

Chesapeake Bay Program Blind Audit
2001 Final Report

November 2002

Carl Zimmermann
Carolyn Keefe
Nutrient Analytical Services laboratory
Chesapeake Biological Laboratory
Solomons, MD 20688

FINAL REPORT

INTRODUCTION:

The purpose of this Blind Audit Program is to provide samples of specific nutrient analytes at concentrations commonly found in estuarine systems for analysis by laboratories who analyze water samples collected from the Chesapeake Bay and its tributaries. The concentrations of these samples, which are unknown to the recipient analysts, are compared to their prepared concentrations.

In the early years of the Chesapeake Bay Program, the U.S. EPA provided blind audit samples on an irregular basis to laboratories analyzing Chesapeake Bay water samples. However, these audit samples were designed for waste water/drinking water applications rather than estuarine water applications. Consequently, the concentrations were much higher than normally occur in the Bay and did not provide a reasonable estimate of accuracy for low level nutrient concentrations. For example, a blind audit concentration of 1.0 mg NH₄-N/L would be comparable to NPDES water samples but would be at least an order of magnitude greater than concentrations normally occurring in most parts of Chesapeake Bay.

The only continuous program providing an estimate of laboratory performance has been the Chesapeake Bay Coordinated Split Sample Program (CSSP). Data generated from this program provide the only long term QA/QC data base to compare nutrient measurements provided by laboratories analyzing water samples collected from Chesapeake Bay and its tributaries. Samples for the CSSP are natural water samples collected from Chesapeake Bay or a tributary. Briefly, a common unfiltered water sample is distributed to the various field/laboratory personnel who in turn subsample into dissolved and particulate fractions. These are analyzed and the results compared to those of other participating laboratories. Resulting data analysis can show how field filtration techniques and/or laboratory practices affect data variability. The CSSP samples are each subject to cumulative errors of analytical determinations from variation in both field and laboratory procedures. Also, these data sets cannot definitively determine the accuracy of laboratory analyses.

The current Blind Audit Program was designed to complement the CSSP. Blind Audit particulate samples distributed to participants have few cumulative errors associated with field filtering and subsampling procedures. Prepared concentrates of dissolved substances, whose concentrations are unknown to the analysts, are provided so that laboratory accuracy can be assessed.

This is the fourth year of the Blind Audit Program and it is the continued intent of this program to provide unknown, low level dissolved and particulate nutrient samples to laboratories analyzing Chesapeake Bay Program nutrients, as well as to other laboratories interested in participating in the Blind Audit Program.

MATERIALS AND METHODS:

Blind Audit samples were sent to participating laboratories in January (29 January 2001) and August (29 August 2001) 2001. Those participating laboratories and contact personnel are found in Table 1.

Parameters measured during the February audit were: total dissolved nitrogen, total dissolved phosphorus, nitrate+nitrite, ammonium and phosphate. A high and low concentration sample were provided for each of these analytes. Particulate carbon, nitrogen and phosphorus samples as well as chlorophyll were also provided for those laboratories that routinely analyze these parameters. The chlorophyll samples were natural population samples collected from the mouth

of the Patuxent River.

Dissolved Blind Audit concentrates were prepared by careful dilution of high quality standards using 18.3 megohm deionized water. The concentrates were sealed in 20 mL ampules for shipment to the participants. One ampule contained a concentrate of an organic nitrogen compound and an organic phosphorus compound to be diluted for the analysis of low level total dissolved nitrogen and total dissolved phosphorus. A second ampule contained a concentrate of organic nitrogen and organic phosphorus to be diluted for the analysis of higher level total dissolved nitrogen and total dissolved phosphorus. A third ampule contained a concentrate to be diluted for the analysis of low level inorganic nutrients (ammonium, nitrate and phosphate). A fourth ampule contained a concentrate to be diluted for the analysis of higher level inorganic nutrients. At each participating laboratory, an aliquot from each ampule was diluted and analyzed according to accompanying instructions for preparation and dilution. These diluted Blind Audit samples were then inserted randomly in a typical estuarine sample set. Final concentrations were reported for each diluted concentrate according to the dilution instructions provided.

Particulate analytes are measured by analyzing suspended material concentrated on filter pads. There are no commercially available suspensions of pure carbon, nitrogen or phosphorus compounds, so a natural sample was subsampled onto filter pads for analysis by participating laboratories. A batch water sample was collected off the CBL pier in January and August, and subsampled for particulate samples of carbon, nitrogen and phosphorus. Particulate C/N samples were filtered from the batch sample with care being taken to shake the batch sample before each filtration to ensure homogeneity. Two 25 mm GF/F were sent to each laboratory for analysis. Vacuum filtration was used to process the filters. Samples were dried completely (overnight at 47° C) before shipment.

The same general procedure was followed for particulate phosphorus samples which were concentrated by vacuum filtration on 47 mm GF/F pads.

Filter pads were sent to each laboratory for the analysis of particulate C, N, and P. The volume of sample filtered was noted in the instructions so that each laboratory could report concentrations in mg/L. Chlorophyll results were reported as ug/L.

For both audits, samples were sent in coolers via next day carrier to the participating laboratories. Because chlorophyll samples were sent, a cold temperature was required, so frozen cold packs were packed in those coolers.

RESULTS

Tables and figures summarizing results from 2001 are found at the end of the report. Concentrations were assessed statistically by calculating the mean and standard deviation of each sample set, then calculating how many standard deviations separated each laboratory's reported concentration from that mean. The relative percent difference (RPD) between each laboratory's reported concentration and the true concentration was also calculated.

WINTER 2001

Total Dissolved Nitrogen: The prepared low level concentration was 0.30 mg N/L and reported

concentrations ranged from 0.30-0.40 mg N/L. The prepared high level concentration was 0.945 mg N/L and reported concentrations ranged from 0.926-1.1 mg N/L (Figure 1). Percent coefficient of variation for the low sample was less than 10 % (9.6%) and the coefficient of variation for the high sample was an astounding 5.7%. No laboratory reported a concentration with a difference 0.1 mg N/L greater than the mean. In fact, only one laboratory reported a concentration whose difference from the mean was greater than 0.05 mg N/L. The other eight laboratories' results were less than 0.05 mg N/L from the mean concentration. For the high level sample, only two laboratories reported concentrations greater than 0.1 mg N/L from the prepared value. The remaining seven laboratories reported concentrations within 0.04 mg N/L of the prepared concentration. Statistics shown in table 2 indicate that all laboratories passed at both concentration levels--six of the nine laboratories were within one standard deviation of the mean concentration for the low level sample and seven of the nine laboratories results from the high concentrate were within one standard deviation of that mean.

Total Dissolved Phosphorus (TDP): The prepared low level concentration was 0.017 mg P/L and reported concentrations ranged from 0.011-0.020 mg P/L. Percent recoveries of the low TDP sample ranged from 65-118% of the prepared concentration (Figure 1). The mean reported concentration of this low level sample was 0.0166 mg P/L with a standard deviation of 0.0029 and a coefficient of variation of 17.5%. Only two of the nine laboratories reported concentrations that were more than one standard deviation from the mean. Results of the higher TDP concentration sample showed the same pattern. The prepared high level concentration was 0.043 mg P/L and reported concentrations ranged from 0.0308-0.047 mg P/L. One laboratory received a "warning" designation and one laboratory reported a concentration that was between one and two standard deviations of the mean (Table 2). The remaining laboratories reported results that were within one standard deviation of the mean. The mean concentration determined by all the reporting laboratories was 0.0403 mg P/L with a standard deviation of 0.0047 (11.7% coefficient of variation).

Ammonium: The prepared low level concentration was 0.0297 mg N/L and reported concentrations ranged from 0.0125-0.0444 mg N/L (Figure 2). The overall coefficient of variation for the low level sample was an abysmal 33%. Eight of the eleven laboratories which reported results were within one standard deviation of the mean (0.029 mg N/L) [Table 2] while the other three laboratories were between one and two standard deviations. The prepared high level concentration was 0.35 mg N/L and reported concentrations ranged from 0.262-0.424 mg N/L. The mean concentration was 0.355 mg N/L with a reported standard deviation of 0.0390 (11% coefficient of variation). One of the eleven laboratories received a "warning" for their reported concentration that was between two and three standard deviations from the mean and one lab result was between one and two standard deviations from the mean. Nine laboratories reported results that were within one standard deviation of the mean (Table 2).

Nitrate+nitrite: The prepared low level concentration was 0.0175 mg N/L and reported concentrations ranged from 0.017-0.040 mg N/L (Figure 2). A mean concentration reported by the laboratories was 0.022 mg N/L (Standard deviation 0.0070, coefficient of variation 32%). Nine laboratories' results for this low level sample fell within one standard deviation of the mean, one laboratory's result was within 1-2 standard deviations of the mean and one received a warning (Table 2). The prepared high level concentration was 0.875 mg N/L and reported concentrations ranged from 0.835-0.9 mg N/L (Mean .870 mg N/L, Standard deviation 0.0201, CV 2.3%). All laboratories reported concentrations for the high level nitrate+nitrite unknown that were within 5% of the prepared concentration. Is that great, or what? All laboratories received a pass rating with seven laboratories reporting results within one standard deviation of the mean

Particulate C		1.285	1.284	1.168	1.291	1.254	1.265
Particulate N		0.122	0.156	0.1450	0.159	0.1480	0.1470
Particulate P	0.0099		0.0120	0.0093	0.0065		0.0092

Chlorophyll: Ten laboratories reported chlorophyll results for the Winter blind audit (Figure 3). Concentrations ranged from 3.68-6.12 ug/L with a mean concentration of 4.62 ug/l, a standard deviation of 0.73 and a coefficient of variation of 15.8%.

SUMMER 2001 DISSOLVED FRACTION

Total Dissolved Nitrogen: The prepared low level concentration was 0.315 mg N/L and reported concentrations of digested samples ranged from 0.269-0.398 mg N/L (Figure 4). Seven laboratories reporting data for digested samples were within one standard deviation of the mean concentration (0.336 mg N/L). The coefficient of variation for this low level sample was 11.6%. The other two laboratories' concentrations were between 1 and 2 standard deviations of the mean. The prepared high level concentration was 1.00 mg N/L and reported concentrations of digested samples ranged from 0.9667-1.154 mg N/L. Six of the nine laboratories reporting data for digested samples were within one standard deviation of the mean concentration (1.0481 mg N/L). The other three laboratories' data were 1 to 2 standard deviations from the mean. A coefficient of variation of 5.7% was determined for this high concentration sample. Statistics shown in table 2 indicate that all laboratories passed at both concentration levels.

Total Dissolved Phosphorus: The prepared low level concentration was 0.017 mg P/L and reported concentrations ranged from 0.014-0.022 mg P/L (Figure 4). The prepared high level concentration was 0.045 mg P/L and reported concentrations ranged from 0.043-0.066 mg P/L. Six laboratories reporting data for the low total dissolved P unknown were within one standard deviation of the mean (0.0187 mg P/L). Results for the high concentration of total dissolved P indicated that eight laboratories reported results that were less than one standard deviation of the mean (0.0493 mg P/L) while one laboratory's result necessitated a warning. Coefficients of variation for the low and high dissolved organic phosphorus samples were 14.4% and 14%, respectively.

Ammonium: The prepared low level concentration was 0.024 mg N/L and reported concentrations ranged from 0.0160-0.038 mg N/L (Figure 5). Six laboratories reported results that were within one standard deviation of the mean (0.026 mg N/L) for the low level sample. Four laboratories results were between 1 and 2 standard deviations. The between laboratory variation (%CV) was 25.4%.

The prepared high level concentration was 0.525 mg N/L and reported concentrations ranged from 0.452-0.574 mg N/L (Figure 5). The between laboratory variation of 7.1%, based on the mean (0.528 mg N/L) and standard deviation was quite small. Of the ten laboratories that reported data, one received a warning, two were between 1-2 standard deviations of the mean and seven were within one standard deviation of the mean.

Nitrate+nitrite: The prepared low level concentration was 0.0152 mg N/L and reported concentrations ranged from 0.011-.0200 mg N/L (Figure 5). The prepared high level concentration was 0.782 mg N/L and reported concentrations ranged from 0.668-0.814 mg N/L.

Seven laboratories reported results for the low level nitrate sample that were within one standard deviation of the mean (0.016 mg N/L). Two laboratories provided results for this sample that were between one and two standard deviations of the mean and one lab received a warning (Table 3).

For the high level nitrate sample, eight laboratories' results were within one standard deviation of the mean (0.760 mg N/L), one was between one and two standard deviations of the mean and one received a warning. Coefficients of variation for the low and high concentration samples were 15% and 4.9%, respectively.

Phosphate: The prepared concentrations for the low and high phosphate samples were 0.0093 and 0.0465 mg P/L, respectively. Results for the low level sample ranged from 0.007-0.0114 mg P/L (Figure 5) with eight laboratories' results falling within one standard deviation of the mean (0.010 mg P/L), one laboratory falling into the "warning" category and one laboratory's results between one and two standard deviations of the mean. High level phosphate results ranged from 0.0406-0.0499 mg P/L with eight laboratories' results falling within one standard deviation of the mean (0.047 mg P/L), one lab with a result that was between one and two standard deviations of the mean and one laboratory receiving a "warning" because their result was between 2-3 standard deviations of the mean (Table 3).

SUMMER 2001 PARTICULATE FRACTION

Particulate Nitrogen: The seven laboratories that analyzed for particulate nitrogen reported results that varied slightly, one from another (0.247-0.322 mg N/L, Figure 6). This yielded a mean concentration of 0.2789 mg N/L \pm 0.0251 S.D. All laboratories' results passed. Five laboratories' results fell within one standard deviation of the mean; while the other two were between one and two standard deviations of the mean. The coefficient of variation of 4.3% (N=7) among the participating laboratories was equivalent to the 5.1% variation found for 12 samples analyzed at CBL in January 1998.

Particulate Carbon: Particulate carbon values also varied little among the participating laboratories (1.535-1.790 mg C/L, Figure 6). This yielded a mean concentration of 1.63 mg C/L \pm 0.0969 S.D. All laboratories' results passed. Five laboratories' results fell within one standard deviation of the mean; while the other two were between one and two standard deviations of the mean. The coefficient of variation 9.0% (N=7) among the participating laboratories was slightly less than the 12.1% variation found for 12 samples analyzed at CBL in January 1998.

Particulate Phosphorus: Six laboratories analyzed samples for particulate phosphorus (Figure 6). Concentrations ranged from 0.0069-0.0273 mg P/L with a mean concentration of 0.0219 mg P/L \pm 0.0075 S.D. All laboratories' results passed. Five laboratories' results fell within one standard deviation of the mean; while the other was between one and two standard deviations of the mean. The large coefficient of variation of 34.2% (N=6) among the participating laboratories is in large part due to one laboratory's very low result (compared with the other results). This was brought to the laboratory's attention for investigation. They subsequently reported no analytical or computational anomalies.

Chlorophyll: Seven laboratories reported results that ranged from 8-15.6 ug/L where the mean concentration was 10.92 ug/L, a standard deviation of 2.57 and a coefficient of variation of 23.5%. Since the samples were natural replicate subsamples, there is no true or prepared

concentration with which to compare.

DISCUSSION

Several important issues should be considered when assessing whether individual Blind Audit results are within acceptable limits.

Variation Associated With An Analytical Method: As we have noted in previous Blind Audit Reports, analytical variability is associated with any quantitative determination. The method detection limit (three times the standard deviation of seven low level replicate natural samples) is often used to express that level of variation. Total dissolved nitrogen data provide a good example. The detection limit at CBL has been determined to be 0.02 mg N/L. Any total dissolved nitrogen measurement has a potential 0.02 mg N/L variability associated with it. This variability, when expressed as a percent of the “true” concentration, can be extremely large for low level concentrations and fairly low for higher concentrations. For example, a 0.20 mg N/L concentration has an analytical variability of 10% associated with it; whereas, a 1.20 mg N/L concentration has an analytical variability of 2%.

Acceptance Limits of Provided Samples: Companies that prepare large quantities of performance evaluation samples assign acceptable confidence limits around the “true” value. In one case (SPEX, CertiPrep), the mean recovery and standard deviation are later reported along with the true concentration and the 95% confidence interval (CI). The 95% CI is the mean recovery \pm 2 standard deviations and is developed from regression equations from Water Pollution Performance Evaluation Studies. A recently purchased set of these standards gave a true total P value of 3.00 mg P/L with a 95% CI of 2.47-3.42 mg P/L. The lower end of the 95% CI recovery allows 82% recovery of the true concentration. This type of statistical analysis was not performed on the Blind Audit Program samples prepared for this study prior to their distribution to the participants.

Parameters assessed in the Blind Audit do not have predetermined acceptance limits, so we are following the statistical procedure of ERA, an approved source of wastewater and drinking water proficiency samples, and the State of Wisconsin Proficiency Testing program. They average the results for each parameter and at each concentration, then calculate the standard deviation from the mean. Results that are within 3 standard deviations “pass”, and those greater than 3 standard deviations “fail”. Results between 2 and 3 standard deviations are in the “warning” category.

Most of the data comparisons showed similar characteristics (Table 2); that is, the reported concentrations were similar, and one or two concentrations fell slightly beyond one standard deviation from the mean of all the data for that portion of the study. Apparently, it is a statistical “reality” in small sample sets with little variability between individual points, that at least one point will lie just beyond one standard deviation from the mean. Thus, for most of the data comparisons, all the reported concentrations “passed.” It should also be noted that no data fell in the “fail” category, although a more substantial number were in the “warning” category than in previous years. In 2000, three warnings were identified: one for a low level ammonium result, and one each for low and high level phosphate. Two laboratories were involved. In 2001, eleven warnings were identified: five in the winter and six in the summer, involving six laboratories.

The data sets with relatively small standard deviations yielded more “warning” points, but these points were within 10% of the relative percent deviation from the prepared concentration. For

example, in the Summer 2001 blind audit of high level nitrate concentration, the mean reported concentration was 0.760 mg N/L and reported concentrations ranged from 0.668-0.814 mg N/L. Seven laboratories reported results for the low level nitrate sample that were within one standard deviation (0.0373 mg N/L) of the mean. Since the standard deviation and coefficient of variation (4.9%) were so small, two laboratories' reported results for this sample were between one and two standard deviations of the mean and one laboratory's results were between two and three standard deviations of the mean, so it was labeled as a "warning." This nitrate data comparison points toward a form of circular reasoning in these statistical assessments. The data being evaluated are also the data which were used to calculate the mean and standard deviation to which the data are being compared.

No laboratory reported concentrations for an individual analyte that were consistently different from the range of the other reported concentrations for both concentration ranges tested for that analyte.

For each study, particulate samples were filtered from a common estuarine water sample and, consequently, are not true blind audit samples made from pure constituents. There is no "true" or prepared concentration with which to compare. In all instances, the standard deviation was less than 10% of the mean reported concentration for particulate carbon and nitrogen.

The proportion of the standard deviation to the mean was high for particulate phosphorus in both 2001 blind audits. This contrasted to all previous years of blind audits in which the coefficient of variation for particulate phosphorus was the lowest of the particulate fractions. In both 2001 blind audits, one or two laboratory's reported concentrations were visibly different from the mean, thus increasing the coefficient of variation. In 2001, a greater volume was filtered which would result in less sampling error. The sample sizes were only five or six, so it was not surprising that these differences were insufficient to generate a warning. A visual inspection of the summer particulate phosphorus data would indicate that one data point was clearly different from the others. These particulate phosphorus data comparisons are an obvious example of the danger of circular reasoning in these statistical assessments. The data being evaluated are also the data which were used to calculate the mean and standard deviation to which the data are being compared. No laboratory reported concentrations for particulate phosphorus that were consistently different from the range of the other reported concentrations for both 2001 blind audits.

Comparison With Previous Blind Audit: The same concentration (0.0170 mg P/L) of low level total dissolved phosphorus was prepared for the summer 2000 and summer 2001 blind audits. A comparison of these data (Table 4) shows that the mean concentrations, standard deviations and resulting percent coefficients of variation were nearly identical for both sets. These results demonstrate the consistently reproducible performance of the laboratories in analyzing low level total dissolved phosphorus concentrations.

Table 4. Low level total dissolved phosphorus, summer 2000 and summer 2001.						
Date	Prepared conc. (mg P/L)	Mean (mg P/L)	Std. Dev.	Coefficient of variation	Labs Pass	Labs Warn
Summer 2000	0.0170	0.0187	0.0036	19.2%	8	0
Summer 2001	0.0170	0.0187	0.0027	14.4%	9	0

Reporting Data Accurately: A surprisingly large percentage of results were entered in the wrong

place on the reporting forms, or had “slipped a decimal” or exhibited some other entry error that could have been easily avoided. Contacting the participants usually resolved these reporting discrepancies and also improved their subsequent reporting practices.

The number of significant figures reported in analytical results can significantly affect data comparability in a blind audit study. If a laboratory reports only two significant figures (for whatever reasons) and an audit sample has a prepared concentration expressed in three significant figures, then substantial under or over estimates of the comparative concentration can be reported. For example, if a 0.032 mg P/L sample has been prepared and a laboratory only reports two significant figures, i.e., 0.03 mg P/L, then the results expressed are 86% of the expected prepared value. During the 2000 study, all participants reported three significant digits for most parameters. It is noteworthy that the 2000 study's coefficients of variation were generally smaller than in the previous two years—probably a result of comparisons of data containing the appropriate number of significant digits. Unfortunately, some 2001 participants reported only one significant digit, thus potentially giving substantial under or over estimates for the comparisons.

CONCLUSION

Now that eight rounds of the Blind Audit Program have been completed, some consistent patterns have been observed that warrant action or further investigation:

1. Reported concentrations of analytes were usually similar between laboratories participating in the Blind Audit Program. No laboratory reported concentrations for an individual analyte that were consistently different from the range of the other reported concentrations for both concentration ranges tested for that analyte. This indicates that most participating laboratories execute and report these measurements with accuracy and precision, reporting the appropriate number of significant digits.
2. Care should continue to be taken when completing report forms. During 2001 some results were written in the wrong place, or had “slipped a decimal,” or reported insufficient significant digits, or contained some other error that could have been easily avoided.

Table 1. Participants in the 2001 Chesapeake Bay Blind Audit Program					
Institution	Contact Person	Phone	Dissolved	Particulate	Chl. A

Old Dominion University, AMRL	Suzanne Doughten	757-664-1043	X	X	X
U. Maryland, HPL	Lois Lane	410-221-8252	X	X	
Virginia Institute of Marine Science	Carol Pollard	804-642-7213	X	X	X
Va. Div. Consol. Lab Services	Robert Potts	804-786-7213	X	X	X
Pa. Dept. Environmental Resources	Michelle Clarke	717-783-1998	X		
Va. Tech. Occaquan Lab	Mary Lou Daniel	703-361-5606	X		
Md. Dept. Heath&Mental Hygiene	Deborah Miller-Tuck	410-767-6180	X		
U. Maryland, CBL	Carl Zimmermann	410-326-7252	X	X	X
D.C. Government/DOH	Al Robertson	410-573-2600	X		
Univ. Delaware	Joe Scudlark	302-645-4300	X	X	X
Delaware DNR	Ben Pressly	302-739-4771	X		
Philadelphia Academy of Science	Paul Kiry	215-299-1076	X	X	X