.07
Dieldrin	0.04	0.52	0.08
Heptachlor	0.10	1.47	0.19
Heptachlor epoxide	0.08	1.85	0.19
Hexachlorobenzene	0.03	7.23	0.09
Lindane (gamma-BHC)	0.16	27.5	0.64
Mirex	0.03	1.93	0.08

18 PCB Congeners:

	Minimum	Maximum	Median
PCB 08	0.08	4.46	0.63
PCB 18	0.37	5.89	0.94
PCB 28	0.08	1.03	0.17
PCB 44	0.06	1.50	0.17
PCB 52	0.11	2.70	0.38
PCB 66	0.09	1.01	0.18
PCB 101	0.12	1.39	0.20
PCB 105	0.07	0.60	0.14
PCB 118	0.06	0.65	0.12
PCB 128	0.12	1.62	0.23
PCB 138	0.11	1.31	0.18
PCB 153	0.11	1.03	0.19
PCB 170	0.09	2.15	0.32
PCB 180	0.11	1.30	0.19
PCB 187	0.08	0.72	0.13
PCB 195	0.10	1.23	0.19
PCB 206	0.10	1.38	0.20
PCB 209	0.12	1.09	0.20

Total Organic Carbon analyses

All QC results for the analysis of total organic carbon in the 1990 Virginian Province sediment samples fell within required control limits. The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM. The certified concentration of total carbon in this reference material is 3.69% (percent dry weight). The average percent recovery for TOC in PACS-1 achieved by the laboratory for n = 18 batches of samples (*i.e.*, 18 separate analyses of PACS-1) was 87.2%, with all values falling within the range 85% to 95%. Since the PACS-1 certified concentration includes both organic carbon and a very small fraction of inorganic carbon, the laboratory's percent recovery values for organic carbon are expected to be below 100%. Based on the good overall percent recovery of organic carbon in the Certified Reference Material, the 1990 Virginian Province sediment TOC data were deemed acceptable for use without qualification.

Butyltin analyses

Data users are cautioned that there are deficiencies in the 1990 Virginian Province butyltin analyses which might limit or preclude the use of these data. The main deficiency is related to the laboratory's failure to detect the butyltin compounds of interest (TBT, DBT, MBT) in the majority of samples analyzed. The method detection limits established by the laboratory were 4 ng/g dry weight (as tin) for both TBT and DBT, and 10 ng/g dry weight (as tin) for MBT. It is possible that the butyltin compounds of interest were present in many samples at concentrations below these detection limits, and, therefore, the occurrence of butyltin compounds in Virginian Province sediments may be more widespread than indicated by these data.

The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM for these analyses. Average percent recoveries relative to the certified value for n = 14 analysis sets were 73% for TBT, 57% for DBT and 394% for MBT. These recoveries fall outside the QA Plan-specified accuracy range of 85% to 115% and indicate that TBT and DBT were consistently under-recovered and MBT was grossly over-recovered in this reference material. Therefore, all values reported for TBT, DBT and MBT in samples where these compounds were detected are considered estimates (SC-C code) and should be used with discretion.

3.3.3 1991 QA Results

Major and trace element analyses (except mercury)

For the 1991 Virginian Province analysis of major and trace elements by ICP-AES and GFAA, the laboratory generally met the pre-established acceptability criteria (control limits) for the QC samples (*e.g.*, calibration check samples, laboratory reagent blanks, matrix spikes, and Laboratory Control Materials). For the ICP-AES analyses, which included the metals Al, Cr, Cu, Fe, Mn, Ni, Pb, and Zn, a total of 13 analytical sets or "batches" of samples were analyzed. The Certified Reference Material (CRM) "BCSS-1" (Estuarine Sediment, issued by the National Research Council of Canada) was analyzed along with every batch as the LCM. The 1991 QAPP required the laboratory's percent recovery (relative to the certified concentration in the reference material) to fall within a range of 80% to 120% for each metal. With the exception of Cr and Pb, the average percent recovery of each metal was within this acceptability range (Table 3-5). The average percent recovery for Cr was slightly lower than acceptable, and the average percent recovery for Pb was slightly higher than acceptable. These results suggest that Cr may have been consistently "under-recovered" and Pb may have been consistently "over-recovered" in the actual samples. Therefore, all reported values for these two metals were qualified with the SC-C code in the database.

The GFAA analyses included the metals Ag, As, Cd, Sb, Se, and Sn; a total of 19 analytical sets or "batches" of samples were analyzed. The CRM BCSS-1 also was analyzed along with every sample batch as the LCM. Average percent recoveries for all metals fell within the acceptability range of 80% to 120% (Table 3-5), and no results were flagged in the database. The CRM BCSS-1 does not have a "certified" value for silver, making it difficult to assess laboratory accuracy and precision for this metal. However, the laboratory was able to achieve a lower detection for this metal in 1991 compared to 1990, which resulted in silver being detected in a much higher number of samples in 1991 compared to 1990.

Table 3-5. Summary results for CRM BCSS-1 (Estuarine Sediment) used as a set control for the 1991 Virginian Province sediment inorganic analyses.

ICP-AES METALS (n = 13 analysis sets or "batches"):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Al	95	6.2	6.5	87	109
Cr	70	2.0	2.8	66	73
Cu	105	3.0	2.8	99	110
Fe	95	2.9	3.0	91	100
Mn	93	2.9	3.1	87	97
Ni	91	2.4	2.7	86	94
Pb	122	26.5	21.7	81	185
Zn	89	1.5	1.7	87	91

GFAA METALS (n = 19 analysis sets):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Ag	na	na	na	na	na
As	94	9.0	9.6	76	114
Cd	91	23.6	26.1	39	157
Sb	98	15.4	15.6	78	137
Se	111	32.5	29.3	50	189
Sn	111	14.9	13.4	66	135

[1] Average percent recovery relative to the SRM certified value.
[2] Standard deviation of the percent recovery values.
[3] Coefficient of variation of the percent recovery values.
[4] Minimum percent recovery for n analysis sets
[5] Maximum percent recovery for n analysis sets

Mercury analyses

For the 1991 Virginian Province mercury analyses, the Certified Reference Material BEST-1 (issued by the National Research Council of Canada) was analyzed along with every sample batch as the LCM (n = 9 sample batches). The average percent recovery of 92% for mercury in this reference material fell well within the acceptability range of 80% to 120%. In addition, an average percent recovery of 104% was achieved for the matrix spike samples analyzed in each batch. Overall, these results indicate acceptable accuracy for the mercury analyses, and no "SC-C" codes were used to qualify the data. The 1991 Virginian Province mercury results were deemed acceptable for use without qualification.

Organic analyses

In general, results for reagent blanks and calibration check samples analyzed with each batch of samples fell within control limits and serve to verify that sample contamination did not occur and that all instruments were calibrated properly throughout the analytical runs. Average recoveries of compounds in matrix spike samples generally fell within control limits, although these recoveries tended to be highly variable between different batches. This, in part, reflects the fact that the spiked samples were chosen at random and sometimes had high "background"

concentrations of the spiked analytes. In these cases it was difficult for the laboratory to accurately recover the spiked amount relative to the high background, resulting in zero percent recovery in some samples. Furthermore, it is difficult to evaluate laboratory performance solely on the basis of matrix spike results because it is often equivocal whether low recoveries are due to flawed methodology, poor technique, or a true matrix interference.

Given the above limitations on using the matrix spike results to assess the overall quality of the 1991 Virginian Province organics data, great emphasis was placed on the LCM results. For both the PAH and PCB/pesticide analyses, SRM 1941 (Organics in Marine Sediment, issued by NIST) was analyzed as the LCM along with each batch of field samples. For most of the individual PAH compounds and PCB congeners with "known" concentrations in SRM 1941 (this includes both "certified" and "non-certified" values), the average percent recovery achieved by the laboratory (based on n = 14 batches for PAHs and n = 15 batches for PCB/pesticides) generally fell within the control limit range of 70% to 130% (Tables 3-6 and 3-7). Whenever the laboratory failed to achieve these average recovery rates for a particular compound, all the results in the 1991 database for that compound were flagged with the "SC-C" code to indicate the potential inaccuracy inferred from the SRM analysis. It is important to note that the 70% to 130% recovery criteria only applies to compounds having SRM concentrations greater than 10 times the laboratory's detection limit. When compounds occur at concentrations less than about 10 times the detection limit, a greater amount of analytical uncertainty is expected and the normal control limit "acceptability" criteria do not apply.

Based on the above, the results for the following organic compounds were flagged with the "SC-C" code in the 1991 Virginian Province organics dataset: PCB 101, PCB 138, PCB 153, PCB 18, PCB 187, acenaphthylene, chrysene, 1-methylphenanthrene, and naphthalene. In addition, although the average percent recovery for ideno(1,2,3-c,d)pyrene was within limits (98%), all results for this compound were flagged with the SC-C code because the recoveries between batches exhibited relatively high variability (*i.e.*, 35% coefficient of variation). Although the average SRM percent recoveries for the compounds dieldrin, heptachlor epoxide and PCB 195 also were outside the acceptability range of 70% to 130% (Table 3-7), these compounds occur in the SRM at concentrations less than 10 times the laboratory's detection limit. Therefore, the acceptability criteria do not apply.

Unlike the 1990 analyses, when the laboratory failed to achieve a consistent detection limit for the organic compounds, in 1991 a consistent detection limit of 0.25 ng/g (dry weight) was achieved for each PCB congener and pesticide and 10.0 ng/g (dry weight) was achieved for each PAH compound.

As previously indicated (see Section 3.2), the laboratory used gas chromatography/electronic capture detection (GC/ECD) with dual column confirmation for the analysis of PCB congeners and chlorinated pesticides in the 1991 Virginian Province sediment samples. All values reported in the database for the PCBs and pesticides represent "confirmed" results (*i.e.*, the analyte was detected and could be quantified on both the primary and secondary columns). In general, for all reported PCB congeners except PCB 195, the rate of confirmation was between 95% and 100% (PCB 195 rate of confirmation was 87%). The rate of confirmation exceeded 90% for all the chlorinated pesticides except the following: heptachlor (59%), heptachlor epoxide (57%), mirex (82%), p,p DDT (65%), and o,p DDT (72%). Whenever an analyte was detected on one column, but was not confirmed on the second column, the result was treated as a "not detect" (*i.e.*, the SC-A code is used to flag the result in the database). Whenever there was a significant discrepancy in the amount detected on the two GC/ECD columns (*i.e.*, greater than a factor of three difference), the lower of the two values is reported in the database and flagged with the "SC-D" code. Please note the warning associated with this code discussed in Section 3.2.

EMSL did not participate in the NOAA intercomparison exercise in 1991. They did participate in 1992 and the results showed a continuing problem with pesticide analyses (see Section 3.3.4). Therefore, all pesticide data (dieldrin, heptachlor epoxide, cis-chlordane, trans-nonachlor, all DDT-series compounds) are qualified with the "SC-C" code to inform the user of potential problems with the data. Note that this qualification does not necessarily apply to PCBs as well.

Table 3-6. Results for SRM 1941 (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1991 Virginian Province sediment PAH analyses (n = 14 analysis sets or "batches").

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
Acenaphthene	111	23.2	20.9	67	137
Acenaphthlylene	41	10.6	25.9	27	61
Anthracene	95	26.4	27.7	59	142
Benz(a)anthracene	92	28.2	30.5	54	165
Benzo(a)pyrene	77	15.1	19.7	52	106
Benzo(e)pyrene	101	22.4	22.2	61	138
Benzo(b+k)fluoranthene	121	25.4	21.0	87	174
Benzo(g,h,i)perylene	105	21.1	20.1	64	141
Biphenyl	103	22.7	22.1	63	138
Chrysene	145	30.4	21.0	94	196
2,6-dimethylnaphthalene	113	24.2	21.3	70	145
Fluoranthene	93	20.2	21.7	64	134
Fluorene	105	32.3	30.7	62	179
Ideno(1,2,3-c,d)pyrene	98	34.1	34.7	21	150
1-methylnaphthalene	99	27.8	28.2	59	158
2-methylnaphthalene	109	33.6	30.8	53	158
1-methylphenanthrene	138	50.5	36.5	64	247
Naphthalene	69	27.3	39.5	8	126
Perylene	72	15.2	21.1	47	96
Phenanthrene	111	27.2	24.4	76	160
Pyrene	96	23.9	24.9	56	134

[1] Listed compounds include those having both "certified" and "non-certified" concentrations in SRM 1941.
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 14 analysis sets
[6] Maximum percent recovery for 14 analysis sets

Table 3-7. Results for SRM 1941 (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1991 Virginian Province sediment PCB/pesticide analyses (n = 15 analysis sets or "batches").

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
PCB 18	32	10.0	31.2	20	50
PCB 28	77	11.7	15.2	58	95
PCB 52	102	14.1	13.8	85	122
PCB 66	87	12.8	14.7	68	104
PCB 101	68	10.4	15.2	48	86
PCB 118	93	28.0	29.9	58	170
PCB 153	66	5.2	7.9	55	76
PCB 105	128	19.1	15.0	99	165
PCB 138	68	5.5	8.1	60	77
PCB 187	64	7.7	11.9	52	84
PCB 180	96	9.9	10.3	80	110
PCB 170	75	6.4	8.6	68	89
PCB 195*	142	29.8	20.8	108	199
PCB 206*	76	10.4	13.7	62	92
PCB 209	82	9.7	11.8	69	98
Dieldrin*	143	29.3	20.6	85	182
Heptachlor epoxide*	139	27.3	19.6	99	184
cis-Chlordane*	96	12.7	13.3	71	122
trans-Nonachlor*	89	15.3	17.2	72	127
4,4'-DDE	91	9.5	10.5	75	109
4,4'-DDD	80	9.0	11.2	64	98
4,4'-DDT*	102	22.6	22.2	62	128

[1] SRM 1941 only lists "non-certified" or informational values for this group of PCB congeners and pesticides (* = concentration in the SRM is less than 10 times the target detection limit).
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 22 analysis sets
[6] Maximum percent recovery for 22 analysis sets

Total Organic Carbon analyses

All QC results for the analysis of total organic carbon in the 1991 Virginian Province sediment samples fell within required control limits. The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM. The certified concentration of total carbon in thisreference material is 3.69% (percent dry weight). The average percent recovery achieved by thelaboratory for n = 11 batches of TOC samples (*i.e.*, 11 separate analyses of CRM PACS-1) was 94.0%, with all values falling within the range 88% to 99%. Since the PACS-1 certified concentration includes both organic carbon and a very small fraction of inorganic carbon, the laboratory's percent recovery values for organic carbon are expected to be below 100%. Based on the good overall percent recovery of organic carbon in the Certified Reference Material, the 1991 Virginian Province sediment TOC data were deemed acceptable for use without qualification.

Butyltin analyses

Data users are cautioned that there are deficiencies in the 1991 Virginian Province butyltin analyses which might limit or preclude the use of these data. The main deficiency is related to the laboratory's failure to detect the butyltin compounds of interest (TBT, DBT, MBT) in the majority of samples analyzed. The MDLs established by the laboratory were 5 ng/g dry weight (as tin) for both TBT and DBT, and 12 ng/g dry weight (as tin) for MBT. It is possible that the butyltin compounds of interest were present in many samples at concentrations below these detection limits, and, therefore, the occurrence of butyltin compounds in Virginian Province sediments may be more widespread than indicated by these data.

The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the Laboratory Control Material for these analyses. Average percent recoveries relative to the certified value for n = 12 analysis sets were 79% for TBT, 89% for DBT and 115% for MBT. The percent recovery value for TBT falls slightly outside the acceptable accuracy limits of 80% to 120% and indicates that TBT may have been consistently under-recovered in this reference material. Therefore, all values reported for TBT in samples where this compound was detected are considered estimates (SC-C code) and should be used with discretion.

3.3.4 1992 QA Results

Major and trace element analyses (except mercury)

For the 1992 Virginian Province analysis of major and trace elements by ICP-AES and GFAA, the laboratory generally met the pre-established acceptability criteria (control limits) for the QC samples (*e.g.*, calibration check samples, laboratory reagent blanks, matrix spikes, and Laboratory Control Materials). For the ICP-AES analyses, which included the metals Al, Cr, Cu, Fe, Mn, Ni, Pb, and Zn, a total of 16 analytical sets or "batches" of samples were analyzed. The Certified Reference Material (CRM) "BCSS-1" (Estuarine Sediment, issued by the National Research Council of Canada) was analyzed along with every batch as the LCM. With the exception of Cr, the average percent recovery of each metal (relative to the certified concentration in BCSS-1) was within the QAPP-specified acceptability range of 80% to 120% (Table 3-8). The average percent recovery for Cr (71%) was slightly lower than acceptable, suggesting that this metal may have been consistently "under-recovered" in the actual samples. Therefore, all reported values for this metal were qualified with the SC-C code in the database.

The GFAA analyses included the metals Ag, As, Cd, Sb, Se, and Sn; a total of 16 analytical sets or "batches" of samples were analyzed. The CRM BCSS-1 also was analyzed along with every sample batch as the Laboratory Control Material. Average CRM percent recoveries for all metals fell within the acceptability range of 80% to 120% (Table 3-8), and no results were flagged in the database. The CRM BCSS-1 does not have a "certified" value for silver, but the average recovery for this metal in laboratory spiked samples (matrix spikes) was within quality control limits.

Mercury analyses

For the 1992 Virginian Province mercury analyses, the Certified Reference Material BEST-1 (issued by the National Research Council of Canada) was analyzed along with every sample batch as the Laboratory Control Material (n = 8 sample batches). The average percent recovery of 88% for mercury in this reference material fell well within the acceptability range of 80% to 120%. In addition, an average percent recovery of 102% was achieved for the matrix spike samples analyzed in each batch. Overall, these results indicate acceptable accuracy for the mercury analyses, and no "SC-C" codes were used to qualify the data. The 1992 Virginian Province mercury

results were deemed acceptable for use without qualification.

Table 3-8. Summary results for CRM BCSS-1 (Estuarine Sediment) used as a set control for the 1992 Virginian Province sediment inorganic analyses.

ICP-AES METALS (n = 16 analysis sets or "batches"):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Al	82	3.7	4.5	78	93
Cr	71	3.3	4.7	66	80
Cu	101	3.7	3.7	94	107
Fe	87	3.6	4.2	80	92
Mn	91	3.2	3.5	83	96
Ni	86	2.8	3.3	79	90
Pb	103	15.3	14.9	72	137
Zn	85	3.5	4.1	80	92

GFAA METALS (n = 16 analysis sets):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Ag	na	na	na	na	na
As	111	13.0	11.6	83	135
Cd	102	11.6	11.4	67	119
Sb	101	15.7	15.5	79	130
Se	85	20.8	24.4	45	123
Sn	99	9.8	10.0	83	116

[1] Average percent recovery relative to the SRM certified value.
[2] Standard deviation of the percent recovery values.
[3] Coefficient of variation of the percent recovery values.
[4] Minimum percent recovery for n analysis sets
[5] Maximum percent recovery for n analysis sets

Organic analyses

In general, results for reagent blanks and calibration check samples analyzed with each batch of samples fell within control limits and serve to verify that sample contamination did not occur and that all instruments were calibrated properly throughout the analytical runs. Average recoveries of compounds in matrix spike/matrix spike duplicate samples generally fell within control limits, indicating acceptable analytical performance. However, matrix spike samples are not the most ideal quality control samples because the analytes of interest are not truly incorporated into the matrix in the same manner as an actual field sample. In addition, it can be difficult to evaluate laboratory performance solely on the basis of matrix spike results because it is often equivocal whether low recoveries are due to flawed methodology, poor technique, or a true matrix interference.

Given the above limitations related to the use of matrix spike samples to assess analytical performance, great emphasis was placed on the LCM results. For both the PAH and PCB/pesticide analyses, SRM 1941 (Organics in Marine Sediment, issued by NIST) was analyzed as the LCM along with each batch of field samples. For most of the individual PAH compounds and PCB congeners with "known" concentrations in SRM 1941 (this includes

both "certified" and "non-certified" values), the average percent recovery achieved by the laboratory (based on n = 13 batches for PAHs and n = 13 batches for PCB/pesticides) generally fell within the control limit range of 70% to 130% (Tables 3-9 and 3-10). Whenever the laboratory failed to achieve these average recovery rates for a particular compound, all the results in the 1992 Virginian Province organics dataset for that compound were flagged with the "SC-C" code to indicate the potential inaccuracy inferred from the SRM analysis. It is important to note that the 70% to 130% recovery criteria only applies to compounds having SRM concentrations greater than 10 times the laboratory's detection limit. When compounds occur at concentrations less than about 10 times the detection limit, a greater amount of analytical uncertainty is expected and the normal control limit "acceptability" criteria do not apply.

Table 3-9. Results for SRM 1941 (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1992 Virginian Province sediment PAH analyses (n = 13 analysis sets or "batches").

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
Acenaphthene	127	22.4	17.7	98	167
Acenaphthlylene	57	12.5	21.9	38	79
Anthracene	93	27.8	29.8	59	145
Benz(a)anthracene	88	14.5	16.4	69	109
Benzo(a)pyrene	69	10.4	15.1	52	86
Benzo(e)pyrene	95	19.9	21.0	63	132
Benzo(b+k)fluoranthene	105	17.4	16.6	79	137
Benzo(g,h,i)perylene	111	15.2	13.8	95	146
Biphenyl	118	29.9	25.3	56	153
Chrysene	152	25.6	16.8	118	198
2,6-dimethylnaphthalene	128	32.0	25.0	66	177
Fluoranthene	100	21.5	21.5	70	143
Fluorene	126	21.4	17.0	91	176
Ideno(1,2,3-c,d)pyrene	114	16.2	14.2	84	140
1-methylnaphthalene	115	29.3	25.5	62	153
2-methylnaphthalene	126	40.5	32.0	52	190
1-methylphenanthrene	130	44.6	34.2	69	239
Naphthalene	77	35.9	46.6	8	131
Perylene	71	9.2	13.1	58	89
Phenanthrene	127	28.7	22.7	75	162
Pyrene	113	17.0	15.0	92	156

[1] Listed compounds include those having both "certified" and "non-certified" concentrations in SRM 1941.
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 13 analysis sets
[6] Maximum percent recovery for 13 analysis sets

Table 3-10. Results for SRM 1941 (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1992 Virginian Province sediment PCB/pesticide analyses (n = 13 analysis sets or "batches").

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
PCB 18	47	9.8	20.9	30	57
PCB 28	74	12.0	16.1	49	94
PCB 52	121	21.2	17.5	90	157
PCB 66	94	15.5	16.4	74	122
PCB 101	69	10.4	15.0	56	90
PCB 118	78	11.2	14.2	64	101
PCB 153	71	6.1	8.6	61	80
PCB 105	146	24.1	16.5	102	183
PCB 138	70	6.6	9.4	61	81
PCB 187	67	5.4	8.1	58	77
PCB 180	100	10.1	10.1	88	118
PCB 170	80	7.3	9.2	69	89
PCB 195*	176	27.5	15.7	133	222
PCB 206*	82	11.5	14.0	54	95
PCB 209	86	15.3	17.7	57	104
Dieldrin*	125	52.2	41.9	63	255
Heptachlor epoxide*	160	72.6	45.3	58	281
cis-Chlordane*	108	24.9	23.0	74	150
trans-Nonachlor*	120	28.1	23.4	75	160
4,4'-DDE	91	17.9	19.6	69	125
4,4'-DDD	84	16.4	19.6	65	110
4,4'-DDT*	102	36.0	35.2	41	167

[1] SRM 1941 only lists "non-certified" or informational values for this group of PCB congeners and pesticides (* = concentration in the SRM is less than 10 times the target detection limit).
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 13 analysis sets
[6] Maximum percent recovery for 13 analysis sets

Based on the above, the results for the following organic compounds were flagged with the "SC-C" code in the 1992 database: PCB 101, PCB 105, PCB 18, PCB 187, acenaphthylene, benz(a)pyrene, chrysene, and 1-methylphenanthrene. In addition, although the average percent recovery for naphthalene was within limits (77%), all results for this compound were flagged with the SC-C code because the recoveries between batches exhibited relatively high variability (*e.g.*, 47% coefficient of variation). Although the average SRM percent recoveries for the compounds heptachlor epoxide and PCB 195 also were outside the acceptability range of 70% to 130% (Table 3-10), these compounds occur in the SRM at concentrations less than 10 times the laboratory's detection limit. Therefore, the acceptability criteria do not apply.

The results of the 1992 NIST/NOAA intercomparison exercise suggested that EMSL-Analytical was producing acceptable data quality for PCB and PAH analyses. Pesticide analyses, however, were still questionable. EMSL reported measurable concentrations for two pesticides (heptachlor and 2,4'-DDT) that were not detected by most other laboratories. Concentrations reported for cis-chlordane and trans-nonachlor were considered too high by NIST to be used in calculating the consensus value. EMSL's continuing problems in the analysis of samples for pesticides resulted in all 1992 pesticide data being qualified with the SC-C code, indicating the quality of the data is questionable.

Detection limits of 0.25 ng/g (dry weight) for each PCB congener and pesticide, and 10.0 ng/g (dry weight) for each PAH compound, were achieved in the most of 1992 Virginian Province organics samples.

As previously indicated (see Section 3.2), the laboratory used gas chromatography/electronic capturedetection (GC/ECD) with dual column confirmation for the analysis of PCB congeners and chlorinated pesticides in the 1992 Virginian Province sediment samples. Most values reported in the database for the PCBs and pesticides represent "confirmed" results (*i.e.*, the analyte was detected and could be quantified on both the primary and secondary columns). In general, the rate of second-column confirmation for all reported PCB congeners and chlorinated pesticides was greater than 80%, with the following exceptions (confirmation rate in parenthesis): PCB 195 (75%), heptachlor (26%), heptachlor epoxide (42%), lindane (35%), o,p DDT (77%), and p,p DDT (79%). Whenever an analyte was detected on one column, but was not confirmed on the second column, the result was treated as a "not detect" (*i.e.*, the SC-A code is used to flag the result in the database). Whenever there was a significant discrepancy in the amount detected on the two GC/ECD columns (*i.e.*, greater than a factor of 3 difference), the lower of the two values is reported in the database and flagged with the "SC-D" code. Please note the warning regarding use of this code discussed in Section 3.2.

Total Organic Carbon analyses

All QC results for the analysis of total organic carbon in the 1992 Virginian Province sediment samples fell within required control limits. The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM. The certified concentration of total carbon in thisreference material is 3.69% (percent dry weight). The average percent recovery achieved by thelaboratory for n = 8 batches of TOC samples (*i.e.*, eight separate analyses of CRM PACS-1) was 97.4%, with all values falling within the range 90% to 106%. Since the PACS-1 certified concentration includes both organic carbon and a very small fraction of inorganic carbon, the laboratory's percent recovery values for organic carbon generally are expected to be below 100%. Based on the good overall percent recovery of organic carbon in the Certified Reference Material, the 1992 sediment TOC data were deemed acceptable for use without qualification.

Butyltin analyses

Data users are cautioned that there are deficiencies in the 1992 Virginian Province sediment dataset for butyltin compounds which might limit or preclude the use of these data. The laboratory detected dibutyltin (DBT) in only 18% and monobutyltin (MBT) in only 3% of the samples analyzed in 1992, while tributyltin (TBT) was detected in 73% of the samples. The MDLs established by the laboratory were 5 ng/g dry weight (as tin) for both TBT and DBT, and 12 ng/g dry weight (as tin) for MBT. It is possible that the butyltin compounds of interest were present in many samples at concentrations below these detection limits, and, therefore, the occurrence of butyltin compounds in Virginian Province sediments may be more widespread than indicated by these data.

The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM for these analyses. Average percent recoveries relative to the certified value for n = 10 analysis sets were 77% for TBT, 52% for DBT and 171% for MBT. These values fall outside the acceptable accuracy control limits of 80% to 120%; therefore, all values reported for TBT, DBT and MBT in samples where these compounds were detected are considered estimates (SC-C code) and should be used with discretion.

Acid volatile sulfides analyses

At present there are no Certified Reference Materials available for acid volatile sulfides. For the 1992 Virginian Province samples, the laboratory utilized a laboratory fortified blank sample as the laboratory control material (LCM). The average percent recovery of AVS for n = 68 laboratory fortified blank samples was 93%, suggesting good overall analytical performance. Average percent recoveries for matrix spike samples were somewhat low (55% for n = 9 matrix spike duplicate sets); these low recoveries were attributed to possible matrix effects. In general, the 1992 AVS analyses were deemed acceptable, and no data qualifier codes were applied to these data.

3.3.5 1993 QA Results

A number of significant problems were uncovered during EMAP's QA review of the 1993 sediment chemistry data. These included analytical problems, switched sample IDs, calculation errors, and transcription errors. Several samples were re-analyzed by either EMSL or ERL-N (after EMSL-Analytical's laboratory shut down). All suspected erroneous data have been qualified, corrected, or deleted. Specific discussions are found below.

Major and trace element analyses (except mercury)

For the analysis of major and trace elements by ICP-AES and GFAA, the laboratory generally met the pre-established acceptability criteria (control limits) for the QC samples (*e.g.*, calibration check samples, laboratory reagent blanks, matrix spikes, and Laboratory Control Materials). For the ICP-AES analyses, which included the metals Al, Cr, Cu, Fe, Mn, Ni, Pb, and Zn, a total of 18 analytical sets or "batches" of samples were analyzed. The Certified Reference Material (CRM) "BCSS-1" (Estuarine Sediment, issued by the National Research Council of Canada) was analyzed along with every batch as the Laboratory Control Material. With the exception of Cr, the average percent recovery of each metal (relative to the certified concentration in BCSS-1) was within the acceptability range of 80% to 120% (Table 3-11). The average percent recovery for Cr (73%) was slightly lower than acceptable, suggesting that this metal may have been consistently "under-recovered" in the actual samples. Therefore, all reported values for this metal were qualified with the SC-C code in the database.

The GFAA analyses included the metals Ag, As, Cd, Sb, Se, and Sn; a total of 18 analytical sets or "batches" of samples were analyzed. The CRM BCSS-1 also was analyzed along with every sample batch as the Laboratory Control Material. Average CRM percent recoveries for all metals fell within the acceptability range of 80% to

120% (Table 3-11). The CRM BCSS-1 does not have a "certified" value for silver, but the average recovery for this metal in laboratory spiked samples (matrix spikes) was within quality control limits. Although thepercent recovery of all metals fell within the acceptable range, all values for As, Sb, and Se were qualified with the SC-C code due to high variability of percent recovery in the matrix spiked samples for these metals.

Table 3-11. Summary results for CRM BCSS-1 (Estuarine Sediment) used as a set control for the 1993 EMAP-Estuaries sediment inorganic analyses.

ICP-AES METALS (n = 18 analysis sets or "batches"):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Al	91	5.1	5.6	83	102
Cr	73	1.8	2.5	71	77
Cu	101	2.5	2.5	95	105
Fe	92	1.9	2.1	88	96
Mn	97	1.5	1.5	93	99
Ni	84	2.6	3.1	81	89
Pb	101	18.9	18.7	70	133
Zn	87	2.4	2.8	82	90

GFAA METALS (n = 18 analysis sets):

Element	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
Ag	na	na	na	na	na
As	108	10.9	10.0	84	123
Cd	98	12.6	12.9	71	123
Sb	102	21.2	20.7	67	139
Se	101	26.6	26.3	66	143
Sn	94	10.3	11.0	77	116

[1] Average percent recovery relative to the SRM certified value.
[2] Standard deviation of the percent recovery values.
[3] Coefficient of variation of the percent recovery values.
[4] Minimum percent recovery for n analysis sets
[5] Maximum percent recovery for n analysis sets

During the QA review of the 1993 data, a problem was noted with the metals data associated with event 3145 (Station 188, chemID # 9303471). For Cu, Ni, Pb and Zn the ICP results did not agree well with data generated from a second visit to that station, and the Simultaneously Extracted Metals (SEM) values for these metals were significantly HIGHER than the bulk metals (1993 was the only year SEM metals were analyzed). EMSL checked the ICP results and confirmed them. They then re-extracted and re-analyzed this sample and came up with significantly higher results than for the first analysis; results similar to those from the duplicate sample. It appeared that they may have switched samples, since the ratio of the difference between the original run and re-run was not consistent among metals, ranging from a factor of 2 to a factor of 20. The results of the original analysis and re-analysis are as follows:

Analyte	Original value (μg/g)	Reanalysis value (μg/g)
Cr	22	69
Cu	5.55	46
Mn	210	1,160
Fe	11,800	41,800
Ni	2.44	42
Pb	14.4	51
Zn	44.2	175

The reanalysis values agreed relatively well with the results from a second sample collected at that station and from previous data from that station. The question was then raised "why were the original results so low?" One possibility is that the sample was switched with another one from that batch. To determine if this was the case, data on all samples in the batch were reviewed. Three sediments had similar values to those from the reanalysis of 9303471, so were potential candidates. Since the EMSL chemistry laboratory was no longer active, these samples were sent to ERL-N for analysis. The results of those analyses are as follows:

	Reanalysis results (μg/g: EMSL's original results in parentheses)		
Analyte	Sample 9303328	Sample 9303400	Sample 9303477
Cr	**23.2** (59)	**23.56** (77.9)	**9.13** (68.6)
Cu	**14.48** (34)	**10.30** (38.8)	**1.20** (48.5)
Mn	**299.75** (1,120)	**217.84** (752)	**80.89** (1,220)
Fe	**12,030** (46,700)	**8,834** (36,000)	**4,630** (36,900)
Ni	**11.77** (20)	**12.66** (39.2)	**2.53** (39.1)
Pb	**11.03** (49)	**9.74** (70.8)	**4.01** (47.9)
Zn	**50.09** (123)	**49.39** (174)	**13.48** (181)

Several things are evident. For the first two samples listed, EMSL's values are consistently greater than those determined by ERL-N by at least a factor of two. The results for sample 9303477 are very different. We believe that in EMSLs original analysis, sample 9303477 was switched with 9303471. First, EMSLs reanalysis results for sample 9303471 are very similar to those from 9303477 (which is why it was selected as one for reanalysis at ERL-N). Second, ERL-N's reanalysis of sample 9303477 produced results much lower than the original results for that sample.

Since EMSL reanalyzed 9303471 and came up with more "reasonable" values relative to a second sample taken at that station, the original results were replaced with the reanalysis results. The results from 9393477 are suspected to be erroneous, and have been deleted from the database.

Mercury analyses

For the 1993 mercury analyses, the Certified Reference Material BEST-1 (issued by the National Research Council of Canada) was analyzed along with every sample batch as the LCM (n = 7 sample batches). The average percent recovery of 97% for mercury in this reference material fell well within the acceptability range of 80% to 120%. In addition, an average percent recovery of 95% was achieved for the matrix spike samples analyzed in each batch. Overall, these results indicate acceptable accuracy for the mercury analyses, and no "SC-C" codes were used to qualify the data. The 1993 mercury results were deemed acceptable for use without qualification.

Organic analyses

In general, results for reagent blanks and calibration check samples analyzed with each batch of field samples fell within control limits and serve to verify that sample contamination did not occur and that all instruments were calibrated properly throughout the analytical runs. Average recoveries of compounds in matrix spike/matrix spike duplicate samples generally fell within control limits, indicating acceptable analytical performance. However, matrix spike samples are not the most ideal quality control samples because the analytes of interest are not truly incorporated into the matrix in the same manner as an actual field sample. In addition, it can be difficult to evaluate laboratory performance solely on the basis of matrix spike results because it is often equivocal whether low recoveries are due to flawed methodology, poor technique, or a true matrix interference.

Given the above limitations related to the use of matrix spike samples to assess analytical performance, great emphasis was placed on the LCM results. For both the PAH and PCB/pesticide analyses, SRM 1941 or 1941a (Organics in Marine Sediment, issued by NIST) was analyzed as the LCM along with each batch of field samples. For most of the individual PAH compounds and PCB congeners with "known" concentrations in SRM 1941a (this includes both "certified" and "non-certified" values), the average percent recovery achieved by the laboratory (based on n = 10 batches for PAHs and n = 10 batches for PCB/pesticides) generally fell within the control limit range of 70% to 130% (Tables 3-12 and 3-13). Whenever the laboratory failed to achieve these average recovery rates for a particular compound, all the results in the 1993 database for that compound were flagged with the "SC-C" code to indicate the potential inaccuracy inferred from the SRM analysis. It is important to note that the 70% to 130% recovery criteria only applies to compounds having SRM concentrations greater than 10 times the laboratory's detection limit. When compounds occur at concentrations less than about 10 times the detection limit, a greater amount of analytical uncertainty is expected and the normal control limit "acceptability" criteria do not apply.

Based on the above, the results for the following organic compounds were flagged with the "SC-C" code in the 1993 database: PCB 101, PCB 18 and chrysene. Although the concentration of PCB 18 in the SRM was less than 10x the detection limit, the very high mean percent recovery (343%) and high variability (range of recoveries from 90% to 1,550% with a CV of 138%) resulted in the SC-C code being applied to all values for PCB 18. In addition, although the average percent recovery for PCB 206 was within limits (110%), all results for this compound were flagged with the SC-C code because the recoveries between batches exhibited relatively high variability (*i.e.*, 78% coefficient of variation).

The SC-C code was also applied to several specific samples for which the data were suspect (*e.g.*, poor agreement between field splits). As in previous years, all pesticide data were qualified with this code. The problems identified with the pesticide data from station 725 (see below, Table 3-14) support this action.

Further review of the data illuminated additional concerns regarding the 1993 PCB results. One problem noted concerned the formulation and use of control charts for assessing precision. Of the ten sample batches submitted, the first two contain results for SRM1941 while the remaining eight used SRM1941a (both SRMs

are acceptable for use with marine sediment). The control limits for SRM1941a results were computed using the SRM1941 (wrong SRM) results from the first two batches, the SRM1941a results from two 1993 batches (control limits being set by the same data on which they are to measure precision), and two additional SRM1941a samples prepared earlier. The associated control limits are so wide (24-162%RSD with average 44% for all analytes) that in all of the eight batches to which they apply only once did an analyte exceed them (PCB 18 at >1500% recovery). These data may nothave exceeded control limits but they are not necessarily precise. Coefficients of variation for SRMs calculated in previous reviews tend to suggest the 1993 PCB and pesticide data are the most variable since the 1990 data set, which has the highest variability. As a result, all 1993 PCB data have been assigned the "SC-C" code indicating potential problems with the data.

A detection limit of 0.25 ng/g (dry weight) generally was achieved for each PCB congener and pesticide and a detection of 10.0 ng/g (dry weight) was achieved for each PAH compound in the majority of samples analyzed.

Some problems with specific samples were also noted. As part of a review of the field split data it was noticed that the total of PAHs from 3069030 was 1,438 ng/g and from the split (3006030) all PAHs were non-detected. The response from EMSL was "The pesticide data for these two field duplicates appear fairly high and agree well, even though no PAHs were found in 9302940 [3006030], yet high levels were reported in 9302943 [3069030]. Examination showed that the samples were correctly identified and appeared similar. However, examination of the extracts showed them to be distinctly different. There was an extract in that same run 9302933 [3115030], which exhibited similar results to 9302043. We are checking to see if we have enough sample left to re-extract and verify our results."

EMSL did re-extract and run all three samples. Results showed that they did indeed mis-label the original extracts for 3006030 and 3115030. The re-analysis of sample 3069030 showed poor precision relative to the original run, with differences of concentrations between runs approaching a factor of 3 (*e.g.*, 132 vs 329). The original data have been replaced with the results of the re-analysis.

As part of the QA review of the data, ERL-N scientists noticed that the concentrations of dieldrin and p,p' DDT from station 725 in the Providence River, RI appeared unreasonably high. This sample was shipped to ERL-N and analyzed for pesticides and PCBs. The results can be found in Table 3-14. The dieldrin and DDT results were found to be erroneous. Because of the uncertainty in pesticide results from this station, all 1993 pesticide data from station 725 have been deleted from the database.

Table 3-12. Results for SRM 1941 and 1941a (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1993 sediment PAH analyses (n = 10 analysis sets or "batches"). Note that since the results are presented as percent recovery, 1941 results and 1941a results were not separated.

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
Acenaphthene	118	16.6	14.1	100	133
Acenaphthlylene	73	17.8	24.4	61	94
Anthracene	86	15.2	17.8	57	110
Benz(a)anthracene	119	11.4	9.6	101	137
Benzo(a)pyrene	93	22.7	24.4	68	137
Benzo(e)pyrene	104	22.4	21.5	80	147
Benzo(b+k)fluoranthene	129	23.5	18.2	97	176
Benzo(g,h,i)perylene	99	22.4	22.6	74	146
Biphenyl	73	17.9	24.7	60	112
Chrysene	136	9.7	7.2	120	153
2,6-dimethylnaphthalene	74	20.1	27.0	55	116
Fluoranthene	97	6.8	7.1	88	110
Fluorene	109	8.7	8.0	102	119
Ideno(1,2,3-c,d)pyrene	107	20.1	18.8	82	143
1-methylnaphthalene	70	17.2	24.4	47	106
2-methylnaphthalene	88	18.0	20.5	57	120
1-methylphenanthrene	92	17.4	19.0	69	120
Naphthalene	88	25.3	28.8	49	120
Perylene	84	21.5	25.6	63	131
Phenanthrene	102	10.7	10.5	88	120
Pyrene	93	9.5	10.2	82	109

[1] Listed compounds include those having both "certified" and "non-certified" concentrations in SRM 1941a.
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 10 analysis sets
[6] Maximum percent recovery for 10 analysis sets

Table 3-13. Results for SRM 1941 and 1941a (Organics in Marine Sediment) used as the set control (Laboratory Control Material) for the 1993 sediment PCB/pesticide analyses (n = 10 analysis sets or "batches"). Note that since the results are presented as percent recovery, 1941 results and 1941a results were not separated.

Compound[1]	Average[2]	Stdv[3]	C.V.[4]	Min[5]	Max[6]
PCB 8*	106	51.6	48.5	39	218
PCB 18*	343	473	138	90	1550
PCB 28	77	13.4	17.3	61	99
PCB 44	84	26.0	30.8	59	129
PCB 52	93	22.7	24.4	68	125
PCB 66*	119	58.9	49.6	84	278
PCB 101	69	14.9	21.3	53	96
PCB 118	76	11.7	15.4	64	96
PCB 153	72	16.6	23.1	51	96
PCB 105	84	21.1	25.1	44	117
PCB 128*	77	24.9	32.4	40	113
PCB 138	91	21.9	23.9	66	144
PCB 187	91	23.6	26.1	65	134
PCB 180	121	24.0	19.8	86	172
PCB 170	96	31.5	32.7	38	132
PCB 206	110	86.2	78.4	50	310
PCB 209	81	13.8	17.0	64	101
Dieldrin*	136	94.8	69.6	49	370
cis-Chlordane*	164	66.9	40.7	97	259
trans-Nonachlor*	89	25.8	29.0	43	122
Hexachlorobenzene	76	21.2	27.9	48	114
2,4'-DDE*	250	123	49.1	112	466
4,4'-DDE	110	28.3	25.7	67	169
4,4'-DDD	99	21.0	21.3	72	145
4,4'-DDT*	137	168	122	12	582

[1] Listed compounds include those having both "certified" and "non-certified" concentrations in SRM 1941a (* = concentration in the SRM is less than 10 times the target detection limit).
[2] Average percent recovery relative to the SRM value.
[3] Standard deviation of the percent recovery values.
[4] Coefficient of variation of the percent recovery values.
[5] Minimum percent recovery for 10 analysis sets
[6] Maximum percent recovery for 10 analysis sets

Table 3-14. Results of re-analysis of sediments from Station 725 (results in ng/g dry weight).

Analyte	ERL-N Results	ERL-N Replicate	EMSL Results	EMSL Confirmation column
PCB 8	1.32	1.42	4.42	
HCB	0.30	0.29	0.71	2
PCB 18	3.50	3.18	17.6	
PCB028	6.02	5.35	11.3	
PCB 52	11.9	10.1	15.8	
PCB 44	6.42	5.64	11.2	
PCB 66	8.16	6.02	11.6	
PCB 101	21.7	17.5625	57.8	
PP DDE	10.8	10.4	65.3	**16**
PCB 118	16.6	14	50.3	
PCB 153	25.6	21.7	76.6	
PCB 105	7.52	5.56	10.4	
PCB 138	27.1	20.8	97	
PCB 187	14.4	10.7	18.3	
PCB 128	3.63	2.91	8.17	
PCB 180	19.4	15.9	82	
PCB 170	6.85	5.62	14.6	
PCB 195	6.09	4.53	10.2	
PCB 206	12.2	9.57	13.2	
PCB 209	13.2	10.3	18.6	
LINDANE	0.19	0.69	nd	
CISCHLORDANE	3.76	3.51	17.2	**3.6**
T-NONACHLOR	2.80	2.63	3.06	5.4
PP DDD	20.9674	18	62.4	29
PP DDT	1.37	4.02	251	87
Dieldrin	Not analyzed	8.92	170	58

Total Organic Carbon analyses

All QC results for the analysis of total organic carbon in the 1993 sediment samples fell within required control limits. The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM. The certified concentration of total carbon in this reference material is 3.69% (percent dry weight). The average percent recovery achieved by the laboratory for n = 8 batches of TOC samples (*i.e.*, 8 separate analyses of CRM PACS-1) was 95.8%, with all values falling within the range 90% to 106%. Since the PACS-1 certified concentration includes both organic carbon and a very small fraction of inorganic carbon, the laboratory's percent recovery values for organic carbon generally are expected to be below 100%. Based on the good overall percent recovery of organic carbon in the Certified Reference Material, the 1993 sediment TOC data were deemed acceptable for use without qualification.

Butyltin analyses

Data users are cautioned that there are deficiencies in the 1993 sediment dataset for butyltin compounds which might limit or preclude the use of these data. The MDLs established by the laboratory were 5 ng/g dry weight (as tin) for both TBT and DBT, and 12 ng/g dry weight (as tin) for MBT. It is possible that the butyltin compounds of interest were present in many samples at concentrations below these detection limits, and, therefore, the occurrence of butyltin compounds in Virginian Province sediments may be more widespread than indicated by these data.

The Certified Reference Material PACS-1 (issued by the National Research Council of Canada) was utilized as the LCM for these analyses. Average percent recoveries relative to the certified value for n = 11 analysis sets were 74% for TBT, 74% for DBT and 188% for MBT. These values fall outside the acceptable accuracy control limits of 80% to 120%. Therefore, all values reported for TBT, DBT and MBT in samples where these compounds were detected are considered estimates (SC-C code) and should be used with discretion.

Acid volatile sulfides analyses

At present there are no Certified Reference Materials available for acid volatile sulfides. For the 1993 samples, the laboratory utilized a laboratory fortified blank sample as the LCM. The average percent recovery of AVS for n = 60 laboratory fortified blank samples was 94%, suggesting good overall analytical performance. With the exception of two batches the 1993 AVS analyses were deemed acceptable.

The "SC-C" code was applied to all results from AVS run #384. This batch contained one set of blind field splits. The AVS concentration reported for one of the splits was 22.6 mg/kg. The concentration reported for the second was 1.77 mg/kg. The EMSL ran duplicate analyses (as part of routine QA) on the second sample and determined a concentration of 7 mg/kg. All three of these concentrations should be the same. The differences suggest a lack of precision for this batch; therefore all samples analyzed as part of this batch were flagged. It should be noted that the precision of laboratory duplicates for AVS was generally good.

The "SC-C" code was also applied to all results from AVS run #387. As part of the QA protocol the laboratory is required to run one sample in duplicate. The relative percent difference between duplicates must be less than 20%. The reported concentrations of the duplicate (LD1 and LD2) were 2100 and 2240 mg/kg with a RPD of 6.4% HOWEVER, the QA summary provided with that batch stated "The original analysis of the duplicate on 9/2/93 [the day all the samples from that batch were analyzed] did not meet the RPD acceptance criterion. The duplicate analyses were performed again on 9/16/93, and the results reported here. LFM and CLE recoveries associated with the reanalysis of the duplicates were reported in Run 383." The entire batch should have been reanalyzed. This shows that the precision on the day the samples were run was poor. The original data were requested

from EMSL. The original LD1 and LD2 were 1910 and 3950 mg/kg with an RPD of 70%. When EMSL provided us the original numbers they also informed us of a transcription error in the moisture content of the sample. A value of 62.9% was entered instead of 82.9%. Therefore the values reported in the original file of 2100 and 2240 mg/kg should be 4440 and 4740 mg/kg respectively. However, the moisture content reported in the electronic file originally received from Cincinnati was, in fact, 83% (rounded), not 63%. EMSL investigated this and informed us that despite the fact that our file lists 83% as the water content used to calculate the original numbers, and 83% is the correct value, a value of 63% was used in the calculation of the AVS concentrations. They could not explain how this happened.

Section 4
QA Results for Fish Contaminant Analyses

4.1 Background

Measurement Quality Objectives for the analysis of chemical contaminants in EMAP-E tissue samples are specified in the 1991 Province Quality Assurance Project Plan (Valente and Schoenherr 1991). This plan requires each EMAP-E laboratory to analyze the following types of quality control samples along with every batch or "set" of field chemistry samples: laboratory reagent blanks, calibration check standards, laboratory fortified sample matrix (matrix spike), laboratory fortified sample matrix duplicate (matrix spike duplicate), laboratory duplicate, and Laboratory Control Material. Results for these QC samples must fall within certain pre-established control limits for the analysis of a batch of samples to be considered acceptable.

Standard or Certified Reference Materials typically are used by EMAP-E laboratories as their Laboratory Control Material. SRMs and CRMs have known or "certified" concentrations of the analytes being measured and therefore are useful for assessing both accuracy and precision. The QA Project Plan requires the laboratory's percent recovery (relative to the certified concentration in the reference material) to fall within certainpre-established control limits to be considered acceptable. If the laboratory consistently fails to meet these acceptability criteria for the CRM or SRM analysis, the values reported for the failed analytes are considered to be suspect (biased) and are flagged in the database, as described in the following section.

Fish were collected from trawls conducted at each station. Individuals of "target" species were selected for contaminant analysis. These individuals were tagged, wrapped in aluminum foil, and frozen. In the laboratory, fish were cleaned, scaled, fileted, and composited by species. An analytical sample consisted of the edible flesh from three to five individuals of a single species from a station.

Because of budget constraints and the poor distribution of target species across the Province, 1991 was the only year in which fish were analyzed for contaminants.

4.2 1991 Results

Major and trace element analyses

For the 1991 Virginian Province analysis of major and trace elements, the laboratory generally met the pre-established acceptability criteria (control limits) for the QC samples (*e.g.*, calibration check samples, laboratory reagent blanks, matrix spikes, and LCMs). The control limits for inorganic analytes is ± 20% of the CRM certified value. These criteria were generally met (Table 4-1). The average percent recovery for Pb (DOLT) was slightly high; however, the value for the DORM CRM was within the acceptable range and the confidence intervals around the DOLT certified value were rather large.

A problem was noted by the laboratory in analysis of selected samples for mercury. The laboratory analyzed 84 composite samples and 40 individual fish. The analytical laboratory experienced a mercury-contamination problem with their freeze-drier, resulting in contamination of all 40 individual-fish samples. As a result, these data had to be deleted from the database. However, EMAP-VP's assessment was focused on the composite samples, and none of these were contaminated.

With the removal of the above-mentioned Hg data from the database, the only flags applied are the "A" and

"B" codes described in the sediment chemistry section (Section 3.2).

Table 4-1. Summary results for CRMs DOLT and DORM (Dogfish liver and muscle tissue, respectively) used as a set control for the 1991 Virginian Province fish tissue inorganic analyses. Average reported values are based on six separate analyses of the CRMs.

Element		Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
As	DOLT	101.2	2.5	2.5	98.0	104.0
	DORM	99.3	2.5	2.5	94.9	101.7
Cd	DOLT	83.3	8.2	9.8	69.6	90.9
	DORM	93.0	10.4	11.2	81.4	104.7
Cr	DOLT	118.8	15.2	12.8	102.5	137.5
	DORM	106.3	8.8	8.3	97.5	117.5
Cu	DOLT	91.4	9.1	10.0	81.7	107.2
	DORM	79.4	5.3	6.7	75.7	88.9
Fe	DOLT	98.9	1.8	1.8	96.2	101.3
	DORM	105.7	11.9	11.2	95.4	125.0
Hg	DOLT	NA	NA	NA	NA	NA
	DORM	91.9	4.9	5.3	86.5	100.3
Ni	DOLT	103.8	52.9	50.9	50.0	188.5
	DORM	89.0	13.5	15.2	76.7	111.7
Pb	DOLT	130.4	25.7	19.7	92.7	164.7
	DORM	89.5	35.3	39.5	55.0	142.5
Se	DOLT	102.1	2.9	2.8	99.7	107.4
	DORM	102.8	4.6	4.4	96.9	108.0
Zn	DOLT	101.7	2.7	2.7	98.0	105.7
	DORM	100.7	3.9	3.8	96.2	106.6

[1] Average percent recovery relative to the CRM certified value.
[2] Standard deviation of the percent recovery values.
[3] Coefficient of variation of the percent recovery values.
[4] Minimum percent recovery for analysis sets
[5] Maximum percent recovery for analysis sets

Organic analyses

Due to a miscommunication within the analytical laboratory, EMAP QA protocols were not followed during the analysis of EMAP-VP 1991 fish tissue samples for organic analytes. However, sufficient data are available for an evaluation of the quality of those samples. First, prior beginning processing of EMAP samples the laboratory participated in a performance evaluation. Based on 11 separate analyses of SRM1974 (Organics in Mussel Tissue), it was determined that the laboratory was sufficiently proficient to begin analyzing EMAP samples. The results of this performance evaluation are listed in Table 4-2. Second, matrix spiked samples were analyzed with each batch, and these results fell well within EMAP's control limits (Table 4-3). Third, during the same time period when the laboratory was processing EMAP samples, they were also processing samples for NOAA's NS&T Program. SRM 1974 was used as the laboratory control material for those samples, and was analyzed with each analytical batch. The laboratory has provided EMAP with those results, which fall within EMAP control limits. Fourth, the QA protocols the lab followed for EMAP samples require the analysis of duplicate samples with each batch. Those results were provided to EMAP and showed excellent precision, with a maximum Relative Percent Difference for an analyte in a given set generally being less than 10%.

The only flags applied are the "A" and "B" codes described in the sediment chemistry section.

Table 4-2. Performance evaluation results for analysis of organic contaminants in tissue. Average reported values are based on 11 separate analyses of SRM 1974 (Organics in Mussel Tissue) performed on different days.

Analyte	Average reported value	NIST non-certified value[1]	Percent difference
alpha-chlordane	21.2	26 ± 1	-15%
trans-nonachlor	17.7	21 ± 5	0%
Dieldrin	11.3	8 ± 4	0%
2,4'-DDE	M[2]	5.8 ± 0.6	NA
4,4'-DDE	41.4	48 ± 2	-10%
2,4'-DDD	5.8	20 ± 7	-55%
4,4'-DDD	46.5	68 ± 3	-28%
2,4'-DDT	5.0	4 ± 1	0%
4,4'-DDT	3.6	3 ± 2	0%
PCB 18	20.9	24 ± 9	0%
PCB 28	85.2	62 ± 3	31%
PCB 44	72.4	65 ± 23	0%
PCB 52	113.7	98 ± 39	0%
PCB 66	98.7	110 ± 5	-6%
PCB 101	127.0	105 ± 11	9%
PCB 105	46.9	45 ± 3	-2%
PCB 118	115.9	110 ± 5	1%
PCB 128	17.3	15 ± 2	2%
PCB 138	122.2	110 ± 11	1%
PCB 153	153.9	145 ± 8	1%
PCB 180	13.3	13 ± 1	0%
PCB 187	27.2	30 ± 1	-6.2%

1 NIST non-certified values with 95% confidence intervals presented in the certificate of analysis for SRM 1974. Reported values falling within these confidence intervals are listed as having a percent difference of 0%.

2 Matrix interference, no peak was found for 2,4'-DDE

Table 4-3. Results of laboratory-fortified matrix spikes analyzed with each batch of fish tissue organic samples analyzed (n=10). Values are percent recovery of the spike.

Analyte	Average[1]	Stdv[2]	C.V.[3]	Min.[4]	Max.[5]
aldrin	95.2	10.9	11.5	83	114
alpha-chlordane	100.2	10.9	10.9	82	112
trans-nonachlor	99.3	12.1	12.2	80	117
Dieldrin	95.2	14.0	14.6	71	118
2,4'-DDE	94.4	9.6	10.2	84	112
4,4'-DDE	99.1	10.5	10.6	86	118
2,4'-DDD	101.6	9.3	9.2	87	112
4,4'-DDD	98.7	12.7	12.9	76	118
2,4'-DDT	101.9	12.5	12.3	79	120
4,4'-DDT	100.5	15.2	15.1	74	118
Total PCBs	99.8	7.7	7.7	87	114

[1] Average percent recovery relative to the concentration of the spike.
[2] Standard deviation of the percent recovery values.
[3] Coefficient of variation of the percent recovery values.
[4] Minimum percent recovery for analysis sets
[5] Maximum percent recovery for analysis sets

Section 5
QA Results for Particle Size Analyses

5.1 Background

At each station crews collected three samples for benthic infaunal analysis and a sediment homogenate which was split for chemistry and toxicity analyses. Associated with each of these samples was an aliquot removed for particle size analysis (percent silt/clay). The annual QA Plans require that approximately 10% of theseanalyses be performed in duplicate and the maximum allowable percent difference for the predominant fraction (silt/clay or sand) is 10%.

5.2 Laboratory Audits

In 1990 sediment particle size analyses were performed at Versar, Inc. in Columbia, MD. This facility was audited by the EMAP-E QAO during the period 15-16 November 1990. No major problems were identified in this audit. The main recommendation in the audit report was the need for minor revisions, mostly in the form of clarifications, to Versar's methods manual and data forms. Versar met all measurement quality objectives in performing the grain size analyses on 1990 samples.

In 1991 and 1993 particle size analyses were performed by SAIC, on-site at the EPA Environmental Research Laboratory in Narragansett, RI. Although no formal auditswere performed, SAIC technicians were closely monitored by the EMAP-VP QA Coordinator (located in the same facility). Prior to the start of analysis, the technician was required to demonstrate proficiency through the analysis of sediments with a variety of grain sizes. Results of these analyses met EMAP QA criteria and the technician was permitted to begin analysis of EMAP samples.

In 1992 particle size analyses were performed by EPA personnel at the EPA Environmental Research Laboratory in Narragansett, RI. Although no formal audits were performed, EPA technicians were closely monitored by the EMAP-VP QA Coordinator (located in the same facility). Prior to the start of analysis, the technician was required to demonstrate proficiency through the analysis of sediments collected by EMAP in 1991 with a variety of grain sizes. Results of these analyses met EMAP QA criteria and the technician was permitted to begin analysis of 1992 EMAP samples.

5.3 Qualifier Codes for Particle Size Data

No codes currently exist for particle size data, indicating all data meet QA criteria and are suitable for EMAP assessment purposes.

5.4 1990 QA Results

All "sediment grain size" and "benthic grain size" samples collected perstation were analyzed forthe determination of percent silt/clay. Approximately 10% of these analyses were performed in duplicate and the percent difference determined as per the EMAP-VP 1990 QA Project Plan. The maximum allowable percent difference for the predominant fraction (silt/clay or sand) is 10%. The mean difference for the samples analyzed was 2.78%, with none exceeding 10% so no remedial action or retesting was required.

5.5 1991 QA Results

Because of budget constraints, not all benthic grain size samples were analyzed. However, at least one sample, of the three collected per station, was analyzed. These data are solely used in the interpretation of benthic community data, and all parties, including those conducting the benthic infaunal analyses agreed on its acceptability. Grain size information presented in the Statistical Summaries is from "sediment grain size" analyses.

All "sediment grain size" and at least one "benthic grain size" sample per station were analyzed for the determination of percent silt/clay. Approximately 10% of these analyses were performed in duplicate and the difference determined as per the EMAP-VP 1991 QA Project Plan. The maximum allowable percent difference for the predominant fraction (silt/clay or sand) is 10%. The mean difference for the samples analyzed was less than 1%, with none exceeding 10% so no remedial action or retesting was required.

5.6 1992 QA Results

All "sediment grain size" and "benthic grain size" samples were analyzed for the determination of percent silt/clay. Approximately 10% of these analyses were performed in duplicate and the difference determined as per the EMAP-VP 1992 QA Project Plan. The maximum allowable percent difference for the predominant fraction (silt/clay or sand) is 10%. The mean difference for the samples analyzed was less than 1%, with none exceeding 10% so no remedial action or retesting was required.

5.7 1993 QA Results

All "sediment grain size" and "benthic grain size" samples were analyzed for the determination of percent silt/clay. Approximately 10% of these analyses were performed in duplicate and the difference determined as per the EMAP-VP 1993 QA Project Plan. The maximum allowable percent difference for the predominant fraction (silt/clay or sand) is 10%. The mean difference for the 50 samples analyzed in duplicate was 1.5%, with none exceeding 5% so no remedial action or retesting was required (the control limit is 10%). In addition, the 1993 QA Project Plan states that if the relative standard deviation (RSD) among the three benthic grain size samples collected from a single station exceeds 20%, the calculations should be checked by the laboratory. The RSD for samples from three stations exceeded 20% and those samples were reanalyzed. The results did not change.

Section 6
QA Results for Sediment Toxicity Testing

6.1 Background

Ten-day laboratory toxicity tests, using the amphipod *Ampelisca abdita*, were performed on sediments collected at each station in the Virginian Province. This test was employed during each year of monitoring of the Province. The QA Plans describe certain requirements for this test to be acceptable, including minimum control survival, physical characteristics (temperature, salinity, dissolved oxygen concentration), and the use of water-only reference toxicant tests.

In addition to the *Ampelisca* test, in 1990 samples from low-salinity waters were tested using the freshwater amphipod *Hyalella azteca* to test the response of *Ampelisca* in low salinity habitats. As a result of this testing, it was determined that the results of the *Ampelisca* test conducted on low-salinity sediments were representative, and the *Hyalella* test was not utilized in subsequent years.

6.2 Data Qualifier Codes For Sediment Toxicity

Ten data qualifier codes, or "flags, exist to describe EMAP-VP sediment toxicity data (Table 6-1), based on the criteria described in the Virginian Province Quality Assurance Plans. Data for some tests that failed QC are included in the dataset because, under certain circumstances, they may be of use to non-EMAP users. These data are flagged with the code describing why they are unacceptable (*e.g.*, ST-D) and the ST-L code, indicating that they were not used in EMAP's assessment of the ecological condition of the Province.

An example of why these data were kept in the database is as follows. Control survival in a particular test was unacceptably low (*e.g.*, 50%), and there was insufficient sediment available to repeat the test. All data associated with this test would automatically receive the ST-D and ST-I flags. However, survival in some of the treatments (*i.e.*, test chambers with sediment from individual stations) was high (*e.g.*, 95%). The conclusion could be drawn that, because of the high survival in the experimental treatment, the sediment from that station is <u>not</u> toxic. Users interested in data from particular stations may find this information useful. However, when control survival is low there is no means by which to classify sediment as toxic. Therefore, treatments associated with an unacceptable control may be classified as non-toxic or unknown: they can never be classified as toxic. This results in an inherent bias which makes these data unacceptable for use by EMAP; therefore they are flagged with the ST-L code.

6.3 Laboratory Audits

Sediment toxicity testing was conducted at the SAIC Environmental Testing Center. This facility was audited during the period 20-23 August 1990 by a team consisting of the EMAP-E QAO, the EMAP QA Coordinator, and technical representatives of the EMSL-Cincinnati laboratory. No major problems were identified; however, the audit report noted that SOP's needed to be developed and in-house audits should be performed more frequently. Corrective actions were implemented in response to the audit recommendations as described in a memo from the EMAP-E QA Coordinator to the EMAP QA Coordinator dated 16 October 1990.

Table 6-1. QA Qualifier Codes associated with sediment toxicity data.

Code	Description
ST-A	No QA Comment
ST-C	Fewer than 5 replicates were tested
ST-D	Mean control survival < 85 %
ST-E	Sample held for >30 days prior testing
ST-G	No reference toxicant test was run
ST-H	Hardness and alkalinity not measured (1990 only)
ST-I	Control survival in one replicate <80%
ST-J	Physical parameters out of bounds
ST-K	<20 animals used per replicate
ST-L	Not Used in Province Assessment

One of the findings of this audit was that the two week maximum holding time for samples created logistical problems for the laboratory, and data demonstrate that extending this to fourweeks does not result in degradation of the sample. As a result, the holding time for sediment toxicity samples was increased from two to four weeks.

A follow-up audit was conducted by the EMAP-VP and EMAP-E QA Coordinators on 5 September 1991. Their findings showed the laboratory to have corrected any short-comings noted in the earlier audit, and they were particularly impressed by the lab's in-house documentation and tracking programs. One recommendation made by the auditors was that the representativeness of preserving amphipods in formalin at the completion of a test, rather than picking them live, be assessed. Data showing the representativeness of this methodology were provided by the laboratory and were satisfactory.

6.4 1990 QA Results

As per the QA Project Plan, the laboratory was required to maintain a control chart for toxicity testing using a reference toxicant. The laboratory used cadmium chloride as their reference material, running a standard 96-hour water-only toxicity test whenever EMAP samples were run. The control chart shows that the LC50 for cadmium chloride ranged from approximately 0.4 to 1.2 mg/L, with all values falling within two standard deviations of the mean as required in the QA Plan.

Of the 126 samples from base stations analyzed (including those duplicated for the *Hyalella* test), only one was assigned the "ST-L" code.

6.5 1991 QA Results

As per the QA Project Plan, the laboratory was required to maintain a control chart for toxicity testing using a reference toxicant. This provides an indication of the "quality" of the test organisms relative to those previously used. The laboratory used SDS (sodium dodecyl sulfate) as their reference material, running a standard 48-hour water-only toxicity test with SDS whenever EMAP samples were run. The control chart shows that the LC50 for SDS ranged from 4.0 to 8.37 mg/L, with all values falling within two standard deviations of the mean as required in the QA Plan.

Several tests failed to meet EMAP QA requirements for control organism survival. Field crews recollected sediment from those stations included in the failed tests. Of the 19 tests run, three exhibited control organism survival less than the required 85% (this was following repeating all tests that failed on the first attempt). These tests were assigned the "ST-L" flag or were deleted from the database and were not included in the dataset utilized in EMAP's assessment of the ecological condition of the Province. As a result of these failures, the volume of sediment collected at each station was increased in 1992 to allow for retesting without the need to redeploy crews for additional sediment collection.

6.6 1992 QA Results

As per the QA Project Plan, the laboratory was required to maintain a control chart for toxicity testing using a reference toxicant. The laboratory used SDS (sodium dodecyl sulfate) as their reference material, running a standard 48-hour water-only toxicity test with SDS whenever EMAP samples were run. The control chart shows that the LC50 for SDS ranged from < 2.57 to 11.2 mg/L, with all but the lowest value falling within two standard deviations of the mean as required in the QA Plan (one in 20 tests would be expected to fall outside of two standard deviations). Results of the one reference toxicity test falling outside two standard deviations of the mean were examined, as were all tests performed during the same time period. No anomalies in the tests were apparent and no re-testing was performed.

No samples were assigned the "ST-L" code.

6.7 1993 QA Results

As per the QA Project Plan, the laboratory was required to maintain a control chart for toxicity testing using a reference toxicant. The laboratory used SDS (sodium dodecyl sulfate) as their reference material, running a standard 48-hour water-only toxicity test with SDS whenever EMAP samples were run. The control chart shows that the LC50 for SDS ranged from 5.32 to 8.59 mg/L, with all the values falling within two standard deviations of the mean as required in the QA Plan. Several treatments contained fewer than five replicates (ST-C code), but no infractions were serious enough to warrant discarding data.

No samples were assigned the "ST-L" code.

Section 7
QA Results for Macrobenthic Community Assessments

7.1 Background

Three replicate sediment samples were collected by field crews at each station for macrobenthic community assessments, including species composition, abundance, and biomass. Two QAsteps were required by the EMAP-VP 1990-1993 QA Project Plans: in-house QC checks (i.e., resorts, recounts, and ID confirmation) on 10% of each technician's work, and verification of species identification by an independent laboratory. Both laboratories performing these analyses, as well as the experts contracted for the independent verification of species taxonomy, have a long record of performing benthic infaunal analyses.

7.2 Laboratory Audits

Macrobenthic community assessments for freshwater samples were performed at Versar, Inc. in Columbia, MD. Versar has subcontracted another laboratory (Cove Corporation in Lusby, MD) to perform the macrobenthic community analyses on samples from saline environments. The two facilities were audited by the EMAP-E QAO during the period 15-16 November 1990. No major problems were identified in this audit. The main recommendation in the audit report was the need for minor revisions, mostly in the form of clarifications, to Versar's methods manual and data forms. Both Versar and Cove Corporation met all measurement quality objectives in performing the grain size and benthic community analyses on 1990 samples.

No subsequent audits were conducted; however, voucher specimens were evaluated by independent laboratories as described in Sections 7.4 and 7.5.

7.3 Data Qualifier Codes for Benthic Community Analyses

No codes currently exist for benthic community analyses, indicating all data meet QA criteria and are suitable for EMAP assessment purposes.

7.4 1990 QA Results

Two QA steps were required by the EMAP-VP 1990 QA Project Plan: 10% recounts and independent verification of species identification. The recounts (multiple types - see Table 7-1) and preliminary species verification were performed by the laboratory responsible for the analyses. These in-house QC measures met the requirements established in the QA Plan. Definitive verification of species identification was performed by an independent laboratory and the results are described below.

External reviews of the taxonomic reference collections (*i.e.*, voucher specimens) maintained by both Versar and Cove were completed in 1990. Taxonomic experts at SAIC's Woods Hole office performed the review of the Cove Corporation reference collection of marine macroinvertebrates. This review disclosed that less than 5% of the total number of species had been misidentified. The species misidentifications subsequently were corrected in the EMAP-E database and the taxonomic experts at Cove Corp. used these results to improve their future accuracy for the species in questions.

Table 7-1. Results of recounts performed by the laboratory processing benthic infauna samples in 1990. Approximately 10% of all samples were processed in duplicate.

Measurement	Mean Error	Range of Error
Benthic sorting	3.06%	0 - 10%
Species identification and enumeration	1.37%	0 - 7.7%
Biomass	0.23%	0 - 1.24%
Weighing blanks for biomass	<0.0001g	0 - 0.0013g

7.5 1991 QA Results

Two QA steps were required by the EMAP-VP 1991 QA Project Plan: in-house QC checks (i.e., resorts, recounts, and ID confirmation) on 10% of each technician's work, and independent verification of species identification. The recounts (multiple types - see Table 7-2) and preliminary species verification were performed by the laboratory performing the analyses. Most of these met the requirements established in the QA Plan. Definitive verification of species identification was performed by an independent laboratory and the results are described below.

A total of 137 specimens collected from oligohaline stations were sent to the Aquatic Resources Center in Franklin, TN for independent taxonomic verification. Eleven (8%) were mis-identified, representing 8 species. The identification of an additional 15 specimens could not be confirmed because of the condition of the specimen (*e.g.*, key taxonomic features missing or destroyed, or male needed for identification and only females sent).

The identification of many of these species is difficult. Misidentified species were closely related taxonomically to the "true" species. In general, the report on species verification was "largely favorable" indicating the analytical laboratory performed well. Suggestions were made regarding identification of tubificid oligochaetes and mollusks prior to the next season.

Table 7-2. Results of recounts performed by the laboratory processing benthic infauna samples in 1991. Approximately 10% of all samples were processed in duplicate.

Measurement	Mean Error	Range of Error
Benthic sorting	4.5%	0 - 20.5%
Species identification and enumeration	2.4%	0 - 14%
Biomass	0.13%	0 - 1.6%
Weighing blanks for biomass	0.0001g	0 - 0.0023g

7.6 1992 QA Results

Two QA steps were required by the EMAP-VP 1992 QA Project Plan: in-house QC checks (i.e., resorts, recounts, and ID confirmation) on 10% of each technician's work, and independent verification of species identification. The recounts (multiple types - see Table 7-3) and preliminary species verification were performed by the laboratory performing the analyses. Most of these met the requirements established in the QA Plan. Both of the laboratories performing these analyzes were evaluated by independent laboratories in 1990 or 1991; therefore, the use of such an independent evaluation in 1992 was deemed unnecessary.

Table 7-3. Results of recounts performed by the laboratory processing benthic infauna samples in 1992. Approximately 10% of all samples were processed in duplicate.

Measurement	Mean Error	Range of Error
Benthic sorting	1.7%	0 - 18%
Species identification and enumeration	1.8%	0 - 12%
Biomass	1.2%	0 - 1.4%
Weighing blanks for biomass	7×10^{-5} g	0 - 7×10^{-4} g

7.7 1993 QA Results

Two QA steps were required by the EMAP-VP 1993 QA Project Plan: in-house QC checks (i.e., resorts, recounts, and ID confirmation) on 10% of each technician's work, and independent verification of species identification. The recounts (multiple types - see Table 7-4) and preliminary species verification were performed by the laboratory performing the analyses. Most of these met the requirements established in the QA Plan. Both of the laboratories performing these analyzes were evaluated by independent laboratories in 1990 or 1991; therefore, the use of such an independent evaluation in 1993 was deemed unnecessary.

Table 7-4. Results of recounts performed by the laboratory processing benthic infauna samples in 1993. Approximately 10% of all samples were processed in duplicate.

Measurement	Mean Error	Range of Error
Benthic sorting	2.9%	0 - 8.9%
Species identification and enumeration	0.75%	0 - 6.7%
Biomass	0.07%	0 - 0.8%
Weighing blanks for biomass	1.1×10^{-4} g	0 - 9×10^{-4} g

Section 8
QA Results for Fish Community Structure and Pathology

8.1 Background

At each Base Sampling Site crews conducted a standard trawl (10 ± 2 minutes at 2-3 knots speed over bottom) to collect fish for community structure analysis. Fish were identified, measured, counted, examined for the presence of selected external pathologies, and selected individuals of a set of 10 target species processed for chemical residue analysis. As part of EMAP-VP's QA program, in 1990 and 1991 the first individual of every species collected by each crew was preserved in formalin and sent into the laboratory for verification by an expert taxonomist. In 1992 and 1993 crews were instructed to save the first two individuals of each species collected. Fisheries experts within EMAP and the National Marine Fisheries Service (NMFS) were employed in 1990 to 1993, and on two occasions when identification was difficult, specimens were sent to outside experts. Preserved fish were archived for use during training in subsequent years..

To verify each crew's ability to properly identify pathologies, fish identified as having an external pathology by the field crews were shipped to ERL-Gulf Breeze (1990) or ERL-Narragansett (1991, 1992, and 1993) for verification by the laboratory's pathologist. It is important to note that this verification in1990 to 1992 was "blind" (*i.e.*, the pathologist did not know which fish the field crews believed to have a pathology). This provided an estimate of the percentage of "false positives". In addition, in order to develop an estimate of the rate of "false negatives" (*i.e.*, number of pathologies missed, therefore never sent in for verification), crews collected andshipped up to 25 individuals of each target species and 10 from any other species (which they determined to be free from external pathologies) caught at Indicator Testing and Evaluation stations. These steps were necessary because in 1990 through 1992 fish were also collected for chemical residue analysis, which took priority over pathology QA. (Note: only fish collected in 1991 were actually analyzed for residues). Because of this, a fish observed by the crew to have a pathology may have been sent in for chemical analysis rather than pathology verification. Therefore, the assessment produced by EMAP on the prevalence of gross external pathologies in fish is based on field observations, not laboratory observations. An error rate is then associated with these data based on the results of the QA review. Because of poor agreement between field and laboratory examinations, this protocol changed in 1993 (described in Section 8.7).

Following a review of the 1990 and 1991 pathology QA data, and in consultation with experts from NMFS, EMAP-VP elected to condense field observations for fish pathologies to four basic categories: lumps, growths, ulcerations, and fin erosion. It was hoped that by making the examination more simple the success rate (*i.e.*, proper identification) would increase.

8.2 Audits

No laboratory audits were conducted for these indicators. Field performance reviews and audits were conducted as described in Section 2. The QA Coordinator or Field Coordinator visited each crew both during trial runs and the field season. One activity observed by the reviewer was the measurement process, with the reviewer remeasuring selected fish. The reviewer also observed and checked the examination for pathologies conducted by the crew.

8.3 Data Qualifier Codes for Fish Community Structure and Pathology Data

No QA codes currently exist for fish community structure data, indicating all data meet QA criteria and are suitable for EMAP assessment purposes. Codes do exist for pathology data. These codes pertain to whether the pathology was verified by an expert pathologist, and are listed in Table 8-1. These codes were used to provide estimates of the percentage of false positives (crew identified a pathology which was not verified in the laboratory) and false negatives (pathologist identified a pathology on a "reference" fish which, by definition, the crew believed to be pathology-free). These codes only pertain to the four "EMAP" pathologies crews focused on beginning in 1992: lumps, growths, ulcers and fin erosion. These are a subset of the list of pathologies targeted in 1990 and 1991; therefore, the codes were applied to 1990/1991 data as well.

8.4 1990 QA Results

To verify each crew's ability to correctly identify fish species for the community structure indicator, the first individual of each species collected by each crew was shipped to ERL-N or Versar for verification by an expert taxonomist.

Three types of errors were detected: misspelled or incomplete species names (in the database), misidentifications, and fish that could not be identified in the field. Errors falling into the first category were easily detected, corrected in the database, and documented. An example of this type of error can be found looking at the "Atlantic tomcod". Records were received from the field for "Atlantic tomcod", "tomcod", and "tom cod" (two words). Each was listed by the computer as separate species.

The second type of error was mis-identifications. Of the 136 fish sent in for taxonomic verification, nine were misidentified, representing seven species. In all cases the crew identified a closely-related species, such as longspine porgy instead of scup, brown bullhead catfish instead of the yellow bullhead, and lizardfish instead of inshore lizardfish. An additional 16 individuals (12 species) were sent in as unknowns or partial unknowns (*e.g.*, herring uncl.).

All errors were corrected in the database. If a QA fish was misidentified by the crew, all other fish **in the same size class** of that species from the same trawl were changed to the correct ID.

Results of laboratory pathology examinations reveal that the crews were generally conservative, classifying "borderline" conditions as pathologies so the fish would be examined by an expert rather than being discarded. Table 8-2 presents results of the laboratory review for the four final pathology categories EMAP-VP selected for continued use.

Table 8-1. Qualifier codes for fish pathology data

Code	Description
FP-A	LUMPS NOT Observed in Field / NOT Observed by Quality Assurance Laboratory
FP-B	LUMPS NOT Observed in Field but was Observed by Quality Assurance Laboratory
FP-C	LUMPS Observed in Field but NOT Observed by Quality Assurance Laboratory
FP-D	LUMPS Observed in Field and Confirmed by Quality Assurance Laboratory
FP-E	LUMPS NOT Observed in Field but NOT Looked for by Quality Assurance Lab.
FP-F	LUMPS Observed in Field but NOT Looked for by Quality Assurance Laboratory
FP-G	GROWTHS NOT Observed in Field / NOT Observed by Quality Assurance Laboratory
FP-H	GROWTHS NOT Observed in Field but was Observed by Quality Assurance Laboratory
FP-I	GROWTHS Observed in Field but NOT Observed by Quality Assurance Laboratory
FP-J	GROWTHS Observed in Field and Confirmed by Quality Assurance Laboratory
FP-K	GROWTHS NOT Observed in Field but NOT Looked for by Quality Assurance Lab.
FP-L	GROWTHS Observed in Field but NOT Looked for by Quality Assurance Laboratory
FP-M	ULCERS NOT Observed in Field / NOT Observed by Quality Assurance Laboratory
FP-N	ULCERS NOT Observed in Field but was Observed by Quality Assurance Laboratory
FP-O	ULCERS Observed in Field but NOT Observed by Quality Assurance Laboratory
FP-P	ULCERS Observed in Field and Confirmed by Quality Assurance Laboratory
FP-Q	ULCERS NOT Observed in Field but NOT Looked for by Quality Assurance Lab.
FP-R	ULCERS Observed in Field but NOT Looked for by Quality Assurance Laboratory
FP-S	FINROT NOT Observed in Field / NOT Observed by Quality Assurance Laboratory
FP-T	FINROT NOT Observed in Field but was Observed by Quality Assurance Laboratory
FP-U	FINROT Observed in Field but NOT Observed by Quality Assurance Laboratory
FP-V	FINROT Observed in Field and Confirmed by Quality Assurance Laboratory
FP-W	FINROT NOT Observed in Field but NOT Looked for by Quality Assurance Lab.
FP-X	FINROT Observed in Field but NOT Looked for by Quality Assurance Laboratory
FP-Y	Fish Not Examined for Gross External Pathology in the Field

Table 8-2. 1990 Pathology QA results based on laboratory examination of fish crews believed to have a pathology and reference, "pathology-free" fish (n=769).

Pathology Type	False Positives[1]	False Negatives[2]
Body Ulcerations	9/20 (45.0%)	8/749 (1.1%)
Body Lumps/Growths	3/12 (25.0%)	26/757 (3.4%)
Fin Erosion	8/17 (47.1%)	16/752 (2.1%)

[1] False Positives: The denominator in this column is the total number of fish identified by the field crews as having a given pathology. The numerator is the number of these fish for which the pathology was not confirmed by the pathologist.

[2] False Negatives: The denominator in this column is the total number of fish identified by the field crews as not having a given pathology. The numerator is the number of these fish for which the pathology was observed by the pathologist.

8.5 1991 QA Results

To verify each crew's ability to correctly identify fish species for the community structure indicator, the first individual of each species collected by each crew was shipped to ERL-N or Versar for verification by an expert taxonomist. Three types of errors were detected: misspelled or incomplete species names (in the database), misidentifications, and fish that could not be identified in the field. Errors falling into the first category were easily detected, corrected in the database, and documented. An example of this type of error can be found looking at the "Atlantic tomcod". Records were received from the field for "Atlantic tomcod", "tomcod", and "tom cod" (two words). Each was listed by the computer as separate species.

Of the 187 fish sent in for taxonomic verification, 14 were misidentified, representing nine species. In all cases the crew identified a closely-related species, such as longspine porgy instead of scup, brown bullhead catfish instead of the yellow bullhead, and lizardfish instead of inshore lizardfish. An additional 14 individuals (five species) were sent in as unknowns or partial unknowns (*e.g.*, herring uncl.).

The total of 28 incomplete identifications or misidentifications represent 51 fish records in the database (including other fish of the same species caught in the same trawl). A total of 7,134 fish were collected in standard trawls during the 1991 field season representing 69 species. The percentage of errors detected was therefore less than one percent. All errors were corrected in the database. If a QA fish was misidentified by the crew, all other fish **in the same size class** of that species from the same trawl were changed to the correct ID.

Results of laboratory pathology examinations reveal that the crews were generally conservative, classifying "borderline" conditions as pathologies so the fish would be examined by an expert rather than being discarded. Of the six fish sent in for verification of a pathology (four additional fish were not shipped), only three were verified by the pathologist. Of the "reference" fish shipped, the pathologist determined that none had a pathology. Fin erosion was not included in these statistics as damage was incurred due to the method of shipping fish (packaged in mesh onion bags) prohibiting accurate examinations by the laboratory staff. These results are for all types of pathologies. Table 8-3 presents results of the laboratory review for only the four final pathology categories EMAP-VP selected for continued use.

Table 8-3. 1991 Pathology QA results based on laboratory examination of fish crews believed to have a pathology and reference, "pathology-free" fish (n=195).

Pathology Type	False Positives[1]	False Negatives[2]
Body Ulcerations	2/5 (40.0%)	0/190 (0.0%)
Body Lumps/Growths	1/1 (100.0%)	0/194 (0.0%)
Fin Erosion	not available	not available

[1] False Positives: The denominator in this column is the total number of fish identified by the field crews as having a given pathology. The numerator is the number of these fish for which the pathology was not confirmed by the pathologist.

[2] False Negatives: The denominator in this column is the total number of fish identified by the field crews as not having a given pathology. The numerator is the number of these fish for which the pathology was observed by the pathologist.

8.6 1992 QA Results

To verify each crew's ability to correctly identify fish species for the community structure indicator, the first two individuals of each species collected by each crew was shipped to ERL-N for verification by an expert taxonomist. Three types of errors were detected: misspelled or incomplete species names (in the database), misidentifications, and fish that could not be identified in the field. Errors falling into the first category were easily detected, corrected in the database, and documented.

Of the 397 fish sent in for taxonomic verification, 36 were misidentified. In all cases the crew identified a closely-related species, such as longspine porgy instead of scup, or brown bullhead catfish instead of the yellow bullhead. An additional eight individuals were sent in as unknowns or partial unknowns (*e.g.*, herring uncl.). Most mis-identified or partially identified individuals were juveniles.

The total of 44 incomplete identifications or misidentifications represent 116 fish records in the database (including other fish of the same species caught in the same trawl). A total of 14,704 fish were collected in all trawls (both standard and non-standard) from all station types during the 1992 field season representing 78 species. The percentage of errors detected was therefore less than one percent. All errors were corrected in the database. If a QA fish was misidentified by the crew, all other fish **in the same size class** of that species from the same trawl were changed to the correct ID.

Results of the pathologist's review of fish collected by field crews in 1992 are illustrated in Table 8-4. Crews appeared to be overly conservative, classifying fish as having a pathology when, in fact, they did not in almost all cases. It is also possible that by requiring the pathologist to blindly examine hundreds of fish, some of the few with a pathology might be missed. Onion bags were no longer used for containing fish starting in 1992. This was the cause of the damage in 1991 which prevented verification of fin erosion; therefore, QA data on fin erosion are included in Table 8-4.

Table 8-4. 1992 Pathology QA results based on laboratory examination of fish crews believed to have a pathology and reference, "pathology-free" fish (n=427).

Pathology Type	False Positives[1]	False Negatives[2]
Body Ulcerations	9/9 (100.0%)	1/418 (0.2%)
Body Lumps/Growths	3/4 (75.0%)	0/423 (0.0%)
Fin Erosion	5/5 (100.0%)	1/422 (0.2%)

[1] False Positives: The denominator in this column is the total number of fish identified by the field crews as having a given pathology. The numerator is the number of these fish for which the pathology was not confirmed by the pathologist.

[2] False Negatives: The denominator in this column is the total number of fish identified by the field crews as not having a given pathology. The numerator is the number of these fish for which the pathology was observed by the pathologist.

8.7 1993 QA Results

As a result of the 1990-1992 data, and the fact that chemistry fish were no longer to be collected, the QA process for pathology data changed after the 1992 field season. Starting in 1993, the results on the prevalence of pathologies in fish of the Virginian Province are based on the laboratory examination, NOT the field exam. Crews were instructed to examine all fish and ship every one suspected of having a pathology to the laboratory for confirmation. In 1993, the examination by the pathologist was no longer "blind". Fish received at the laboratory were coded as "pathology" or "reference" fish. If the pathologist disagreed with the crew's observation (*i.e.*, he felt a pathology fish did not have a pathology or a reference fish was found to have one), a second pathologist was consulted and their collective decision entered into the database. Although data from 1990 through 1992 show the crews to be efficient at not missing many pathologies (*i.e.*, low incidence of false negatives), the pathologist's review of reference fish continued. The results of the laboratory examinations are presented in Table 8-5. The high rate of "false positives" is likely the result of the crews being overly conservative following instruction to ship any fish SUSPECTED of having a pathology.

Table 8-5. 1993 Pathology QA results based on laboratory examination of fish crews believed to have a pathology and reference, "pathology-free" fish (n=620).

Pathology Type	False Positives[1]	False Negatives[2]
Body Ulcerations	10/12 (83.3%)	1/608 (0.2%)
Body Lumps	5/5 (100.0%)	0/615 (0.0%)
Body Growths	4/11 (36.4%)	2/609 (0.3%)
Fin Erosion	1/4 (25.0%)	0/616 (0.0%)

[1] False Positives: The denominator in this column is the total number of fish identified by the field crews as having a given pathology. The numerator is the number of these fish for which the pathology was not confirmed by the pathologist.

[2] False Negatives: The denominator in this column is the total number of fish identified by the field crews as not having a given pathology. The numerator is the number of these fish for which the pathology was observed by the pathologist.

8.8 Lessons Learned

In 1990-1992 samples were sent in "blind". In general, samples had been stockpiled until the end of the field season and then examined. Because only a few fish with pathologies were inter-mixed with hundreds of reference fish, it is possible that some true pathologies identified by the crews may have been missed in the laboratory examination. In 1993 fish were no longer sent in to the laboratory as blind samples. Each fish was identified to the analyst as being a fish which the crew believed to have a pathology, or as a pathology-free fish.

The incidence of gross external pathologies reported for 1990 to 1992 is based solely on field operations, with the rate qualified by the reported rates of false positives and false negatives (see Sections8.4 to 8.6). Because the potential existed for a fish with a pathology to be saved for chemical residue analysis instead of laboratory verification of the pathology, the incidence rate could not be reported based on the laboratory results. Therefore, the incidence rates are based on the QA codes beginning with "observed in the field" regardless of whether or not the pathology was observed in the laboratory. The incidence reported for 1993 was based on the laboratory results and need not be qualified.

It should be noted that the incidence rates for all four years are similar when using uncorrected results (*i.e.*, 1990 to 1992 results are not adjusted based on the rates of false positives or negatives).

Section 9
QA Results for Water Quality Measurements

9.1 Background

During the four years of EMAP monitoring in the Virginian Province two different approaches have been utilized to characterize water quality: 1) point-in-time water column profiling using a CTD (Conductivity, Temperature, Depth logger), and 2) continuous, long-term near-bottom measurements using a moored datalogger. The Seabird SBE 25 Sealogger CTD has been used to obtain vertical profiles of temperature, salinity, dissolved oxygen (DO), pH, light transmission, fluorescence and photosynthetically active radiation (PAR). The Hydrolab DataSonde3 datalogger has been used to record long-term time series of temperature, salinity, dissolved oxygen and pH in the near-bottom waters. During the 1990 Demonstration Project, CTD casts were conducted at all station classes once per sampling interval, while DataSonde3 instruments were repeatedly deployed for approximately 10-day durations at 23 long-term dissolved oxygen (LTDO) stations throughout sampling intervals 1 and 2. The Seabird CTDs have continued to be used in each subsequent year of sampling in the Virginian Province. In 1991, the deployment interval for the DataSonde3's was shortened to one to three days, and units were deployed at all stations. Prior to the 1992 season, the decision was made to cease deploying these instruments.

The Field Methods Manuals prepared for each year of Virginian Province sampling provide detailed descriptions of the field protocols for use of the various water quality instruments. The following sections describe the QA/QC protocols that have evolved and the subsequent QA results achieved for each year of Virginian Province water quality monitoring over the period 1990-1993.

9.2 1990 Calibration and Calibration Check Procedures

Seabird CTD

In 1990, the Seabird CTDs were calibrated prior to sampling and throughout the field season as needed (Table 9-1). The dissolved oxygen calibrations were checked against Winkler titrations and/or saturation table values and the pH calibrations checked against standard pH buffer solutions. Field QC checks of the CTD temperature, conductivity (salinity), dissolved oxygen, and pH sensors were conducted daily. For these checks, real-time CTD readings from just below the surface were compared to sample measurements taken with a mercury thermometer, refractometer, and Winkler titrations from a water sample collected with a *Go-Flo* water sampling bottle. It should be noted that the use of a refractometer for verifying the CTD's salinity sensor simply serves as a crude check to determine if the sensor has suffered an electronic problem resulting in gross errors. The pH readings were checked using a pH 7 standard buffer solution. If any of the parameters did not fall within the acceptable QC limits (Table 9-2), the instrument was checked and, if necessary, recalibrated.

For deployment of the CTD, the instrument was first turned on while on deck then lowered to just below the surface and allowed to equilibrate for two minutes. The unit was then lowered slowly to one meter above the bottom and again allowed to equilibrate for two minutes. The vessels were not equipped with a meter wheel; therefore, at times too much cable was paid out and the CTD came in contact with the seafloor. When this occurred, the CTD was immediately brought up to one meter off the bottom. After the two-minute bottom soak, the unit was hauled back on deck. Electronic files containing the CTD cast data were usually downloaded to the on-board computer while the vessel was anchored on station. The field crew quickly reviewed the temperature, salinity and dissolved oxygen data. On a number of occasions problems with the on-board computer prevented the field crews from downloading and reviewing the CTD cast data.

Table 9-1. Summary of calibration procedures used for Virginian Province water quality instruments in 1990.

Instrument	Sensor	Calibration Procedure
Seabird SBE 25 CTD	Temperature	Calibrated by manufacturer prior to sampling season
	Conductivity	Calibrated by manufacturer prior to sampling season
	DO	Two point (zero & air-saturated water)
	pH	Three point (pH 4, 7 & 10 std. solns.)
	LightTrans.	Calibrated by manufacturer prior to sampling season
	Fluorescence	Calibrated by manufacturer prior to sampling season
	PAR	Calibrated by manufacturer prior to sampling season
	Pressure	Calibrated by manufacturer prior to sampling season
Hydrolab DataSonde 3	Temperature	Calibrated by manufacturer prior to sampling season
	Salinity	0.5 M potassium chloride (KCl) solution
	DO	Water-saturated air
	pH	Two point (pH 7 & 10 std. solns.)
	Depth	Zeroed at water's surface (sealevel)

Table 9-2. Field calibration checks performed during the 1990 Virginian Province Demonstration Project.

Instrument	Frequency of check	Parameter	Checked against	Maximum acceptable difference
Seabird SBE 25 CTD	Daily	Temperature	Thermometer	± 2 °C
		Salinity	Refractometer	± 2 ppt
		DO	Winkler titration	± 1.0 mg/L
		pH	pH 7 std. solution	± 0.5 pH units
Hydrolab Datasonde 3	Pre- and post-deployment (each use)	Temperature	Thermometer	± 2 °C
		Conductivity	0.5 M KCl std.	± 5 mS/cm
		DO	Water-saturated air	± 12.5% saturation
		pH	pH 7 std. solution	± 0.5 units

Hydrolab DataSonde3

The DataSonde3 units were calibrated prior to each deployment using the manufacturer's recommended procedures (Table 9-1). QC checks were conducted at the dock on the morning that the instruments were to be deployed or onboard the vessel just prior to deployment. The QC checks procedures were similar to the calibrations: dissolved oxygen percent saturation was compared to expected readings of 102.5% in water-saturated air (102.5% is used instead of 100% saturation because Hydrolab's Lo-Flo membrane was installed on the instruments), specific conductivity was compared to a standard reading using a 0.5 M KCl solution, pH was compared to pH 7 standard buffer solution and temperature was compared to thermometer readings. If any of the parameters did not fall within acceptable QC limits (Table 9-2), the crews re-calibrated the sensor prior to deployment.

Individual units, housed inside a protective PVC casing were moored approximately one meter above the bottom. They were programmed to record data internally at 30-minute intervals throughout their ten-day deployments. Upon retrieval, the units were examined for evidence of biological fouling of the probes. They underwent a post-deployment QC check that was identical to the pre-deployment calibration check. The data files were downloaded either on board the vessel or in the mobile laboratory. The raw data files were quickly reviewed, paying particular attention to the dissolved oxygen values. If the dissolved oxygen dropped to zero at any time during the record, no replacement unit was deployed at that station for fear of "poisoning" the DO probe.

9.3 Data Qualifier Codes for Water Quality

Because of the number of parameters monitored and the overall complexity of the water quality datasets, a rather large number of data qualifier codes, or "flags", are incorporated into these datasets. These codes are listed in Tables 9-3 and 9-4. In order for the codes to make sense, it is important to understand the data manipulations employed in the analysis stage.

CTD Qualifier Codes

The first step in assessing the quality of CTD profiles was for the crew chief to examine the profile on the on-board field computer as soon as the data were collected. As described in Section 9.4, problems were encountered in 1990 with this procedure. Upon receipt of the electronic file in the Information Management Center, the first step is verification. An analyst examined each cast to determine if it was associated with the correct station. The CTD depth was compared to the depth recorded from the boat's fathometer, and individual measurements were compared to those expected (*e.g.*, low salinity would not be expected from a station in Long Island Sound) or measured via other mechanisms (*e.g.*, from the Hydrolab or ambient checks).

The next step in the data assessment process is validation. Each CTD file consists of an entire profile of measurements made through the water column from the surface to the bottom and back up to the surface. For ease of analysis, each CTD file was split into separate components which were stored as individual files. Upon submersion, the CTD was allowed to sit approximately one meter below the surface for several minutes to allow it to come to thermal equilibrium after being on the hot deck. The section of the profile from the point of immersion until the unit is lowered through the water column is the "surface soak". Data from this file were not used other than to ensure that the crew allowed sufficient time for equilibration. The section of the profile starting when the unit is lowered and ending when it reaches the bottom (actually one meter off the bottom) is the "downcast". The unit was then allowed to sit at the bottom and record data for several minutes. This is the "bottom" file. Finally, the point from the start of the unit's ascent until it reaches the surface is the "upcast".

Table 9-3. Data qualifier codes attached to 1990 - 1993 CTD water quality data. Only C-A through C-H were applied to the 1990 data.

CODE	DESCRIPTION
C-A	Reject entire CTD cast (all parameters).
C-B	Accept entire CTD cast (all parameters).
C-C	Bottom file acceptable; downcast file rejected; no surface values; reported bottom values are the average of all bottom records.
C-D	Downcast file acceptable; bottom file rejected; first and last records of downcast file used for reported surface and bottom values, respectively.
C-E	Downcast and bottom files rejected; however, first and last records of downcast file appeared reasonable and were used for surface and bottom values, respectively.
C-F	Downcast file acceptable; bottom file rejected; reported surface values are the first record of the downcast file; bottom values are the first record of the bottom file (appeared acceptable).
C-G	Downcast and bottom files rejected; reported surface values are the first record of the downcast file (appeared reasonable). No bottom values reported.
C-H	Downcast and bottom files rejected; bottom values are the last record of downcast file (appeared reasonable). No surface values reported.
C-IA	Reject surface values (all parameters)
C-IB	Reject pre-deploy. soak, accept post-deployment soak (all parameters)
C-IC	Reject entire bottom soak, no bottom values available (all parameters)
C-ID	Reject entire downcast file (all parameters)
C-IE	Reject bottom soak, use last value of downcast (all parameters)
C-IF	Reject average of bottom soak but accept last value (all parameters)
C-IG	Shallow station with pre-deployment soak and bottom soak only (no profile)
C-IH	Shallow station: surface and bottom values equal. Bottom file used for both.
C-II	Depth values questionable
C-IJ	Reject surface dissolved oxygen (pre and post)

(continued)

Table 9-3. continued.

CODE	DESCRIPTION
C-IK	Reject pre cast dissolved oxygen but accept post cast dissolved oxygen
C-IL	Reject downcast dissolved oxygen
C-IM	Reject bottom dissolved oxygen
C-IN	Reject bottom soak dissolved oxygen but use last value of downcast
C-IO	Reject average bottom dissolved oxygen but use last value of bottom file
C-IP	Reject surface salinity (pre and post)
C-IQ	Reject pre cast salinity but accept post cast salinity
C-IR	Reject downcast salinity
C-IS	Reject bottom salinity
C-IT	Reject bottom soak salinity but use last value of downcast
C-IU	Reject average bottom salinity but use last record of bottom file
C-IV	Reject surface temperature (pre and post cast)
C-IW	Reject pre cast temperature but accept post cast temperature
C-IX	Reject downcast temperature
C-IY	Reject bottom temperature
C-IZ	Reject bottom soak temperature but use last value of downcast
C-JA	Reject average bottom temperature but use last value of bottom file
C-JB	Reject surface pH (pre and post)
C-JC	Reject pre cast pH but accept post cast pH
C-JD	Reject downcast pH
C-JE	Reject bottom pH
C-JF	Reject bottom soak pH but use last value of downcast file

(continued)

Table 9-3. continued.

CODE	DESCRIPTION
C-JG	Reject average bottom pH but use last value of bottom file
C-JH	Reject surface PAR (pre and post soak)
C-JI	Reject pre cast PAR but accept post cast PAR
C-JJ	Reject downcast PAR
C-JK	Reject bottom PAR
C-JL	Reject bottom soak PAR but use last value of downcast
C-JM	Reject average bottom PAR but use last value of bottom file
C-JN	Reject surface transmissometry (pre and post)
C-JO	Reject pre cast transmissometry but accept post cast transmissometry
C-JP	Reject downcast transmissometry
C-JQ	Reject bottom transmissometry
C-JR	Reject bottom soak transmissometry but use last value of downcast
C-JS	Reject average bottom transmissometry but use last value of bottom file
C-JT	Reject surface fluorescence (pre and post)
C-JU	Reject pre cast fluorescence but accept post cast fluorescence
C-JV	Reject downcast fluorescence
C-JW	Reject bottom fluorescence
C-JX	Reject bottom soak fluorescence but use last value of downcast
C-JY	Reject average bottom fluorescence but use last value of bottom file
C-JZ	Fluorescence off-scale

Each component of the profile (*i.e.*, each file) was examined to determine if it was reasonable. In 1990 only the dissolved oxygen (DO) records were examined. Files were classified as acceptable or not solely based on the DO record. In subsequent years each parameter was assessed independently, resulting in a significant increase in the number of codes. Components of the profile could be classified as unacceptable for a number of reasons. For example, the downcast could contain unexplained spikes (it should be relatively smooth), or the shape of the bottom soak might suggest the unit impacted the bottom and mud was sucked up into the pumping system, clogging it. Also, the upcast may not match the downcast.

As part of EMAP's data assessment activities, "surface" and "bottom" values are reported for key parameters such as DO and salinity. These values were extracted from the profile. In general, the surface value is the first record of the downcast, and the bottom value is the average of all values in the bottom file. As shown in the CTD qualifier codes listed in Table 9-3, when a section of the profile is determined to be unacceptable, other alternatives are employed. If the bottom soak is determined to be unacceptable, the last record of the downcast is generally used as the reported bottom value. If the appearance of the bottom section of the profile suggests mud clogged the pump, and this clog cleared itself during the bottom soak; or if there was a significant near-bottom oxycline resulting in a change in DO between the start and end of the bottom soak, the last value of the bottom file is used in place of the average. Similarly, the downcast file might be deemed unacceptable because of severe spiking in the middle of the downcast, suggesting a temporary clog or an intermittent electronic problem. However, both the beginning and end of this file may appear reasonable. In such cases the file may be classified as unacceptable but the first record is still used for the surface value, and, if the bottom file is unacceptable, the last record may be used for the bottom value.

These codes generally are not of interest to users requesting the summary data, *i.e.*, reported surface and bottom values. The flags simply point to documentation on how those values were determined. The codes are likely of greater importance to users requesting the actual profiles as they point to potential problems with those data.

Hydrolab Qualifier Codes

Codes describing Hydrolab files are listed in Table 9-4. As described in Section 9.2, QC checks of individual units were performed both before deployment and following retrieval. The qualifier codes indicate the results of those checks. For example the code "H-K" means that the DO at the end of the file may underestimate the actual DO concentration. This would indicate that the Hydrolab unit failed QC upon retrieval, likely due to fouling of the DO sensor. Fouling is a gradual process, making it difficult to determine at what point during thedeployment the readings become unreliable.

The code "H-H" requires explanation. Prior to deployment each unit is set to log for a certain period of time at a selected interval. In 1990 when the units were deployed for 10-day periods at selected stations, the units were set to log at 30-minute intervals. In 1991 when they were deployed at every station for one to three days they were set to log at 15-minute intervals. "Autolog" is a back-up which automatically logs data every hour regardless of how the logging run is set up. In the event that the crew accidentally set the unit incorrectly *(e.g.*, set it to start logging on 4/5/96 instead of 4/5/91), Autolog would still log data hourly. As a result, data are collected for the duration of the deployment; however, the logging interval would be different from other files collected that year.

Table 9-4. Data qualifier codes attached to Hydrolab water quality data.

CODE	DESCRIPTION
H-A	No file available.
H-B	Acceptable file.
H-C	Dissolved oxygen not acceptable.
H-D	Salinity not acceptable.
H-E	Temperature not acceptable.
H-F	pH not acceptable.
H-G	Discontinuous record due to power loss.
H-H	Autolog file.
H-I	Total record less than 24 hours.
H-J	DO at start of file may overestimate actual ambient DO concentration.
H-K	DO at end of file may underestimate actual ambient DO concentration.
H-L	DO at start of file may underestimate actual ambient DO concentration.
H-M	DO at end of file may overestimate actual ambient DO concentration.
H-N	Data not available for entire deployment.
H-O	Depth not acceptable.
H-P	Salinity at file end may underestimate actual ambient salinity concentration.
H-Q	pH at start of file may underestimate actual ambient pH value.
H-R	Battery Power not acceptable.

9.4 1990 QA Results

Seabird CTD

The CTD data were affected by several procedural problems that came to light during and after the sampling season. First, the QC checks for both dissolved oxygen and pH did not perform satisfactorily during the Demonstration Project. The field crews were not prepared to identify unacceptable CTD casts in the field, and as a result, many of the casts were later flagged for containing unacceptable data. In addition, many CTD data files were lost in the beginning of the field season as a result of a computer software problem.

Performance of the dissolved oxygen probe was checked by comparing the CTD sensor's reading to that calculated using a digital titrator that was part of a HACH Winkler titration field kit. The results of the Winkler titrations were not as reliable as initially expected: the difference between two replicate dissolved oxygen water samples exceeded 0.5 mg/L in over 11% of the field QC checks conducted throughout the sampling season (Table 9-5). This large amount of variability between replicates made it difficult to assess whether the QC checks were reliable enough to evaluate the performance of the dissolved oxygen sensor. It was unknown whether the 60 CTD QC checks that exhibited differences between the dissolved oxygen sensor and the Winkler titrations in excess of 1 mg/L were a result of faulty sensors or poor QC check (*i.e.*, Winkler titration) procedures.

The field QC check of the pH sensor involved comparing the CTD sensor's reading to a pH 7 standard buffer solution. This procedure was implemented to insure that the sensor's calibration did not drift; however, it proved to be an inappropriate check. Post-sampling-season review of the CTD casts revealed that one of the pH sensors was malfunctioning for most of the summer, a problem that was never detected in the field because upon malfunctioning, the reading defaulted to a constant value of pH 7. This was not detected because the pH 7 buffer solution was used for the check, and crews generally did not review the pH data collected with each profile.

A total of 480 CTD casts were conducted during the 1990 Demonstration Project. Data from 9% of these casts were lost and are not included in the database. The remaining 437 casts were carefully reviewed for acceptability based solely on the performance of the dissolved oxygen sensor (Table 9-6). Of the reviewed casts, 68% had acceptable dissolved oxygen profiles (see Section 9-3) and 23% yielded unacceptable dissolved oxygen profiles, although individual surface and/or bottom values were accepted in some of the casts where the profile itself was rejected. The profiles of other parameters were used to assess the validity of the dissolved oxygen data (*e.g.*, high fluorescence would be expected in areas of supersaturated dissolved oxygen); however, the acceptability of the data recorded by these other sensors was not determined.

Table 9-6 also reports QA results specific to Base Sampling Sites sampled during Interval 2. These are the data that are currently being used in the assessment of the ecological condition of the Virginian Province. Many stations were visited on more than one occasion; however, only one dissolved oxygen value is reported per station. The table shows that for 92% of the stations used in this assessment, acceptable bottom DO values were obtained in 1990.

Table 9-5. Results of CTD dissolved oxygen field QC checks used during the 1990 Demonstration Project. Dissolved oxygen readings from the CTD sensor were compared to HACH Winkler titrations.

# of CTD field QC checks	174
# of Winkler replicates w/ DO differences > 0.5 mg/L	20 (11%)
# of CTD/Winkler QC checks w/ DO differences > 1.0 mg/L[1]	60 (34%)
Range of CTD/Winkler DO differences (mg/L)	-4.7 to +4.0

[1] It was unknown whether the 60 CTD QC checks that exhibited differences between the dissolved oxygen sensor and the Winkler titrations in excess of 1 mg/L were a result of faulty sensors or poor QC check (*i.e.*, Winkler titration) procedures.

Table 9-6. Results of 1990 post-sampling season CTD data review. (Percentages are based upon # of total casts attempted and total number Base Stations sampled in Interval 2). Acceptability of casts in 1990 was based solely upon performance of the DO sensor whereas performance of all sensors were considered in subsequent years.

	Total Casts	Base Stations
# Total casts/Base Stations	480	111
# Events w/ lost files	43	3
# Casts accepted for all parameters (C-B)	298 (62%)	98 (88%)
# Casts rejected for all parameters (includes lost casts)	142 (30%)	11 (10%)
# Casts w/ acceptable surface DO	315 (66%)	99 (89%)
# Casts w/ acceptable bottom DO	337 (70%)	102 (92%)

Hydrolab DataSonde3

Evaluation of the Hydrolab QA activities during the Demonstration Project revealed several concerns regarding field procedures. The calibration of the conductivity cell, for measuring salinity, produced unsatisfactory results. In addition, the dissolved oxygen QC check was deemed to be inappropriate. The ten-day deployment period often resulted in extensive biological fouling of the probes and unacceptable dissolved oxygen records.

Calibration of the specific conductivity parameter resulted in inaccurate salinity measurements. It was determined that the 0.5 M KCl solution was not standardized properly and thus did not have the assumed standard reading of 58.64 mS/cm. The actual conductivity of the standard utilized was determined using an Auto-Salinometer calibrated to Copenhagen seawater. All of the raw data files were then re-processed using a sub-routine that compensated for the incorrect calibrations.

The dissolved oxygen QC check using water-saturated air was not appropriate since the sensor membrane had to be wiped dry during the process. This may have had a large effect on the retrieval QC checks since this procedure removes biological and physical fouling that may have altered the sensor's performance during the deployment. A substantial amount of biological fouling appeared on the instrument casing and probes during the ten-day deployments. This fouling may have been responsible for underestimating dissolved oxygen values (towards the end of the datafile) in over 60% of the DataSonde3 records (Table 9-7).

DataSonde3 units were deployed a total of 123 times during the 1990 Demonstration Project. Data from 18% of these deployments are not part of the database because of lost units, incorrect data logging setups or missing datafiles (H-A code). The results of post-retrieval QC checks are summarized in Table 9-7. Of the 104 files which were reviewed, only 19% were totally acceptable for all parameters throughout the entire record. While failures of the conductivity, temperature and pH sensors occurred sporadically, this low overall percentage is due mainly to fouling which resulted in poor performance of the dissolved oxygen sensor during the later half of the records, as discussed earlier.

Table 9-7. Results of calibration checks following retrieval of Hydrolab Datasonde 3 instruments for 1990 Virginian Province monitoring. Percentages indicate the number of times that acceptance criteria were met.

Parameter	Acceptance Criteria	Percent Accepted
Temperature	± 1 °C	99% (103/104)
Salinity	± 2 ppt	96% (100/104)
DO	± 0.5 mg/L	38% (40/104)
pH	± 0.5 units	97% (101/104)

9.5 1990 Lessons Learned/Changes for 1991

Seabird CTD

The post-season evaluation of CTD casts led to a restructuring of the quality control criteria for these instruments. During the 1990 Demonstration Project, the field crews were expected to re-calibrate the CTD sensors in the field. This proved to be a difficult task, particularly for the dissolved oxygen sensor calibration which requires the CTD to be placed in a large tank of air-saturated water. The field crews were constantly on the move and rarely had the opportunity to set up a proper calibration tank. In short, the 1990 experience served to demonstrate that accurate calibration of the dissolved oxygen sensor requires a controlled environment and experienced personnel. Procedural changes implemented in 1991 required the field crews to send back to the field operations center any CTD that failed a calibration QC check. At the field operation centers, trained technicians were on hand to perform a more complete evaluation and, if necessary, re-calibration of any malfunctioning instruments under controlled laboratory conditions.

The daily QC checks conducted during the 1990 Demonstration Project helped to identify sensor drift and the need for re-calibration; however, they could not be used to determine if the instrument performed properly during a specific CTD cast. Field QC checks were changed to include two components: QC checks on the sensor calibrations and QC checks on each deployment. Much investigation went into determining the most appropriate methods for conducting these tests, and the most important findings are summarized below.

Very little sensor calibration drift was observed in the CTDs throughout the 1990 field season; therefore, it was determined that weekly calibration checks (outlined in Table 9-8) would be sufficient for the 1991 field season. The criterion for acceptance of DO data was re-evaluated and the acceptable difference during QA checks was reduced from 1.0 to 0.5 mg/L. The dissolved oxygen and pH QC checks used during the Demonstration Project yielded unsatisfactory results and had to be modified for the 1991 field season.

The dissolved oxygen sensor needed to be compared to a reliable dissolved oxygen value. Winkler titrations were the first choice; however, they had produced unacceptable results in 1990. The performance of the HACH kits were evaluated in the laboratory. These tests identified three faulty titrators that showed excessive variability in amount of titrant delivered when running replicate samples. Further testing demonstrated that the Hach kits could accurately measure dissolved oxygen concentrations under the following conditions: use of properly calibrated titrator, daily standardization of sodium thiosulfate prior to titrating samples, and properly trained technicians who were familiar with conducting Winkler titrations.

Table 9-8. Summary of water quality instrument field calibration checks for 1991-93 Virginian Province monitoring.

Instrument	Frequency of check	Parameter	Checked against	Maximum acceptable difference
Seabird SBE 25 CTD	Each station	Temperature Salinity DO	Thermometer Refractometer YSI DO meter	± 2 °C ± 2 ppt ± 0.5 mg/L
Seabird SBE 25 CTD	Once each week (in concert with YSI check)	Temperature Salinity DO pH	Thermometer Refractometer YSI DO meter pH buffer solution	± 2 °C ± 2 ppt ± 0.5 mg/L ± 0.5 units
Hydrolab Datasonde 3	Pre- and post-deployment (each use)	Temperature Salinity DO pH	Thermometer Refractometer YSI DO meter pH buffer solution	± 1 °C ± 2 ppt ± 0.5 mg/L ± 0.5 units
YSI Model 57 DO meter	Once each week	Temperature DO	Thermometer Winkler titration	± 1 °C ± 0.5 mg/L

This last point, that Winkler titrations require experienced personnel in order to produce accurate results, was a great concern. The field crews are only exposed to a minimal amount of training in many different topics prior to sampling and once in the field, have many demands placed upon them. It was unrealistic to depend upon all of the crew members having the needed experience and available time to conduct daily titrations. Assorted instruments were evaluated and it was determined that the hand-held dissolved oxygen meter manufactured by Yellow Springs Instruments (YSI) provided reliable DO measurements that could be used in the QC checks. However, weekly titrations were still performed, but only by selected individuals who were provided with additional training. Also, beginning in 1991, field crews utilized a potassium iodide/iodate solution to determine the true normality of the thiosulfate solution prior to the weekly QC check of the YSI instrument.

A post-season check of all CTDs revealed that a faulty pH sensor went undetected throughout much of the field season. Consultation with Seabird electronics identified the fact that broken pH sensors will default to a reading of 7; therefore, the field QC check did not detect the broken sensor. Since the crews only reviewed temperature, salinity and dissolved oxygen data in the field, they did not identify the faulty sensor during their normal sampling routine. The field QC check was modified for the 1991 season to require comparing the pH sensor's reading to a standard pH 10 solution. Additionally, the field computer system was modified to include vertical profiles of all parameters, including pH, to be reviewed when the data were downloaded.

Review of the CTD casts obtained during the Demonstration Project revealed several deployment problems that affected the performance of the dissolved oxygen sensor. The most commonly encountered problems were: 1) air bubbles trapped in the dissolved oxygen plumbing loop, 2) mud being sucked through the conductivity cell and into the plumbing loop upon instrument contact with the bottom, and 3) insufficient thermal equilibration time of the dissolved oxygen sensor. Research scientists at Seabird Electronics Inc. were extremely helpful in assessing the CTD datafiles from field and tank tests and in identifying these deployment problems (Report on Dissolved Oxygen Data by Nordeen Larson, March 1991).

CTD deployment procedures were modified for the 1991 field season in hopes of minimizing these problems (Strobel and Schimmel 1991). The instruments were not turned on until just prior to entering the water, to allow

all the air to purge from the plumbing loop during the two-minute pumpdelay. In order to allow the units to thermally equilibrate, the CTDs soaked in the surface waters for a minimum of three minutes prior to being lowered through the water column. The crews were instructed to keep the instruments from coming in contact with the seafloor. This was accomplished through a buoy/counter weight system or a well marked pay-out cable. The CTDs remained at depth (ca. one meter off the bottom) for at least two minutes. The units were hauled back to just below the surface, held there for a one minute surface soak, then brought back on board.

Additional emphasis was placed upon deployment QC checks of CTD casts. The hand-held YSI meter was used to measure dissolved oxygen concentration in water collected in a *Go-Flo* bottle from approximately one meter off the bottom at each station. This measurement was taken at approximately the same time as the CTD cast and provided a check on the operation of the CTD dissolved oxygen sensor during deployment. It also provided redundant data in case the data were lost or deemed unacceptable during the post-season review.

The CTD component of the field computer system was modified so the crews could view vertical profiles of each parameter along with the raw data file. Each CTD cast data file was reviewed in the field for evidence of deployment problems. A standard check on the data file was comparison of the downcast versus the upcast for all parameters, with particular attention to dissolved oxygen, salinity and light transmission. The dissolved oxygen profile was further evaluated by comparing the surface dissolved oxygen values at the beginning and end of the cast, and by comparing the bottom dissolved oxygen value to that recorded by the hand-held YSI meter. If either of these dissolved oxygen differences exceeded 0.5 mg/L, the field crews recalibrated the YSI and redeployed the CTD to obtain a second profile.

It was suggested that, as part of the pre-season calibration, all units be tested side-by-side in a controlled tank test. The results of that test are shown in Table 9-9.

Hydrolab DataSonde3

Hydrolab DataSonde3 deployments conducted during the 1990 Demonstration Project along with post-sampling season tank tests revealed the need for modification of certain calibration, QC check and deployment procedures.

DataSonde 3 evaluation tests resulted in a new salinity calibration procedure for the 1991 field season. It was found that salinity should be calibrated using a seawater standard rather than calibrating specific conductivity which is converted to salinity units. Tanks tests showed that it was better to calibrate salinity using a 30 ppt standard and deploy the instrument in nearly freshwater than to calibrate salinity with a 15 ppt standard and deploy the unit in a high salinity environment. In 1991, the conductivity cell was calibrated using a secondary seawater standard, the salinity of which was determined using a Guidline laboratory salinometer calibrated with Copenhagen seawater. It was decided to use a single standard throughout the entire Province rather than using assorted calibrationstandards for deployment in different salinity waters; therefore, a secondary seawater standard of approximately 30 ppt was used throughout the field season. The salinity of the standard was measured with the laboratory salinometer prior to being sent out in the field, throughout the summer and at the end of the sampling season. In all cases, the salinity drifted by less than 0.1 ppt over the three-month period.

Calibration and retrieval QC check procedures were modified to include immersing the DataSonde3 unit in a bucket of local seawater or freshwater, and comparing its temperature, salinity and dissolved oxygen readings to those recorded by a thermometer, refractometer and YSI dissolved oxygen meter, respectively. This appeared to be a better field check, because it eliminated the problem of wiping the membrane dry and possibly removing some of the biological fouling that may have affected the dissolved oxygen probe's performance.

Table 9-9. Summary of test in which all water quality instruments were placed in a well-mixed tank of seawater prior to the 1991 field season. Values are means ± 95% confidence limits for the simultaneously-recorded readings from "n" number of instruments of each type. Testing was conducted over several hours, during which time dissolved oxygen was varied to give high, medium and low concentrations; the other parameters did not vary significantly.

	Dissolved Oxygen					
	High (mg/L)	Medium (mg/L)	Low (mg/L)	Salinity (ppt)	Temperature °C	pH
Hydrolab Datasonde 3 (n = 34[a])	7.3 ± 0.4	4.4 ± 0.5	2.1 ± 0.5	24.6 ± 0.5	21.9 ± 0.1	8.6 ± 0.2
Seabird CTD (n = 4)	7.1 ± 0.2[b]	4.4 ± 0.7	2.1 ± 0.4	24.4 ± 0.1	21.9 ± 0.1	8.6 ± 0.1
YSI meter (n = 3)	7.1 ± 0.1	4.2 ± 0.1	2.0 ± 0.2	NA	21.8 ± 0.1	NA
Winkler titration (n = 16)	7.4 ± 0.4	4.4 ± 0.2	2.2 ± 0.2	NA	NA	NA

[a] Dissolved oxygen values for Hydrolab Datasonde 3's are based on n = 31 instruments; DO sensors on three instruments failed due to improperly installed membranes.

[b] The value from one instrument was omitted as an outlier in calculating this mean.

The most critical lesson learned regarding the DataSonde3 instruments was that ten-day deployments are not appropriate for the waters encountered throughout the Virginian Province. The general impression was the dissolved oxygen sensor produced reliable readings for the first five days of deployments, but then underestimated the dissolved oxygen concentration towards the end of the records. It was impossible to determine what sections of the records were acceptable and when the sensor became too fouled to produce accurate readings. A major change for the 1991 sampling season was that DataSonde3 units were deployed at all base stations for a single deployment rather than long-term servicings at a selected group of stations. The deployment period for these continuous near-bottom records were reduced from ten to three days during the 1991 field season.

The field computer system was modified to standardize the format of the data files being recorded and to streamline calibration and QC check procedures. In addition, the software included a more detailed data review routine, including time series plots for all parameters. Unfortunately, there were problems interfacing the Hydrolab software module with the boat computer system; therefore, all communication with the DataSonde3 instruments had to be done in the mobile laboratory rather than onboard the vessel. Specifics of the Hydrolab software module are documented and on file with the EMAP-VP data management group.

A series of controlled experiments were conducted to answer questions regarding the performance of the Hydrolab DataSonde3 instruments. These performance evaluations included an experiment that was conducted during 1991 crew chief training where 34 DataSonde3 units, 4 CTD instruments, 4 YSI meters and individual Winkler titrations

were used to measure the concentration of dissolved oxygen in a 500 gallon test tank. Results of this experiment are summarized in Table 9-9 and highlighted below.

-Hydrolab dissolved oxygen measurements are normally precise to within ±0.5 mg/L of the mean value; however, they are less reliable when exposed to low dissolved oxygen concentrations. This could be due to a longer response time of the *Lo-Flo* membranes to low dissolved oxygen levels.

-Approximately 10% of the instruments deployed experienced some sort of sensor malfunction; this was mostly due to a faulty calibration (*e.g.*, insufficient stabilization time prior to calibration, air bubbles beneath *Lo-Flo* membrane, etc.) rather than a malfunctioning sensor.

These experiments helped to better understand the performance and expectations of the instruments. Many practical lessons were learned that were passed on to the 1991 field crews during their training sessions.

YSI Dissolved Oxygen Meter

Incorporation of the use of the YSI dissolved oxygen meter required an additional quality control check on its performance. The YSI meters were calibrated immediately prior to use at each station using the water-saturated air calibration procedure recommended by the manufacturer. Calibration QC checks were conducted at weekly intervals in the mobile laboratories. Following calibration, the YSI probe was immersed into a bucket ofair-saturated water and allowed to stabilize. The dissolved oxygen of the water bath was determined by Winkler titration and compared to the YSI reading. If the dissolved oxygen difference exceeded 0.5 mg/L (Table 9-8), the instrument was checked thoroughly and the probe was either recalibrated or replaced. Because the unit was air-calibrated prior to use at each station, this served as a check on the overall performance of the unit and on the air-calibration method.

9.6 1991 QA Results

The 1991 sampling season yielded more reliable water column measurements than the 1990 Demonstration Project. Many lessons were learned during the Demonstration Project that led to improved calibration, field quality control check and deployment procedures. These changes, along with improved training of the field crews and more elaborate data review protocols, resulted in a significant increase in acceptable water column profiles and continuous near-bottom records.

One of the most significant improvements in 1991 was the addition of the YSI dissolved oxygen meter. The YSI meter was used for comparisons in the CTD and DataSonde3 field QC checks of dissolved oxygen. It was also used to measure bottom dissolved oxygen at all stations, which resulted in three separate bottom dissolved oxygen measurements (CTD, DataSonde3 and YSI) for most stations. These values were compared during the post-sampling season data reviews and helped to identify acceptable versus unacceptableCTD casts and DataSonde3 records.

Seabird CTD

Procedural problems regarding CTD calibrations, QC checks and deployments were minimal during the 1991 sampling season.

All calibrations were conducted at the Virginian Province instrumentation facility in Narragansett, RI. A calibration tank with air-saturated freshwater was always on hand to perform dissolved oxygen calibrations. A downfall of this system is that the faulty CTD had to be shipped to the testing facility and a replacement unit sent to the field crew. This often created a hiatus in the collection of CTD data for that field crew. On one occasion, the CTD unit was damaged during shipment, which resulted in further loss of CTD data.

Weekly calibration QC checks were an appropriate method for evaluating the performance of the sensors and recognizing any calibration drifts. The side-by-side comparisons with the YSI dissolved oxygen meter were a simple check that produced reliable results (Table 9-10).

Modified deployment procedures and more elaborate QC checks helped to increase the quality of data being collected. The on-station comparison of the CTD sensor's bottom dissolved oxygen value with the YSI bottom dissolved oxygen measurements proved to be a valuable sensor performance check. The improved data review procedures, using the updated CTD software module, allowed the field crews to recognize unacceptable casts while they were anchored on-station, providing them the opportunity to conduct another cast when needed. These checks improved the number of acceptable CTD casts (see Table 9-11). In 1991, 80% of the casts had acceptable bottom dissolved oxygen values, compared to only 70% in 1990. Acceptable bottom DO values from the CTD were measured at 91% of the stations used in EMAP's assessment of the ecological condition of the Province (*i.e.*, Base Stations); and, because redundant measurements were taken with the YSI meter, bottom dissolved oxygen concentration data are available for those stations where the CTD failed to pass QC.

The CTD software on the field computer system was a great improvement over that used in the 1990 Demonstration Project. The biggest improvement was the data archiving system, which resulted in cast data being lost from only 11 events compared to 43 in 1990. The CTD units still experienced intermittent problems of hanging up which prevented on-station downloading of data to the field computer. When this occurred, the casts could not be reviewed and the risk of unacceptable data increased, along with field crew frustration levels.

Table 9-10. Results of weekly calibration checks of water quality instruments used in the Virginian Province, 1991.

Instrument	Parameter	Checked against	Acceptance Criteria	Percent Accepted
YSI meter	Temperature	Thermometer	± 2 °C	100% (27/27)
	DO	Winkler titration	± 0.5 mg/L	89% (24/27)
Seabird CTD	Temperature	Thermometer	± 2 °C	100% (27/27)
	Salinity	Refractometer	± 2 ppt	100% (27/27)
	DO	YSI meter	± 0.5 mg/L	89% (24/27)
	pH	Standard buffer	± 0.5 units	100% (27/27)

Table 9-11. Results of 1991 post-sampling season CTD data review. (Percentages are based upon # of reviewed casts or number of Base Stations). Note: different criteria were used for accepting and rejecting CTD cast data in 1990 vs. 1991. Acceptability of casts in 1990 was based solely upon performance of the DO sensor, the performance of all sensors were considered in 1991.

	Total Casts	Base Stations
Total casts/Base Stations	291	101
# Casts accepted unqualified for all parameters (C-B code)	166 (57%)	80 (79%)
# Casts rejected for all parameters (includes lost casts)	40 (14%)	5 (5%)
# Casts w/ acceptable surface DO	236 (81%)	94 (93%)
# Casts w/ acceptable bottom DO	233 (80%)	92 (91%)

Hydrolab DataSonde3

Continuous long-term near-bottom records collected from the Hydrolab DataSonde3units were greatly improved in 1991. This was a direct result of shorter deployment periods and improved quality control procedures.

The changes in pre- and post-deployment QC checks resulted in improved field checks. Theunits were immersed into a bucket of local seawater and real-time readings compared to those from the YSI(DO), refractometer (salinity), and thermometer (temperature). This provided a more realistic assessment of the DataSonde3 dissolved oxygen probe's performance than the saturated air method employed in 1990 by eliminating the problem of potentially removing biological fouling that may have affected the dissolved oxygen records.

DataSonde3 units were deployed for a single three-day or less period at 113 stations throughout the Province (includes other than base sampling sites). This decreased deployment period resulted in a significant increase in acceptable records, particularly dissolved oxygen. In 1991, 87% of the datafiles were accepted in their entirety, compared to only 19% in 1990. A total of 94% of the 1991 retrieval QC checks for dissolved oxygen met the acceptance criteria (see Table 9-12).

The modified Hydrolab software module in the field computer system helped to decrease the number of lost data files and standardized the datafile format. The field crews were able to review time series plots for all parameters and identify any malfunctioning units in the field. On several occasions, the crews were unable to establish communications and download data from a retrieved unit. These units were returned to the field operations center and usually the data were retrieved; however, there was not a post-deployment QC check for these records.

Table 9-12. Results of calibration checks following retrieval of Hydrolab Datasonde 3 instruments for 1991 Virginian Province monitoring(base sampling sites only). Percentages indicate the number of times that acceptance criteria were met.

Parameter	Acceptance Criteria	Percent Accepted
Temperature	± 1 °C	100% (106/106)
Salinity	± 2 ppt	99% (105/106)
DO	± 0.5 mg/L	94% (100/106)
pH	± 0.5 units	99% (105/106)

YSI DO Meter

The YSI dissolved oxygen meter provided a useful QC comparison for the CTD and DataSonde3 instruments and an additional point-in-time measurement of bottom dissolved oxygen at most stations.

The YSI probes were calibrated prior to use at each station and calibration QC checks were conducted weekly. The water-saturated air calibration method appeared to yield acceptable results. During the weekly calibration QC checks, the YSI meter measured slightly lower dissolved oxygen levels than Winkler titrations and expected saturation table values, but the 0.5 mg/L acceptance criteria was met 89% of the time (Table 9-10). The YSI dissolved oxygen values agreed closely with the CTD dissolved oxygen measurements during their side-by-side checks (Table 9-10).

The results of the HACH Winkler titrations were greatly improved in 1991. In no cases did the dissolved oxygen measured in two replicate water samples exceed 0.5 mg/L; in fact the maximum difference was only 0.3 mg/L.

9.7 1992 QA Results

SeaBird CTD

All calibrations were conducted at the instrumentation facility in Narragansett, RI. A calibration tank with air-saturated freshwater was always on hand to perform dissolved oxygen calibrations. A downfall of this system is that the faulty CTD had to be shipped to the testing facility and a replacement unit sent to the field crew. This often created a hiatus in the collection of CTD data for that field crew.

Weekly calibration QC checks were an appropriate method for evaluating the performance of the sensors and recognizing any calibration drifts. The side-by-side comparisons with the YSI dissolved oxygen meter were a simple check that produced reliable results (Table 9-13).

Results of the review of 1992 CTD files are presented in Table 9-14. Acceptable bottom DO values from the CTD were measured at 92% of the stations used in EMAP's assessment of the ecological condition of the Province (*i.e.*, Base Stations). And, because redundant measurements were taken with the YSI meter, bottom dissolved oxygen concentration data are available for those stations where the CTD failed to pass QC.

YSI DO Meter

The YSI dissolved oxygen meter provided a useful QC comparison for the CTD and an additional point-in-time measurement of bottom dissolved oxygen at most stations. In addition, surface YSI values were also collected beginning in 1992. This provided a better check on the CTD than bottom measurements because the crew could better assure that both measurements were made at the exact same depth.

The YSI probes were calibrated prior to use at each station and calibration QC checks were conducted weekly. The water-saturated air calibration method appeared to yield acceptable results. During the weekly calibration QC checks, the YSI meter measured slightly lower dissolved oxygen levels than Winkler titrations and expected saturation table values, but the 0.5 mg/L acceptance criteria was met 100% of the time (Table 9-13). The YSI dissolved oxygen values agreed closely with the CTD dissolved oxygen measurements during their side-by-side checks (Table 9-13).

Table 9-13. Results of weekly calibration checks of water quality instruments used in the Virginian Province, 1992.

Instrument	Parameter	Checked against	Acceptance Criteria	Percent Accepted
YSI meter	Temperature	Thermometer	± 2 °C	100% (16/16)
	DO	Winkler titration	± 0.5 mg/L	100% (16/16)[a]
Seabird CTD	Temperature	Thermometer	± 2 °C	100% (17/17)
	Salinity	Refractometer	± 2 ppt	100% (17/17)
	DO	YSI meter	± 0.5 mg/L	100% (17/17)[b]
	pH	Standard buffer	± 0.5 units	100% (17/17)

[a] One check barely passed with a difference of 0.5 mg/L.
[b] Two tests barely passed with a difference of 0.5 mg/L

Table 9-14. Results of 1992 post-sampling season CTD data review. (Percentages are based upon # of reviewed casts or number of Base Stations). Note: different criteria were used for accepting and rejecting CTD cast data in 1990 vs. 1991 and 1992. Acceptability of casts in 1990 was based solely upon performance of the DO sensor whereas performance of all sensors were considered in 1992.

	Total Casts	Base Stations
Total casts/Base Stations	144	103
# Casts accepted unqualified for all parameters (C-B code)	58 (40%)	57 (55%)
# Casts rejected for all parameters (includes lost casts)	3 (2%)	0 (0%)
# Casts w/ acceptable surface DO	126 (88%)	98 (95%)
# Casts w/ acceptable bottom DO	123 (85%)	95 (92%)

9.8 1993 QA Results

SeaBird CTD

All calibrations were conducted at the Virginian Province instrumentation facility in Narragansett, RI. A calibration tank with air-saturated freshwater was always on hand to perform dissolved oxygen calibrations. A downfall of this system is that the faulty CTD had to be shipped to the testing facility and a replacement unit sent to the field crew. This often created a hiatus in the collection of CTD data for that field crew.

Weekly calibration QC checks were an appropriate method for evaluating the performance of the sensors and recognizing any calibration drifts. The side-by-side comparisons with the YSI dissolved oxygen meter were a simple check that produced reliable results (Table 9-15).

Results of the review of 1993 CTD files are presented in Table 9-16. Acceptable bottom DO values from the CTD were measured at 97% of the stations used in EMAP's assessment of the ecological condition of the Province (*i.e.*, Base Stations). And, because redundant measurements were taken with the YSI meter, bottom dissolved oxygen concentration data are available for those stations where the CTD failed to pass QC.

YSI DO Meter

The YSI dissolved oxygen meter provided a useful QC comparison for the CTD and an additional point-in-time measurement of surface and bottom dissolved oxygen at most stations.

The YSI probes were calibrated prior to use at each station and calibration QC checks were conducted weekly. The water-saturated air calibration method appeared to yield acceptable results. During the weekly calibration QC checks, the YSI meter measured slightly lower dissolved oxygen levels than Winkler titrations and expected saturation table values, but the 0.5 mg/L acceptance criteria was met 100% of the time (Table 9-15). The YSI dissolved oxygen values agreed fairly closely with the CTD dissolved oxygen measurements during their side-by-side checks (Table 9-15).

Table 9-15. Results of weekly calibration checks of water quality instruments used in the Virginian Province, 1993.

Instrument	Parameter	Checked against	Acceptance Criteria	Percent Accepted
YSI meter	Temperature	Thermometer	± 2 °C	100% (26/26)
	DO	Winkler titration	± 0.5 mg/L	100%(26/26)[a]
Seabird CTD	Temperature	Thermometer	± 2 °C	100% (25/25)
	Salinity	Refractometer	± 2 ppt	92% (23/25)
	DO	YSI meter	± 0.5 mg/L	96% (24/25)[b]
	pH	Standard buffer	± 0.5 units	100% (25/25)

[a] Two checks barely passed with a difference of 0.5 mg/L.
[b] One test barely passed with a difference of 0.5 mg/L

Table 9-16. Results of 1993 post-sampling season CTD data review. (Percentages are based upon # of reviewed casts or number of Base Stations). Note: different criteria were used for accepting and rejecting CTD cast data in 1990 vs. 1991 -1993. Acceptability of casts in 1990 was based solely upon performance of the DO sensor whereas performance of all sensors were considered in 1993.

	Total Casts	Base Stations
Total casts/Base Stations	147	111
# Casts accepted unqualified for all parameters (C-B code)	105 (71%)	84 (76%)
# Casts rejected for all parameters (includes lost casts)	0 (0%)	0 (0%)
# Casts w/ acceptable surface DO	139 (95%)	105 (95%)
# Casts w/ acceptable bottom DO	143 (97%)	108 (97%)

Section 10
QA Results for Total Suspended Solids Analyses

10.1 Background

In 1990 water samples for total suspended solids (TSS) analysis were collected only at Indicator Testing and Evaluation Sites. The intent was to use these data to evaluate transmissometer data collected from those stations. Because these data are not used in the assessment of the ecological condition of the Province, no QA results for 1990 TSS analyses are presented in this document.

Surface water samples were collected at all Base Sampling Sites beginning in 1991. A 250cc plastic bottle was filled with water, refrigerated, and shipped to the laboratory for filtering and analysis according to standard EPA methods.

A problem with the QC process for TSS samples was discovered during the preparation of this report. No QA criteria were in place in 1991 by which the data could be evaluated against. A retrospective evaluation of the 1991 data using the criteria set in place in 1992 show that large percentage of the samples should have been flagged as failing QC when, in fact, they were not. However, in reviewing data from all three years it was discovered that the QC requirements in the QA Plans were unrealistic and frequently not met. This is discussed below in Sections 10.4 and 10.7.

10.2 Data Qualifier Codes for Total Suspended Solids Data

Data qualifier codes for the suspended solids dataset are listed in Table 10-1. The SOP called for filtering a large enough volume of water to ensure the residue weight was atleast one milligram. The SS-C code was applied to those samples with low TSS concentrations for which samples were refiltered using a larger volume of water, and the weight of the residue was still less than one milligram.

Table 10-1. Data qualifier codes for total suspended solids data (NOTE: These codes may change - see Section 10.7).

Code	Description
SS-A	Sample failed to meet EMAP-Estuaries QA requirements. Relative percent difference between duplicates exceeded 10%. Data should be used with caution.
SS-B	No QC samples were run on the day this sample was analyzed. These data can not be evaluated relative to EMAP-Estuaries QA standards.
SS-C	Residue weight was less than 1.0 mg even with larger volume filtered. Value reported was associated with the largest volume filtered.

10.3 Audits

All TSS analyses conducted in 1991 and 1992 were performed by SAIC's Environmental Testing Center (ETC). As described in Section 6.3, this laboratory was audited in 1990 and 1991. The results of the 1991 audit included TSS analysis, and were generally favorable, with no QA infractions noted. Samples collected in 1993 were analyzed by the Marine Ecosystem Research Laboratory (MERL) of the University of Rhode Island. This laboratory has extensive experience in TSS analyses; therefore, no audits were deemed necessary.

10.4 1991 QA Results

TSS samples were introduced as a research indicator in 1991, and, as such, no QA requirements were included in the QA Plan. However, as part of routine analysis approximately 10% of the samples were reanalyzed. Subsequent QA Plans required that at least 10% of all samples analyzed for TSS concentration be analyzed in duplicate. To pass QA, the RPD between the duplicates must be less than 10%. If it exceeds 10%, all samples analyzed since the last successful QC check must be repeated.

The mean RPD for the 14 sets of duplicates was 10.4%, with a maximum of 32.6%. Six of the fourteen sets exceeded 10%; however, none of the data were assigned QA codes because control limits were not in place at the time of the review. See Section 10.7 for additional discussion.

10.5 1992 QA Results

The QA Plan required that at least 10% of all samples analyzed for TSS concentration be analyzed in duplicate. The RPD between the duplicates was then calculated. To pass QA, this value must be less than 10%. If it exceeds 10%, all samples analyzed since the last successful QC check must be repeated.

Due to an apparent mis-communication at the analytical laboratory, the first group of samples did not have the appropriate QA samples run. Therefore, the quality of the resultant data cannot be evaluated and are "flagged" in the EMAP database with the SS-B code. A sufficient number of duplicate analyses were performed with the remainder of the samples; however, several failed QA, with the RPD exceeding 10%. Unfortunately this was not discovered until several months after the analyses were completed, and the original samples (degradable) had been discarded. As a result, approximately 44.4% of the data have been flagged as being of questionable quality (SS-A or SS-B).

10.6 1993 QA Results

The QA Plan required that at least 10% of all samples analyzed for TSS concentration be analyzed in duplicate. The RPD between the duplicates was then calculated. To pass QA, this value must be less than 10%. If it exceeds 10%, all samples analyzed since the last successful QC check must be repeated.

The analytical laboratory chose to analyze all of the samples in duplicate. The RPD for these analyses ranged from 0 to 35.7% with a mean of 10.5%. The median RPD was 8.5%. Because duplicate analyses were available for nearly all samples (duplicate data for 17 of the samples were not available due to analytical problems), and the mean RPD slightly exceeded EMAP control limits, we chose to report the mean of the duplicates and not to assign QA qualifier codes to any of the results. See Section 10.7 for additional discussion.

10.7 Lessons Learned and Changes Suggested

Results for total suspended solids generated in 1991 to 1993 suggest a problem with the stated QA process. The RPD for nearly half of the 1991 and 1992 duplicate pairs fell outside of the control limits. Evaluation of the 1993 results, which were generated by MERL (an academic laboratory with extensive experience in TSS analyses), showed that half of the samples had an RPD greater than 8.5%, with a mean RPD of 10.5%.

In the analysis of TSS samples, water is filtered and small masses of sediment weighed. The relatively large tare weight of the pans when compared to the small weight of the samples likely results in the errors shown. The 1993 results suggest that a better method would be to analyze ALL samples in duplicate and report the mean of the measurements. We recommend that in the future this methodology be employed for all TSS samples.

In the interim we suggest that a new QA Qualifier Code be applied to all 1991 and 1992 samples which simply states that the value reported represents results from a single measurement rather than the mean of two measurements.

Section 11
Summary of Data Collection Success

Data completeness goals are provided in the annual Quality Assurance Project Plans. Generally a minimum completeness goal of 90% is listed for each indicator. Table 11-1 provides summary information regarding data completeness. Of the 446 Base Sampling Sites originally selected, 21 were deemed unsampleable due to inaccessibility, obstructions, or water depth and could not be moved in accordance with the design. The completeness rate for most indicators was above or close to the 90% mark. The notable exception is suspended solids; however, the collection of samples for TSS analyses at all Base Sampling Sites did not begin until 1991.

Table 11-1. Summary of collection and processing status of samples collected in 1990-1993 (Base Sampling Sites only).

Sample Type	# Stations Expected to be Sampled[a]	# Stations Sampled With Data Passing Final QC[b] (%)	
Dissolved Oxygen	446	420	(94%)
Light Attenuation Coefficient (CTD cast)	446	408	(91%)
Suspended Solids	446	298	(67%)[c]
Sediment Chemistry[d]			
Organics	446	397	(89%)
Metals	446	394	(88%)
Sediment Toxicity	446	373	(84%)
Sediment Grain Size	446	394	(88%)
Benthic Infauna	446	404	(91%)
Fish Community Data (successful trawl)	446	390	(87%)

[a] A total of 446 Base Sampling Sites were originally selected for sampling. Of these, 21 were found to be unsampleable due to obstructions or inadequate water depth prior to the sampling season.

[b] This value takes into account samples not collected, damaged or lost during shipping or processing, or failing to pass final Quality Control checks.

[c] Samples for TSS analyses were not collected in 1990. Note that QA Criteria did not exist for 1991 samples.

[d] The success rate denotes percent of stations with some valid data. However, as discussed in Section 3, not all stations successfully sampled have valid data for all analytes.

Section 12
References

Holland, A.F., ed. 1990. *Near Coastal Program Plan for 1990: Estuaries.* EPA 600/4-900/033. Narragansett, RI: U.S. Environmental Protection Agency, Environmental Research Laboratory, Office of Research and Development.

Latimer, R.W. 1992. FY 1992 Annual Quality Assurance Report and FY 1993 Work Plan for the Environmental Monitoring and Assessment Program Estuaries Resource Group. Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. September 1992.

Schimmel, S.C., B.D. Melzian, D.E. Campbell, C.J. Strobel, S.J. Benyi, J.S. Rosen, and H.W. Buffum. 1994. *Statistical Summary: EMAP-Estuaries Virginian Province - 1991*. EPA/620/R-94/005 Narragansett, RI: U.S. Environmental Protection Agency, Environmental Research Laboratory, Office of Research and Development.

Strobel, C.J. and S.C. Schimmel. 1991. *EMAP-Estuaries 1991 Virginian Province Field Operations and Safety Manual*. Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, June 1991.

Summers. 1993. FY 1993 Annual Quality Assurance Report and FY 1992 Work Plan for the Environmental Monitoring and Assessment Program Near Coastal Resource Group. Gulf Breeze, FL: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. September 1993.

Valente, R.M. 1991a. FY 1990 Annual Quality Assurance Report and FY 1991 Work Plan for the Environmental Monitoring and Assessment Program Near Coastal Component. Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. February 1991.

Valente, R.M. 1991b. FY 1991 Annual Quality Assurance Report and FY 1992 Work Plan for the Environmental Monitoring and Assessment Program Near Coastal Resource Group. Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. September 1991.

Valente, R. and J. Schoenherr. 1991. *EMAP-Estuaries Virginian Province Quality Assurance Project Plan*. Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. July 1991.

Valente, R., C.J. Strobel J.E. Pollard, K.M. Peres, T.C. Chiang and J. Rosen. 1990. *Quality Assurance Project Plan for EMAP Near Coastal: 1990 Demonstration Project.* Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. July 1990.

Valente, R., and C.J. Strobel. 1993. Environmental Monitoring and Assessment Program - Estuaries: 1993 Virginian Province Quality Assurance Project Plan. U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory, Narragansett, RI. May 1993.

Valente, R., C.J. Strobel and S.C. Schimmel. 1992. *EMAP-Estuaries Virginian Province 1992 Quality Assurance Project Plan*. Narragansett, RI: U.S. Environmental Protection Agency, Office of Research and Development, Environmental Research Laboratory. July 1992.