Acquisition Directorate

Research & Development Center

Report No. CG-D-03-09

Evaluation of Representative Sampling for Rare Populations Using Microbeads

Distribution:

This document is available to the U.S. public through the National Technical Information Service, Springfield, VA 22161.

August 2008

Homeland Security

N O T I C E

This document is disseminated under the sponsorship of the Department of Homeland Security in the interest of information exchange. The United States Government assumes no liability for its contents or use thereof.

The United States Government does not endorse products or manufacturers. Trade or manufacturers' names appear herein solely because they are considered essential to the object of this report.

This report does not constitute a standard, specification, or regulation.

James E. Fletcher
SRC Branch Chief
United States Coast Guard
Research & Development Center
1082 Shennecossett Road
Groton, CT 06340-6048

Technical Report Documentation Page

1. Report No. CG-D-03-09	2. Government Accession Number ADA493426	3. Recipient's Catalog No.
4. Title and Subtitle EVALUATION OF REPRESENTATIVE SAMPLING FOR RARE POPULATIONS USING MICROBEADS		5. Report Date August 2008 6. Performing Organization Code Project No. 4125
7. Author(s) E.J. Lemieux, S. Robbins, K. Burns, S. Ratcliff, and Penny Herring (USCG)		, 8. Performing Report No. R&DC UDI # 756
9. Performing Organization Name and Address Naval Research Laboratory Bldg. F-14, Fleming Key Trumbo Annex Key West, FL. 33040	U.S. Coast Guard Research and Development Center 1082 Shennecossett Road Groton, CT 06340-6048	10. Work Unit No. (TRAIS) 11. Contract or Grant No. HSCG32-05-XR00014
12. Sponsoring Organization Name and Address U.S. Department of Homeland Security United States Coast Guard Office of Operating and Environmental Standards Washington, DC 20593-0001		13. Type of Report & Period Covered Final 14. Sponsoring Agency Code Commandant (G-5224) U.S. Coast Guard Headquarters Washington, DC 20593-0001
15. Supplementary Notes The R&D Center's technical point of contact is Penny Herring, 860-441-2868, email: penny.r.herring@uscg.mil.		

16. Abstract (MAXIMUM 200 WORDS)

Determining a way to adequately sample sparse populations remains a challenge in ballast water testing. In an effort to quantify variability in counting, the Naval Research Laboratory conducted an experiment which eliminated as much variability as possible by use of standardized micro beads in artificial seawater. Concentrations counted were based on Environmental Technology Verification Program, International Maritime Organization and additional proposed guidelines, all of which involved sparse populations. Inert 10-micron beads at 1, 5, 10, 50, 100, 500 and 1000 beads per milliliter (mL) were used to represent phytoplankton while 100-micron beads at 10, 30 and 60 beads per 500 mLs of water represented zooplankton. Combinations of initial sample collected and level of concentration were tested as well as observer bias. The results showed that for phytoplankton beads, the 50, 100 and 1000 concentrations were not significantly different and had accuracy and precision measurements below the 10% and 20% benchmarks respectively. The zooplankton populations all showed poor precision with all measurements well above the 20% benchmark. In both experiments, no significant differences were found between observer counts.

17. Key Words ballast water, sparse populations, micro beads, low concentrations, minimum detection		18. Distribution Statement This document is available to the U.S. public through the National Technical Information Service, Springfield, VA 22161.	
19. Security Class (This Report) UNCLASSIFIED	20. Security Class (This Page) UNCLASSIFIED	21. No of Pages 46	22. Price

(This page intentionally left blank.)

EXECUTIVE SUMMARY

The invasion of non-indigenous species (NIS) into new aquatic and marine ecosystems is a significant problem worldwide. In the United States, NIS such as zebra mussels, quagga mussels, Chinese mitten crabs, and Eurasian water milfoil clog waterways and industrial piping, out-compete native species, upset local foodchains, and cause severe economic and ecological damage. It is estimated that zebra mussels and quagga mussels together have caused more than 1.5 billion dollars damage between 1989 and 2007 in the U.S. alone.

One of the major pathways for the introduction and subsequent transport of NIS is in ballast water. Efforts to reduce the number of NIS in ballast water include regulations and discharge standards at state, federal, and international levels. Generally regulations call for either mid-ocean exchange of ballast water or treatment by an approved ballast water treatment (BWT) system prior to discharge. The discharge criteria vary among the different regulatory agencies, but all criteria define extremely sparse concentrations for all size classes. Criteria range from non-detectable to less than 10 live organisms/cubic meter (m^3) for the class greater than 50 µm and from less than 0.1 to less than 10 organisms/milliliter (mL) for the class greater than 10 µm but less than 50 µm. While these objectives are a significant step forward in reducing the number of introductions of NIS, they present a substantial and complicated challenge in terms of sampling and enumeration to achieve sufficient statistical rigor and confidence in both regulatory measurements and testing of treatment systems for the evaluation of performance.

In order for BWT systems to be approved by the regulating agencies, the technologies must be tested to verify that they can treat ballast water to the level of the proposed discharge standards. Several land-based facilities are currently operating or under construction that will test BWT systems in a standardized manner using rigorous test protocols. These protocols include the International Maritime Organization (IMO) G-8 Guidelines for Type Approval of Ballast Water Management Systems and the U.S. Environmental Protection Agency's Environmental Technology Verification (ETV) Program's Protocol for the Evaluation of Ballast Water Treatment Systems. Both protocols specify taking whole water samples of at least 2 m^3 for organisms greater than 50 µm and at least 1 liter for organisms greater than 10 µm but less than 50 µm.

Early work performed on natural and augmented water samples at the Naval Research Laboratory in Key West showed considerable variability in estimates of organism concentrations because the analysis methodologies were at or near their detection limits. While test precision measurements ranged from 15.0% to 30.0% in terms of the variability about the mean, accuracy was indeterminate since the actual concentration was unknown. There were cases of large standard deviations, and enumeration was a potential source of the variability. A further finding was that enumeration and evaluation had to be completed within 6 – 8 hours of the sample being taken to avoid sample degradation.

The present study was conducted to make a baseline assessment of the current test protocol accuracy and precision and to assess how these might be optimized by either sample concentration or increased analytical samples. Specifically, this work was intended to evaluate the accuracy and precision of analyses of the two size classes (a) and (b) above. Enumeration methods used were from the IMO G-8 protocol and the ETV protocol. The concentrations of the test matrices were based on the IMO D-2 discharge standard, and the sample sizes were from the IMO G-8 and ETV protocols.

In order to simplify the experiment and eliminate bias, appropriately sized inert polymer beads were used in place of organisms. Artificial sea water replaced natural sea water with its attendant organic and inorganic particles. The beads in the test solutions were prepared from a traceable stock solution. As a result, solutions could be prepared with known concentrations for subsequent assessment of accuracy. Standard microscope procedures were used in the evaluations. For each analytical sample density, the percent difference of the observed mean from the expected mean was used as a measure of enumeration accuracy, while the coefficient of variation (CV) (ratio of the standard deviation to the mean) was related to precision. A percent difference of 10% and a CV=0.2 (20%) were used for benchmarks for acceptable accuracy and precision, respectively.

For the greater than 10 µm but less than 50 µm sized class (phytoplankton), beads representing densities of 1 – 1000 organisms per milliliter (mL) were placed in 1-mL Sedgwick Rafter slides, examined under a microscope, and tallied. Accuracy was assessed by comparing tallied counts to expected concentrations. Operator bias was evaluated by comparing results from two different operators. Comparison of analysis results of the full 1-mL Sedgwick Rafter sample to results from one half the Sedgwick Rafter sample indicated no significant difference. The number of complete 1-mL slides that could be enumerated in 8 hours was determined.

For the greater than 50 µm size class (zooplankton) a discharge concentration of 10 organisms/m^3 was used. Individual beads were counted to prepare the primary samples from which sub-samples were taken for analysis. Concentrations ranging from 20 to 120 beads/liter represented concentrations based on filtering 1 – 6 m^3. Sub-samples were examined microscopically and number of beads counted.

Several statistical tests were run on data from these simplified experiments. For the greater than 10 µm but less than 50 µm size class (phytoplankton), it was found that evaluating one-half of the Sedgwick Rafter slide was not statistically different from analyzing the full slide. That means that up to twice as many samples could be analyzed before samples degrade. Thus, for a given test of a BWT system, more sub-samples representing a greater volume could be analyzed thereby providing better statistical rigor. In general, an increasing concentration and increasing sub-samples analyzed resulted in improved precision, though similar trends in accuracy were not as apparent. Improvements in the accuracy of the observed mean were not achieved after five replications for concentrations above 50 beads/mL.

For the greater than 50 µm size class (zooplankton), although adequate accuracy (within 10%) was achieved for the most concentrated solutions (20, 60 and 120 beads/mL) after sufficient 1-mL aliquots were examined to result in a stable mean, precision remained inadequate at each of these same concentrations. The sample densities represented concentration factors of 1000, 3000 and 6000 for a total whole water sample volume of 6 m^3 with an assumed discharge concentration equivalent to the IMO D-2 standard (<10 organisms/m^3). Current test protocols designed to sample discharge concentrations at or below 10 organisms/m^3 with triplicate 1 m^3 samples or duplicate 2 m^3 samples (as proposed in the most recent version of the ETV protocol) will result in inadequate precision even when concentrated by 6000 times with this enumeration method. On the basis of these data, concentration of 6 m^3 by 6000 times and 450 1-mL aliquots analyzed would provide improved precision but would remain below the 20% (as CV) identified as the benchmark value.

Major conclusions from this work include the following:

Phytoplankton Size Organisms (more than 10 but less than 50 µm) –

- Samples should be concentrated by at least a factor of five and at least four replicate Sedgwick Rafter slides should be counted for accuracy and precision.
- Evaluation of 10 random rows of the sample slide is statistically the same as full evaluation of all 20 rows.

Zooplankton Size Organisms (larger than 50 µm) -

- The ETV protocol recommendations for samples size will result in inadequate precision. This study recommends a sample size of 6 m^3 concentrated by a factor of 6000 with analysis of 450 1-mL aliquots to achieve an arbitrary 20 per cent coefficient of variation.

Caution –

- This study was conducted under "ideal conditions". The conclusions and analyses should therefore be considered cautiously when preparing to work with natural assemblages of organisms and in natural water.

(This page intentionally left blank.)

TABLE OF CONTENTS

EXECUTIVE SUMMARY ..... **v**

LIST OF FIGURES ..... **x**

LIST OF TABLES ..... **xi**

LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS ..... **xii**

1 Introduction **1**

1.1 Background 1
1.2 Objectives 2

2 Experimental Approach **2**

3 Materials & Methods **5**

3.1 Equipment, Sample and Subsample Preparation 5
3.1.1 Simulated Phytoplankton Stock 5
3.1.2 Simulated Zooplankton Stocks 6
3.2 Enumeration 7
3.2.1 Simulated Phytoplankton Samples 7
3.2.2 Simulated Zooplankton Samples 7
3.3 Comparative Analysis 8
3.3.1 Accuracy of Initial Sample Stock: % Difference 8
3.3.2 Sampling Statistics 8
3.3.2.1 Observed vs. Expected Counts 9
3.3.2.2 Mean and Standard Deviation 9
3.3.2.3 Analysis of Variance 9
3.3.2.4 Whole Chamber Counting vs. Partial Chamber Counting 10
3.3.2.5 Effect of Observers 10

4 Results and Discussion **11**

4.1 Accuracy of Initial Sample Phytoplankton Stock Preparations 11
4.2 Accuracy and Precision of Enumeration 11
4.2.1 Phytoplankton Enumerations – Mean/Standard Deviation vs. Sample Size/Concentration 11
4.2.2 Zooplankton Enumerations – Mean/Standard Deviation vs. Sample Size/Concentration 15
4.2.3 Observed vs. Expected Counts 17
4.2.4 Accuracy Evaluated as Percent Difference 23
4.2.5 Precision Evaluated as Coefficient of Variation 23
4.2.6 Statistical Tests for Sample Concentration and Volume Differences 25
4.2.6.1 Phytoplankton 25
4.2.6.2 Zooplankton 26
4.3 Distribution Analyses for Simulated Phytoplankton 27
4.3.1 Whole Versus Partial Phytoplankton Counting Chamber Enumeration 30
4.3.2 Differences Between Observers 1 & 2 for Sub-samples 32

5 Conclusions **34**

5.1 Overall Conclusions/Recommendations 36

6 References **37**

LIST OF FIGURES

Figure 1. Diagram of modified ETV protocol sampling methodology for phytoplankton. 3
Figure 2. Diagram of modified ETV protocol sampling methodology for zooplankton. 4
Figure 3. Fluorescent, inorganic beads manufactured from fluorescent polymer microspheres. 5
Figure 4. Chromosphere beads; the red colored bead was used for these experiments. 6
Figure 5. Phytoplankton matrix (1 bead/mL, 20 rows counted). 12
Figure 6. Phytoplankton matrix (5 beads/mL, 20 rows counted). 12
Figure 7. Phytoplankton matrix (10 beads/mL, 20 rows counted). 13
Figure 8. Phytoplankton matrix (50 beads/mL, 20 rows counted). 13
Figure 9. Phytoplankton matrix (100 beads/mL, 20 rows counted). 14
Figure 10. Phytoplankton matrix (500 beads/mL, 20 rows counted). 14
Figure 11. Phytoplankton matrix (1000 beads/mL, 20 rows counted). 15
Figure 12. Zooplankton matrix (1 bead/50 mL). 16
Figure 13. Zooplankton matrix (3 beads/50 mL). 16
Figure 14. Zooplankton matrix (6 beads/50 mL). 17
Figure 15. Phytoplankton matrix (1 bead/mL, 10 and 20 rows counted). 18
Figure 16. Phytoplankton matrix (5 beads/mL, 10 and 20 rows counted). 18
Figure 17. Phytoplankton matrix (10 beads/mL, 10 and 20 rows counted). 19
Figure 18. Phytoplankton matrix (50 beads/mL, 10 and 20 rows counted). 19
Figure 19. Phytoplankton matrix (100 beads/mL, 10 and 20 rows counted). 20
Figure 20. Phytoplankton matrix (500 beads/mL, 10 and 20 rows counted). 20
Figure 21. Phytoplankton matrix (1000 beads/mL, 10 and 20 rows counted). 21
Figure 22. Zooplankton matrix: (10 bead per 50 mL). 21
Figure 23. Zooplankton matrix: (30 bead per 50 mL). 22
Figure 24. Zooplankton matrix: (60 bead per 50 mL). 22
Figure 25. Comparison of the coefficient of variation (CV) at a range of discharge concentrations in relation to volume sampled (Number of Sedgwick-Rafter slides counted). 24
Figure 26. A closer comparison of the coefficient of variation (CV) for the highest simulated phytoplankton sample concentrations analyzed. 25
Figure 27. Zooplankton matrix: comparison of the coefficient of variation (CV) to volume sampled for three concentrations (20, 60 and 120 beads/mL). 27
Figure 28. Variance versus mean for simulated phytoplankton enumerations. 28
Figure 29. Observer comparison for high microbead concentrations. 29
Figure 30. Comparison of the coefficient of variation (CV) relative to volume sampled.. 31
Figure 31. Comparison of the coefficient of variation (CV) relative to volume sampled for concentrated samples. 31

LIST OF TABLES

Table 1. Test matrix for phytoplankton discharge concentrations & possible sample volumes for filtering into a 1-L flask. 4
Table 2. Test matrix for zooplankton investigations with variations in sample volume and concentration factor. 5
Table 3. The absolute percent difference calculated for counts performed for initial vendor stock assessments. 11
Table 4. Comparison of accuracy in terms of absolute percent difference of the observed from the expected concentration for simulated phytoplankton concentrations. (na = not applicable). 23
Table 5. Phytoplankton matrix (20 rows counted): levene test of equal variances. 95 percent confidence (α=0.05). 25
Table 6. ANOVA test results comparing four bead concentrations (50, 100, 500, 1000) between sample volumes of 2 ml to 6 mls analyzed. 26
Table 7. Phytoplankton matrix (20 rows counted): ANOVA tukey post hoc test of multiple comparisons.. 26
Table 8. Simulated Poisson distribution and corresponding chi-square percentiles. 28
Table 9. Comparison of the test statistic to the chi-square percentile. 29
Table 10. Comparison of accuracy in terms of percent difference for simulated phytoplankton stocks with only 10 random rows counted in the counting chamber. 30
Table 11. Phytoplankton matrix (10 rows counted): Levene test of equal variances with 95 Percent confidence (α=0.05). 32
Table 12. Comparison of CV's between the 20 row and 10 row counts for the 50, 100, 500, and 1000 bead concentrations. 32
Table 13. Phytoplankton matrix (20 rows counted): Independent T-test test results comparing observer counts for seven bead concentrations. 33
Table 14. Zooplankton matrix: ANOVA test results comparing observer counts for three bead concentrations. 33

LIST OF ACRONYMS, ABBREVIATIONS, AND SYMBOLS

ANOVA	analysis of variance
APHA	American Public Health Association
BWT	ballast water treatment
BWTE	ballast water treatment equipment
df	degrees of freedom
EPA	Environmental Protection Agency
ETV	Environmental Technology Verification
F	Calculated random-effects analysis of variance
F crit	Critical random-effects analysis of variance values
gal	gallon
ID	identification
IMO	International Maritime Organization
L	liter
m	meter
mean	measure of central tendency
mg	milligram
mL	milliliter
mm	millimeter
MS	mean square
na	not available
NISA	National Invasive Species Act
NIST	National Institute of Standards and Technology
NRLKW	Naval Research Laboratory in Key West, Florida
NRL	Naval Research laboratory
NSF	National Sanitation Foundation
pH	potential of Hydrogen
ppm	parts per million
p-value	the probability of observing a test statistic that is as extreme or more extreme than currently observed (assuming that the null hypothesis is true).
sig	significant
SPSS	Statistical Package for Social Sciences, software created by SPSS Inc.
SS	sum of squares
UK	United Kingdom
USCG	United States Coast Guard
^{o}C	degrees Celsius
µL	microliter
µm	micron
#	number
®	registered sign
α	alpha at 95% confidence interval

1 INTRODUCTION

1.1 Background

The invasion of non-indigenous species (NIS) into aquatic ecosystems has been determined to be a significant ecological and financial problem. Moreover, it has been established that ballast water of commercial vessels is a primary mechanism for movement of these organisms. Many efforts, including technology and policy related efforts, have been undertaken or are underway to mitigate the movement and impact of NIS on ecosystems. In particular, various international, national and state discharge standards and regulations to limit the discharge of organisms in ships' ballast water are either developed, proposed or are under consideration. Each of the proposed standards identifies the discharge criteria on the basis of organism size classes, defining groups as either a) greater than 50 microns (µm), b) greater than 10 µm and less than or equal to 50 µm and c) less than or equal to 10 µm. In each of these size classes, discharge concentration criteria have been specified by various regulatory agencies such that a range of concentrations have been identified. These range from "non-detectable" to less than 10 organisms/m^3 in the 50 µm size class and less than 0.01 to less than 10 organisms/mL for the greater than 10 µm but less than or equal to 50 µm class. These concentrations are very low and will be very difficult to measure with statistical rigor and confidence.

The proposed discharge standards and the ability to measure organism concentrations with adequate accuracy and precision is of significant concern to land-based test facilities of ballast water treatment systems. There are a number of such test facilities which are either in operation or under construction. Their objective will largely be to test treatment systems in a standardized manner to assess the system's ability to routinely achieve the proposed discharge standards. Test facilities are currently utilizing the test protocols (G8) provided by the IMO Guidelines for the Approval of Ballast Water Management Systems or the US Environmental Protection Agency's Environmental Technology Verification (ETV) Protocol for the Verification of Ballast Water Treatment Systems. In each of these test protocols, sample methodologies, designs and volumes are provided for the measurement of organism concentration in the relevant size categories. However, as these protocols and methodologies were derived based on the recommendations of subject matter experts without empirical data or an opportunity for validation, many of these protocols require further evaluation, particularly where non-standard methods are utilized. The Naval Research Laboratory in Key West (NRLKW) in collaboration with the USCG Environmental Standards Division and the USCG Research & Development Center has conducted experiments to test and validate the ETV protocol as a whole and in part. These studies have included the development of test methods, test systems and the use of non-standard test methods.

In this regard, the USCG Research & Development Center commissioned NRLKW to investigate the sampling and enumeration of organisms representing sparse populations as would be expected in treated water intended to meet the proposed discharge standards. The test protocols involve the treatment and collection of large volumes of water (~ 200-300 m^3) from which 1 m^3 to 2 m^3 samples are drawn at various points in the overall test process including before treatment and storage in a simulated ballast tank and upon discharge from the tank. Sub-samples are then taken for the enumeration and live/dead classification of organisms by size class. For organisms larger than 50 µm, the entire sample is concentrated through a 35 micron plankton net into a cod end, then resuspended in a 1 Liter flask to produce a concentrated sample. The 35 µm net has diagonal openings of just under 50 µm and therefore captures 50 µm organisms. For the

size class greater than 10 µm but less than or equal to 50 µm, a 1-L sub-sample of the filtrate (that which passes through the 35 µm net) is acquired for analyses without sample concentration.

During previous testing and validation efforts by NRLKW, it was determined that the sample sizes and analytical methods have resulted in imprecise measurements. From these results, it was evident that more sample concentration and/or larger sample quantities may be necessary to achieve adequate precisions and accuracy. Determining appropriate sample volumes can be affected by factors such as: the species type or size class being analyzed, organism concentration, method of analysis, and observer effects when manual counts are being performed. In addition, the time required to perform the analyses can affect the number of samples which may be analyzed. Since manual counts are time intensive, sample water quality may degrade due to decreasing oxygen levels and the degradation of organic matter resulting in an unintended reduction in viable organism populations. As a result, there is a limited time period in which samples must be analyzed. All of these factors need to be considered and tested in various combinations to determine a sample regime that will optimize the accuracy of sample analysis.

1.2 Objectives

The overarching objectives for the work reported here was to determine for the enumeration methods currently in use under the ETV protocol what sample sizes and subsequent sample concentration are appropriate and achievable for the two size classes defined above. Through the use of inorganic beads of appropriate size, the study was intended to evaluate the accuracy and precision of current enumeration methods over a range of concentrations and sample volumes. Specifically, this research was to result in:

- Optimization and demonstration of methods for the enumeration of phytoplankton and zooplankton in sparse populations
- Determination of the amount of sample concentration necessary to perform analyses that result in scientifically defensible estimates which are representative of the sampled discharge.

2 EXPERIMENTAL APPROACH

To measure the accuracy and precision of the enumeration methods, it was necessary to remove as many variables as possible that could introduce bias and confound the primary analysis. By using inorganic beads, the change in sample make-up due to the death and degradation of organisms was prevented. In addition, with both natural and culture populations, the actual sample concentration is unknown; with the use of inorganic beads the sample concentration could be accurately predicted within the variance supplied and reported by the manufacturer. Finally, the inhomogeneity of natural samples was eliminated through the use of spherical beads. As a result of homogeneity and lack of sample degradation, the best enumeration scenario of those examined was determined. Overall, the enumeration conducted in such highly controlled experiments was expected to provide benchmarks of enumeration accuracy and precision without the effects of full-scale system sampling operations.

Experimental concentrations were based on the allowable concentration of organisms in the discharge of a treatment system that had successfully treated to the IMO D-2 discharge standard. For phytoplankton, concentrations both above and below this standard were added and considered to provide data for the determination of the effect of concentration on measurement accuracy and precision and to permit identification of the appropriate organism concentration which should be achieved for optimum enumeration. A similar approach was taken for zooplankton populations. In this case, however, the range

of concentrations was based on only the IMO D-2 standard and was varied on the basis of sample volume and maximum ability to concentrate the sample, as described later.

Two test matrices were developed, one to represent zooplankton or organisms > 50 µm and one for phytoplankton or organisms greater than 10 µm and less than or equal to 50 µm. The phytoplankton matrix was constructed to represent 1) a range of discharge organism concentrations, as discussed above, and 2) a realistic range of sample volumes which would subsequently be concentrated to 1 L.

Figure 1 provides a schematic for the overall phytoplankton sampling methodology utilized in the ETV protocol. Samples are acquired from the ballast water discharge with a concentration of organisms of either 1, 10, or 100 organisms/mL. The sample is assumed to be homogeneous and representative of the discharge. A 3 m^3 sample volume is acquired on a time average basis. In the case of phytoplankton, no sample concentration is performed during this step and the sample concentration as acquired is equal to the discharge concentration. Finally, sampling enrichment may occur via the filtering of the entire sample through a plankton net mesh, though this is not currently called for in the ETV protocol. The cod end is then rinsed into a 1 Liter flask for subsequent sub-sampling and analysis. Larger sample volume effects were evaluated by selecting large volumes which would result in concentration ratios of 1, 5 and 10. The resultant test matrix of discharge concentration versus sample volume is shown in Table 1.

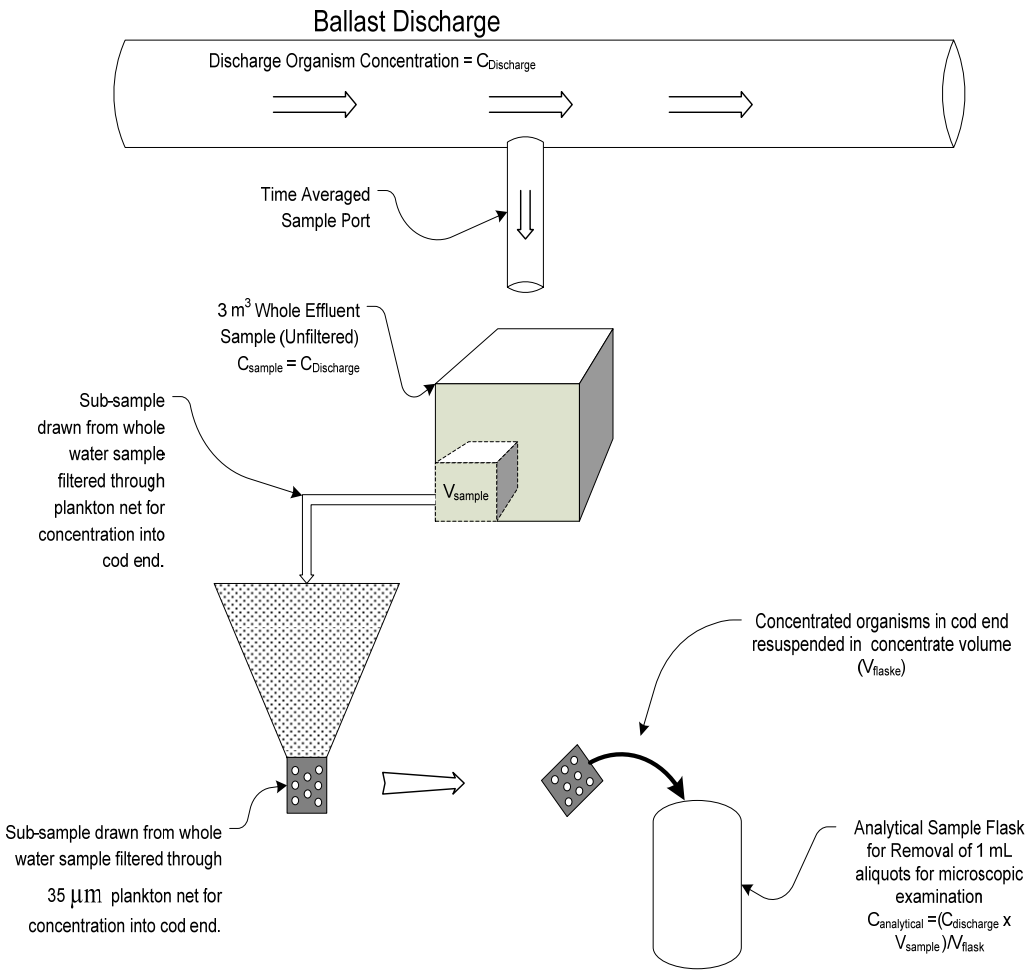

Figure 1. Diagram of modified ETV protocol sampling methodology for phytoplankton.

Table 1. Test matrix for phytoplankton discharge concentrations & possible sample volumes for filtering into a 1-L flask.

Discharge Concentration→		1 bead/mL	10 beads/mL	100 beads/mL
Sample Volume (V_{sample})	Concentration Factor	Analytical Sample Concentration (beads/L)		
1 Liter	1	1.00E+03	1.00E+04	1.00E+05
5 Liters	5	5.00E+03	5.00E+04	5.00E+05
10 Liters	10	1.00E+04	1.00E+05	1.00E+06

The zooplankton methodology cited in the ETV protocol calls specifically for the entire sample to be concentrated using plankton net filtration as shown in the diagram of Figure 2. As with phytoplankton, zooplankton enumeration involved the hypothetical collection of samples from the in-line flow stream. During normal operations, the whole water sample is collected and filtered through a plankton net, and the cod end is then rinsed into a flask for sub-sampling and analyses. The concentration of the resulting sample in the flask depends on initial concentration in the water, the volume of water filtered through the plankton net, and the final volume in the flask. In this case, the matrix (shown in Table 2) provided the opportunity for an evaluation of the effect of larger sample volumes and a variety of concentrate volumes, thereby resulting in larger concentration ratios than normally sampled. The matrix assumes that the discharge concentration will be 10 living organisms (beads)/m^3. There were three different hypothetical sample volumes collected: 1, 3 and 6 cubic meters. All volumes were concentrated in their entirety and then processed to represent three final volumes: 500 mLs, 1 liter or 3 liters. Figure 2 shows the process of sample collection and concentrating represented by the different total bead to volume ratios in the zooplankton test matrix.

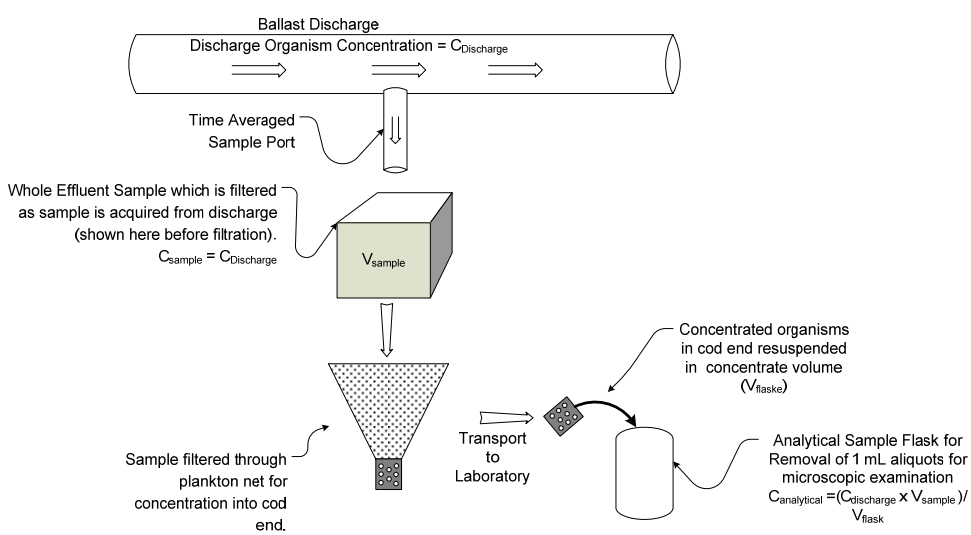

Figure 2. Diagram of modified ETV protocol sampling methodology for zooplankton.

Table 2. Test matrix for zooplankton investigations with variations in sample volume and concentration factor.

Whole Water Sample Volume→		1 m^3	3 m^3	6 m^3
Concentrate/Flask Volume (V_{flask})	Concentration Factor	Analytical Sample Concentration (beads/L)		
3 Liter	1000	3.33	10.00	20.00
1 Liter	3000	10.00	30.00	60.00
0.5 Liter	6000	20.00	60.00	120.00

Finally, stocks of the concentrated solutions identified in Table 1 and Table 2 were prepared and analyzed using the appropriate enumeration method. Analyses and enumeration were conducted on multiple replicates or sub-samples of these solutions as required to either count 300 organisms, cumulatively, or until a six hour time limit was reached. The latter represents the practical limit for enumeration of living organisms without fixatives or sample degradation as determined by the NRL-KW test facility during previous validation work. These results were then compiled and analyzed for both sample accuracy and precision as a function of sample replicates and concentrate volume.

3 MATERIALS & METHODS

3.1 Equipment, Sample and Subsample Preparation

3.1.1 Simulated Phytoplankton Stock

All stock preparation and sampling occurred at the bench top level. A total number of beads representing the phytoplankton class were added to a specific volume of artificial seawater. The beads were purchased from Duke Scientific Corporation/ Thermo Scientific and were fluorescent, polymer, microsphere beads (Figure 3) which came in a 15 milliliter aqueous solution. Their nominal diameter was 10 microns with a measured mean diameter of 9.9 µm. The vendor stock concentration of beads was 1.9 x 10^7 beads per milliliter with 1% solids. The vendor specifications indicated that the aqueous suspension of the beads had a multi-component dispersing system which prevented clumping and aided in the dispersion of the beads.

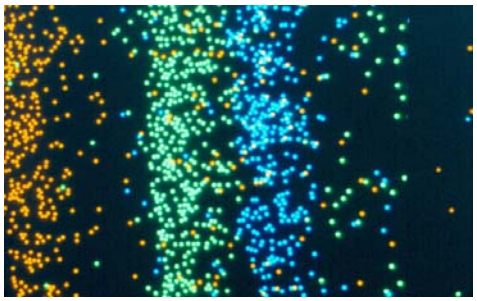

Figure 3. Fluorescent, inorganic beads manufactured from fluorescent polymer microspheres. (Photograph indicates fluorescent colors available. Courtesy of Duke Scientific Corporation)

The stock bottle was mixed and multiple drops then dispensed into a small alternate container. This volume was then sampled using a scientific grade Eppendorf micro-pipettor which was able to draw the stock sample in microliters (µL). Calculations were made to determine the volume of stock required to obtain the desired number of spheres for each sample. This volume was then dispensed into a 1000 mL volume graduated Erlenmeyer flask that had been filled to the 1000 mL mark with artificial seawater (33 ppt salinity (parts per thousand)). Based on the stock volume calculations for 10,000 beads per mL, the multiple bead concentrations for the phytoplankton matrix were created as follows:

0.0526 µL of stock into 1000 mLs artificial seawater = 1 bead per mL
0.2630 µL of stock into 1000 mLs artificial seawater = 5 beads per mL
0.5260 µL of stock into 1000 mLs artificial seawater = 10 beads per mL
2.630 µL of stock into 1000 mLs artificial seawater = 50 beads per mL
5.26 µL of stock into 1000 mLs artificial seawater = 100 beads per mL
26.3 µL of stock into 1000 mLs artificial seawater = 500 beads per mL
52.6 µL of stock into 1000 mLs artificial seawater = 1000 beads per mL

A flask stopper was used to cap the flask so that it could be well mixed by inverting it five times for a count of five seconds each. Once the flask contents were well mixed, the stopper was removed for subsampling.

A scientific Eppendorf pipette, set to draw a 1-mL volume, was then used to subsample the 1000 mL flask. The subsample was dispensed onto a Sedgwick-Rafter counting slide. This slide was marked with a twenty by fifty grid of 1 µL volume squares which results in the entire 1 mL sample being sectioned into 1000 sample units. The coverslip was set into place over the slide at which time the slide contents were allowed to settle for twenty minutes. Enumeration of samples is described in Section 3.2.

3.1.2 Simulated Zooplankton Stocks

The bead type chosen for the zooplankton matrix was also purchased from Duke Scientific/ Thermo Scientific Corporation as Chromosphere-T certified, colored polymer microspheres with NIST (National Institute of standards an Technology) traceable diameters (Figure 4). The beads came in a one gram bottle; their nominal diameter was 150 µm with a certified mean diameter (+/- uncertainty) of 149 µm +/- 3.6 µm. (Note scale in figure.)The vendor's approximate bead count per gram of stock was 5.5 x 10^5 beads.

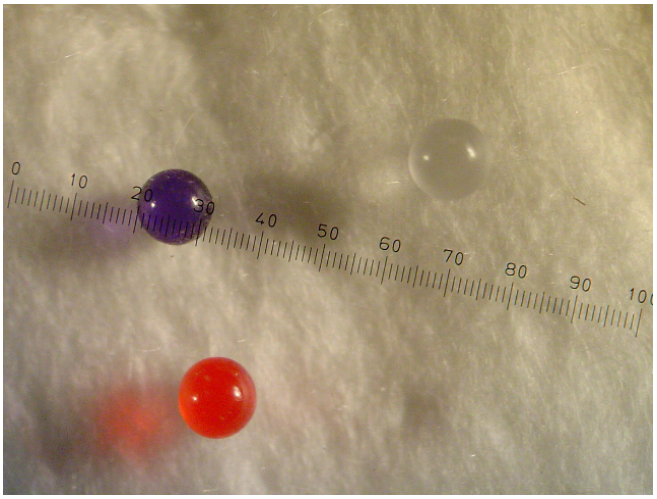

Figure 4. Chromosphere beads; the red colored bead (150 µm) was used for these experiments.

The vendor specifications stated that these beads could be dispersed in aqueous media with the aid of a small amount of surfactant or in lower alcohols such as methanol and ethanol. Ethanol was used as the dispersant for these experiments. Because the beads were packaged dry, a small amount was taken out, dipped in the ethanol, and then suspended into a multiwell plate filled with ASTM Type I deionized water. The well was then placed under a lighted dissection stereoscope. With the help of magnification, beads were individually captured and transferred until the required number of 10, 30 or 60 beads were contained in separate wells. For each concentration, a wide-mouthed, graduated PYREX, 500 mL capacity flask was filled with artificial seawater (33 ppt salinity) to the 500 mL level. The separated beads were then added to the appropriate flasks. Based on the bead concentrations and flask volumes, the multiple bead test concentrations were created as follows:

10 beads into 500 mLs artificial seawater = 0.02 beads per mL
30 beads into 500 mLs artificial seawater = 0.06 beads per mL
60 beads into 500 mLs artificial seawater = 0.12 beads per mL

The flask was then capped with a stopper and mixed well by inverting the flask for a count of five seconds; this was repeated five times. Once mixed, the volume was sampled using a five mL capacity, serontological, graduated plastic pipette with pipette helper. A 1-mL subsample from the flask was then dispensed into a single well of a twelve multiwell plate. This step was repeated until a series of 1-mL subsamples had been individually dispensed. Once a series of subsamples had been generated, the first plate was positioned under the dissection stereoscope for analysis.

3.2 Enumeration

3.2.1 Simulated Phytoplankton Samples

Once the twenty minute settling time expired for the phytoplankton subsamples, the Sedgwick-Rafter counting slide was placed under a contrast-phase microscope with a 20X objective and positioned to view the square in the uppermost left-hand corner. The number of beads within the square were counted and the slide was then moved horizontally along the row. Each square was counted until all fifty squares in the first row were tallied for beads. This number was recorded and represented the number of beads per row (equivalent to a 50-µL volume). The procedure was repeated until all twenty rows of the Sedgwick-Rafter counting slide were tallied. The twenty rows were then summed to arrive at the number of beads per Sedgwick-Rafter (equivalent to 1-mL volume). For each test concentration, the number of Sedgwick-Rafter slide volumes that could be counted in an eight hour period was also determined.

3.2.2 Simulated Zooplankton Samples

For the zooplankton subsamples, 276 1-mL subsamples were individually taken and dispensed into the wells of 23 12-well plates before counting commenced. As soon as the dispensing of subsamples was completed, the first plate of 12 wells was placed under the stereo dissection microscope. The zoom was adjusted so that the well took up the entire field of view. The field of view was then scanned for beads and the number recorded on the data sheet. This process continued sequentially in the order that the subsamples were initially dispensed. After the initial 276 mL had been counted, the remaining test volume was dispensed into multi-well plates and counted. As with phytoplankton, the goal with each test concentration was to count as many 1-mL subsamples as could be completed in an eight hour period.

3.3 Comparative Analysis

3.3.1 Accuracy of Initial Sample Stock: % Difference

Prior to the creation of the simulated phytoplankton concentrate stocks as outlined in Section 3.1.1, the ability to accurately produce sample concentrations using the vendor supplied bead concentrates was examined. As discussed, vendor bead stock was provided in a 15-mL, aqueous solution with a concentration of 1.9×10^7 beads/mL with 1% solids. As such, a volume of 0.526 µL vendor stock was expected to yield 10,000 beads. This stock volume was then scaled to volumes representative of the matrix counts. For example, 1000 mL of artificial seawater was inoculated with 526 µLs of vendor stock to produce an anticipated concentration of 10,000 beads/mL. The mixture was then well mixed, sub-sampled and enumerated as described in Section 3.2.1. No time limit was placed on the enumeration; counting proceeded until the entire stock concentration had been sub-sampled and enumerated. Thus all beads within the concentrate were enumerated directly. The accuracy of the stock mixture relative to the expected concentration was determined by comparing the actual number counted (observed) to the expected number and calculating the percent difference between the two according to Equation 1. Note that the absolute value of the difference was taken to avoid negative percentages.

$$\% \text{ Difference} = ((|x_1 - x_2|)/ x_2)*100 \qquad (1)$$

where,

x_1 = is the observed mean number and
x_2 = is the expected number.

This stock preparation and counting evaluation was completed three times.

On the other hand, due to the relatively low total number of beads required for the zooplankton stock concentrations, the beads for each test were individually selected and dispensed into a separate container, thus eliminating associated vendor stock or laboratory errors. The container was then rechecked once the beads were dispensed to confirm that all were added to the test flask for analysis.

Vendor estimations of concentration accuracy for the two bead types used in this experiment are based solely on weight. In other words, concentration counts were not performed by the vendor on the stock provided. It was determined that the vendor has a product called "EZY-CAL Count Precision Standards" which is used specifically for calibration purposes regarding particle counts. The accuracy and precision of this product is stated as 2000 beads per mL +/- 10% and is based on counts of five 1-mL samples which have a minimum of 2000 beads per mL. These samples are run through a particle counter rather than observed manually using microscopes. The vendor indicated that counting concentrations lower than 2000 beads per mL resulted in variability deemed too high for their product standards. This information indicates that the best precision that the vendor expects to get for minimum counts of 2000 beads per mL is 10% and that 5 mLs of sample volume is sufficient to obtain result.

3.3.2 Sampling Statistics

Several statistical measures and/or tests were used in the analysis of the experimental data. These included: Standard Deviation (**σ**), Coefficient of Variation (CV), Analysis of Variance (ANOVA), Levene Test, Tukey Test, and T-Test. Statistical software from SPSS, Inc was used for a number of these tests.

3.3.2.1 Observed vs. Expected Counts

A comparison of the observed against expected concentration was evaluated graphically, both by a comparison of the measured sample concentration versus number of Sedgewick Rafter chambers or volume analyzed and through a comparison of observed direct counts against expected counts. In both cases, accuracy could be evaluated through comparison of the observed concentration relative to the expected benchmark. Precision was best qualitatively evaluated in the former case through the observation of the standard deviation error bars pictured graphically.

3.3.2.2 Mean and Standard Deviation

The mean and standard deviation of the bead counts were calculated using the cumulative data for each test concentration performed in both the phytoplankton and zooplankton matrices. The mean (Average) calculations were performed using EXCEL with the following equation:

$$\bar{x} = \frac{\sum_{i=1}^{n} x_i}{n} \qquad (2)$$

where

$\bar{x}$ = mean sample concentration across n observations

x_i = the ith obvserved count and

n = the total number of sub-samples analyzed.

The standard deviation, σ, was also calculated according to the following equation which represents the root mean square deviation of the observed values from the observed mean.

$$\sigma = \sqrt{\frac{\sum (x - \bar{x})^2}{n-1}} \qquad (3)$$

These calculations were then used to calculate the coefficient of variation (CV) for each of the bead concentrations using the following equation:

$$CV = \frac{\sigma}{\bar{x}} \qquad (4)$$

The CV for each bead concentration was displayed graphically to identify the effects (if any) that bead concentration and volume sampled had on CV. Any CV results greater than or equal to 0.50 (50%) were to be considered too imprecise for further statistical analysis while results less than 50% would be further evaluated.

3.3.2.3 Analysis of Variance

Analysis of Variance (ANOVA) is an analysis technique to determine if experimental results differ from each other with respect to input variables or characteristics of "Treatments." For concentrations with CV greater than 50%, an Analysis of Variance (ANOVA) with 95% confidence (α = 0.05) was utilized to evaluate whether statistically significant differences were present in measurements. Qualifying data were tested using an ANOVA with the following null hypotheses:

H_{o1} = There is no significant difference in CV with regard to bead concentrations
H_{a1} = There is a significant difference in CV with regard to bead concentrations

The analysis of variance is considered a robust test for comparing data sets when the data used have equal variances and the sample sizes are equal or nearly equal. The Levene statistical test is used to determine if two populations (or samples) have substantially similar or different variances. The variances of the different bead concentrations were tested using the Levene test to determine if an ANOVA was an appropriate statistical test to employ. The null hypothesis for this test is as follows:

H_{o2} = The error variance of the dependent variable is equal across groups.
H_{a2} = The error variance of the dependent variable is NOT equal across groups.

In addition to the ANOVA, the multiple range Tukey test was employed to provide *post hoc*, multiple comparisons of the bead concentrations and to identify which means varied significantly from each other. The null hypothesis is that all means are equal.

3.3.2.4 Whole Chamber Counting vs. Partial Chamber Counting

All of the comparisons described thus far involved Sedgwick Rafter slides that had had all 20 rows counted. Another comparison was made to address the question as to the necessity of counting all 20 rows. Specifically, the question was whether the same variability could be achieved if only 0.5 mL of the 1-mL volume were randomly chosen and counted? To test this, 10 of the 20 Sedgwick-Rafter rows were randomly chosen and the number of beads summed. The mean, standard deviation, and CV were then calculated as described in Section 3.3.2.2. The percent difference was also calculated as described in Section 3.3.1 except that the observed numbers were multiplied by 2 to reflect number of beads per 1 mL rather than 0.5 mL. These 1-mL increments were then used to compare to the 20 row calculations. To determine if there were significant differences between counting 10 and 20 rows, the CV's for each 10 row volume was calculated for the bead concentrations and then compared using an Independent Samples T-Test with 95% confidence ($\alpha = 0.05$). The null hypotheses for the comparisons were:

H_{o3} = There is no significant difference in CV with regard to bead concentrations between the 20 row and 10 row counts.
H_{a3} = There is a significant difference in CV with regard to bead concentrations between the 20 row and 10 row counts.

3.3.2.5 Effect of Observers

For both the phytoplankton and zooplankton test matrices, a comparison between observers was done to determine if there were significant differences in bead counts. One observer was the primary counter and performed a full day of counts. The second observer was brought in for QA/QC purposes and performed counts on at least half of the Sedgwick-Rafter slides counted by observer #1. The counts per Sedgwick-Rafter slide per observer were then compared using an Independent Samples T-Test with 95% confidence ($\alpha = 0.05$). The null hypothesis was:

H_{o4} = Determinations of bead concentration do not differ among observers when 20 rows are counted.
H_{a4} = Determinations of bead concentration differ among observers when 20 rows are counted.

(Note that there are two factors in the above experiment – counts per slide and number of slides counted. Observer #1 is more subject to fatigue during a full day of counting than is observer #2. There is no way to partition out this factor.)

4 RESULTS AND DISCUSSION

4.1 Accuracy of Initial Sample Phytoplankton Stock Preparations

The percent difference for sample stock was always under 10.0% (Table 3). The results shown coincide with the 10% precision specified by the vendor for their calibration beads. When performing counts, the vendor starts with a minimum of 2000 beads per mL^{-1} and considers anything less than that too variable. The count of 10,000 beads per mL counted in this study was well above the 2000 bead per mL minimum required for the vendor counts. The results indicate that a 10% systematic error needs to be considered when making bead concentrations from Duke Scientific/Thermo Scientific microbead stocks. Thus a 10,000 bead per mL count could be a minimum of 9000 beads to a maximum of 11000 beads per mL.

Table 3. The absolute percent difference calculated for counts performed for initial vendor stock assessments.

Sample Stock Number	Sedgwick-Rafter Slide counted (= 1 mL)	Total Number of Beads Counted	% Difference (absolute value)
1	1	10412	4.1
1	2	10263	2.6
2	1	9723	2.8
2	2	10486	4.9
3	1	9781	2.2
3	2	9143	8.6
3	3	9098	9.0

4.2 Accuracy and Precision of Enumeration

4.2.1 Phytoplankton-Sized Bead Enumerations – Mean & Standard Deviation vs. Sample Size & Concentration

The test matrix for phytoplankton-sized beads had a total of seven different test concentrations of beads. Each test flask was subsampled in one milliliter volumes over the duration of eight hours. For the phytoplankton-sized bead concentrations of 1, 5, 10 and 50 beads per mL, nine slides were fully counted for a total of 9 mL volume sampled per bead concentration. Six slides were fully counted for the 100, 500 and 1000 bead concentrations for a total of 6 mL volume sampled per bead concentration. Plots of the observed means as a function of the volume of sub-samples analyzed (or in this case, the number of Sedgwick-Rafter slides counted) are shown in Figure 5 through 11. The associated error bars reflect one standard deviation in the sample counts. The red line reflects the expected concentration of the prepared stock. Each of these can be preliminarily evaluated for accuracy by comparison of the observed mean to the expected concentration (red line). The precision can be estimated by the width of the associated error bars.

Overall, an increasing bead concentration and increasing sub-samples analyzed resulted in reduced standard deviation as expected. Similar trends in accuracy, however, were not as apparent in that the mean did not consistently improve relative to an increase in expected bead concentration or volume sampled.

Figure 5 and Figure 6 show the results for the 1 and 5 bead concentrations, respectively, and these demonstrated the greatest variability in terms of accuracy and precision. Furthermore, clear relationships between decreased standard deviation in relation to increased volume sampled were inconsistent. While the standard deviation at any given volume analyzed was relatively constant across all sample concentrations, the most significant feature is the CV which will be discussed in a later section.

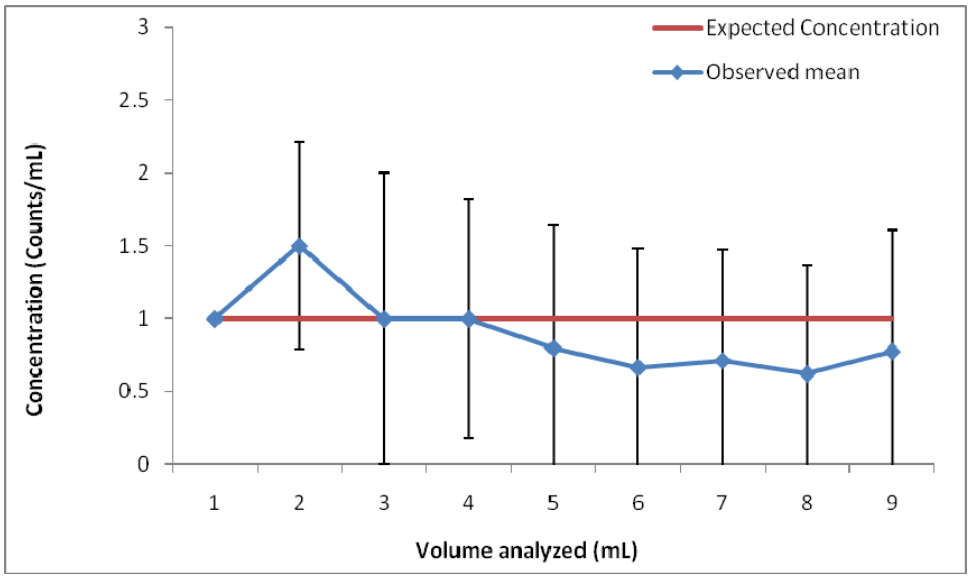

Figure 5. Phytoplankton matrix (1 bead/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 1 bead per mL concentrations.

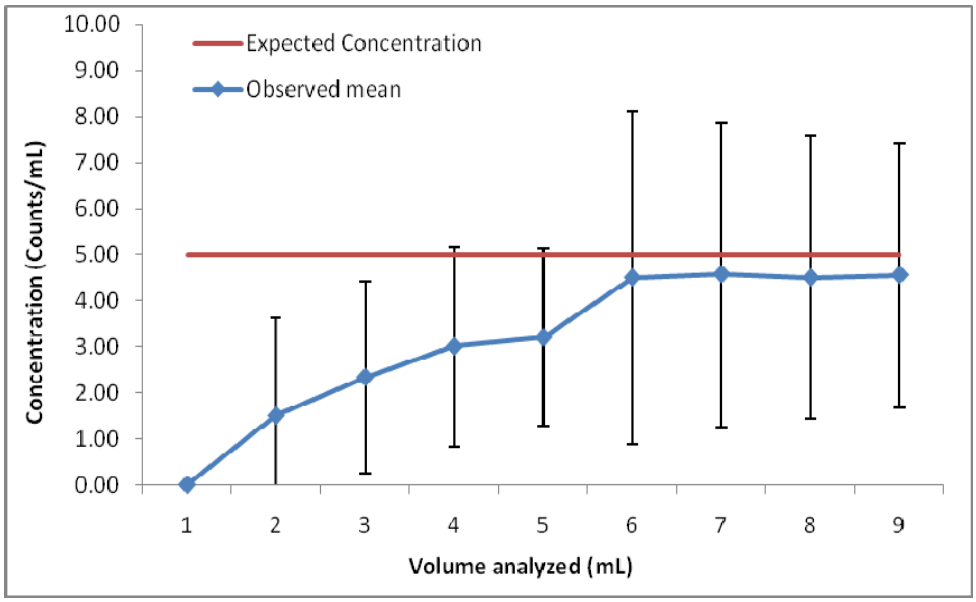

Figure 6. Phytoplankton matrix (5 beads/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 5 beads per mL concentrations.

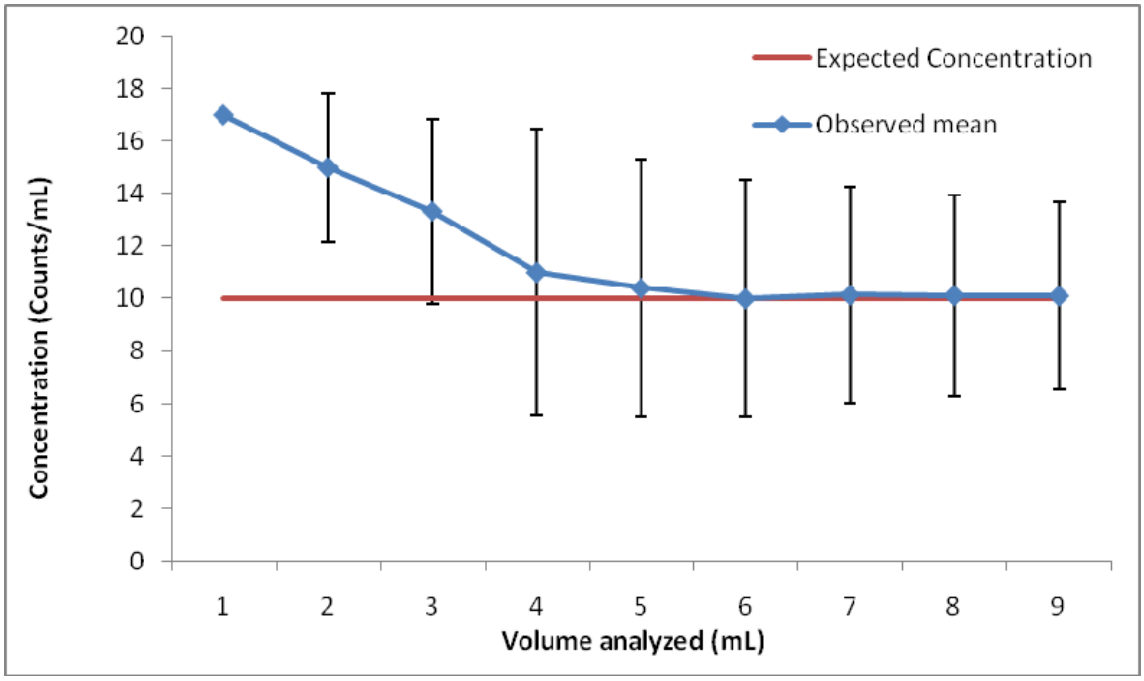

Figure 7. Phytoplankton matrix (10 beads/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 10 beads per mL concentrations.

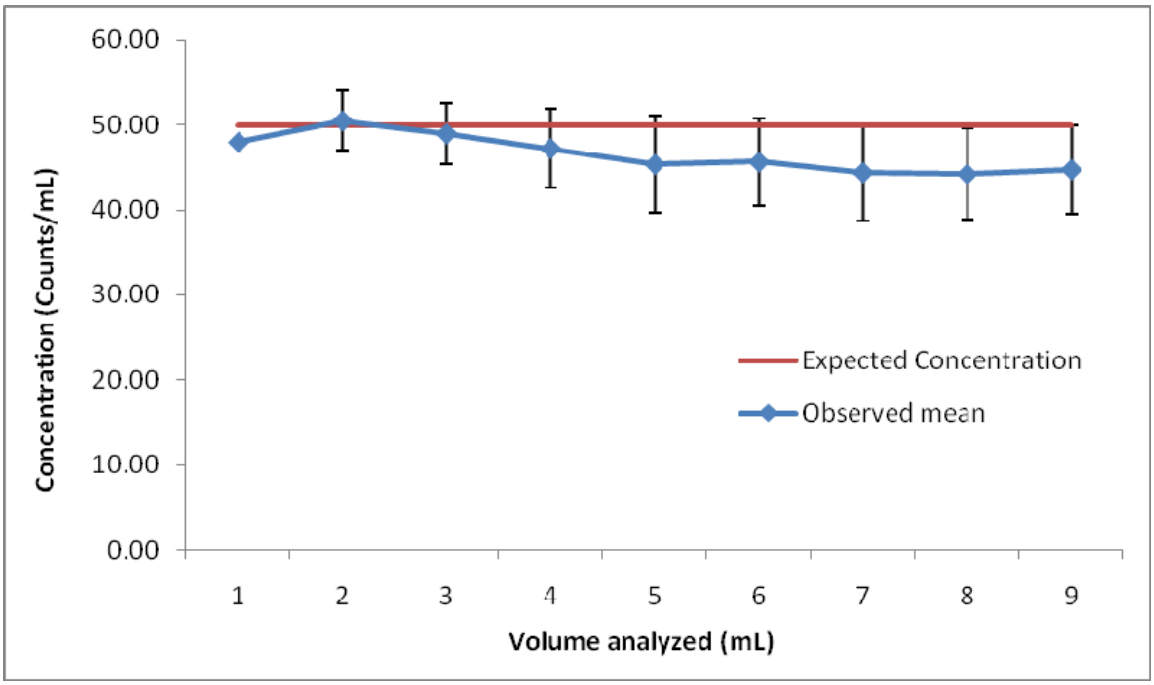

Figure 8. Phytoplankton matrix (50 beads/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 50 beads per mL concentrations.

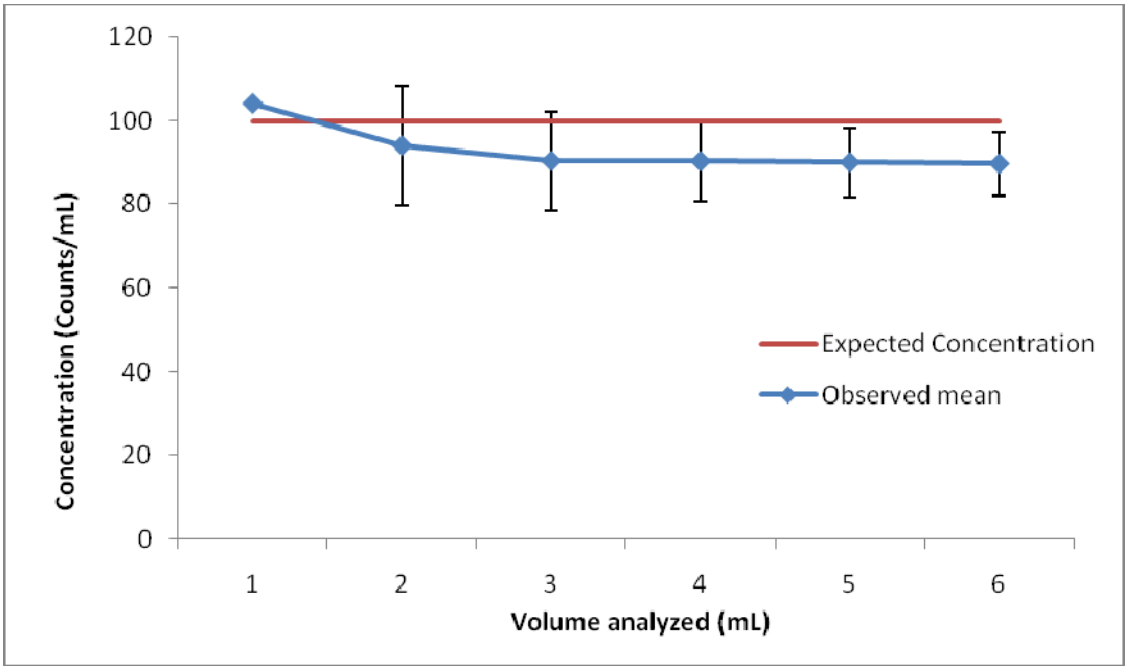

Figure 9. Phytoplankton matrix (100 beads/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 100 beads per mL concentrations.

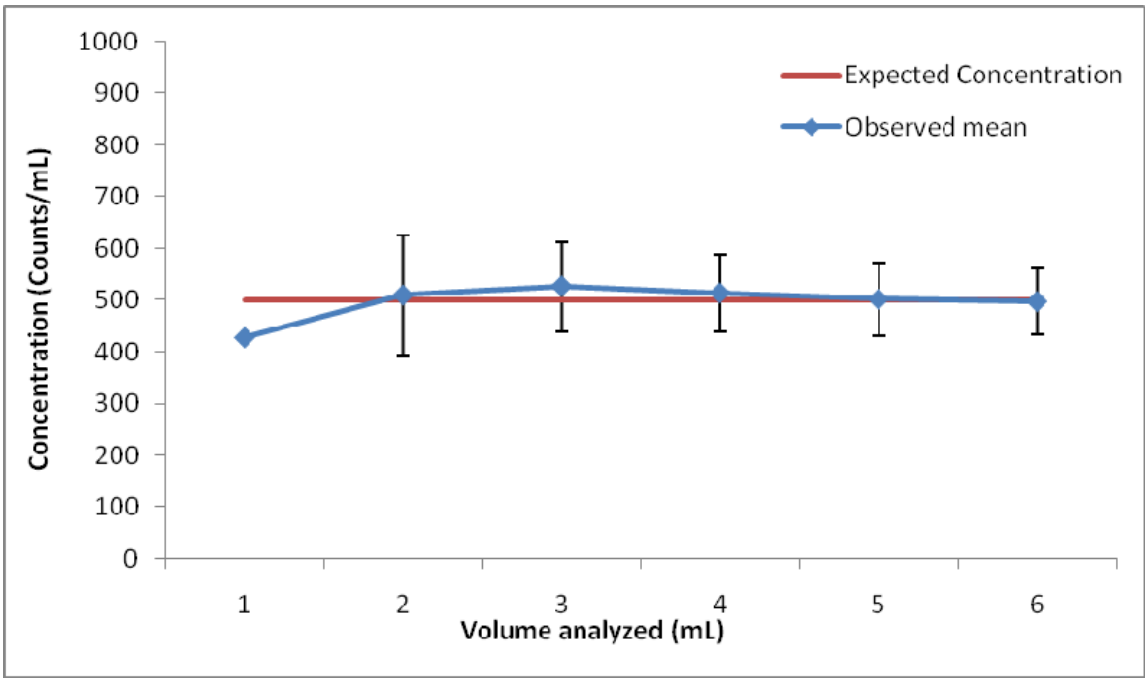

Figure 10. Phytoplankton matrix (500 beads/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 500 beads per mL concentrations.

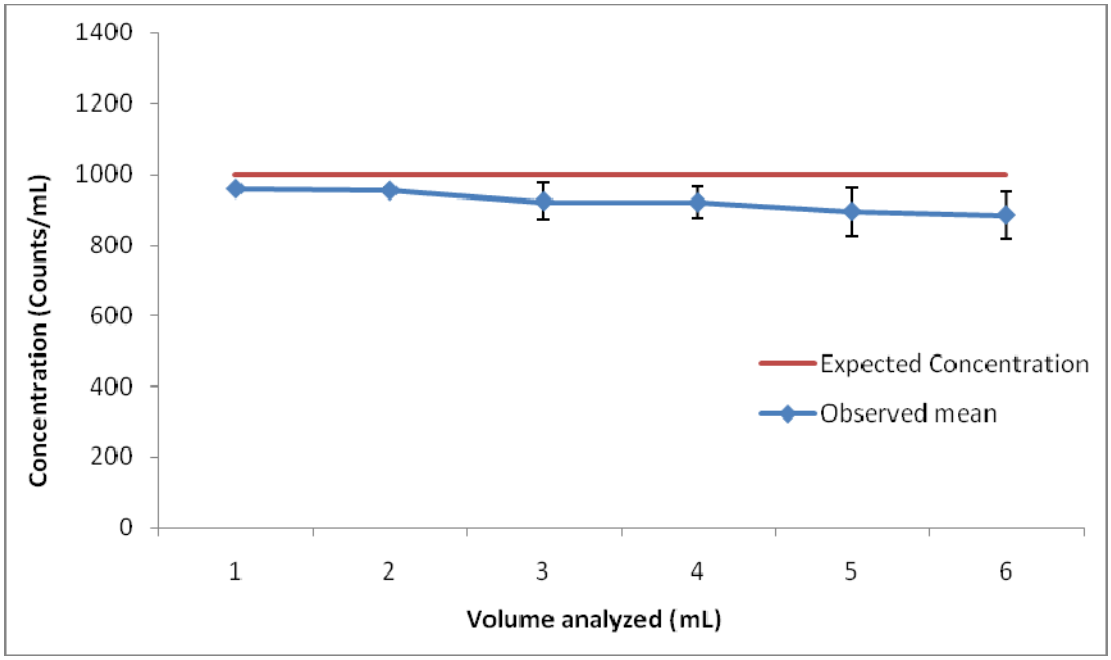

Figure 11. Phytoplankton matrix (1000 beads/mL, 20 rows counted). A comparison of means and standard deviation for measured counts of 1000 beads per mL concentrations.

4.2.2 Zooplankton Enumerations – Mean & Standard Deviation vs. Sample Size & Concentration

A total of three different test concentrations of zooplankton-sized beads were created and counted (20, 60 and 120 beads/L). The entire stock volume was analyzed for each test concentration which resulted in 500 1-mL samples (n = 500). The graphs in Figure 12 through Figure 14 show the means and corresponding standard deviations for each concentration as a function of total sub-sample volume (in this case, the subsamples have been grouped in volumes of 50 mLs). All of the test concentrations displayed relatively large standard deviations about the mean, though accuracies were very good after sufficient replicates were analyzed or organisms counted. However, at the 20 beads/L concentration, accuracy was relatively poor at nearly all volumes or sub-sample quantities analyzed, and equilibrium was not approached until the final samples. Counts performed on stock solutions of 60 beads/L, shown in Figure 13, reflected improved accuracy in terms of agreement between measured and expected, however the standard deviations were still relatively large. Finally, the data presented in Figure 14 for the 120 beads/L concentration showed both an increase in accuracy and precision compared to Figure 12 and Figure 13.

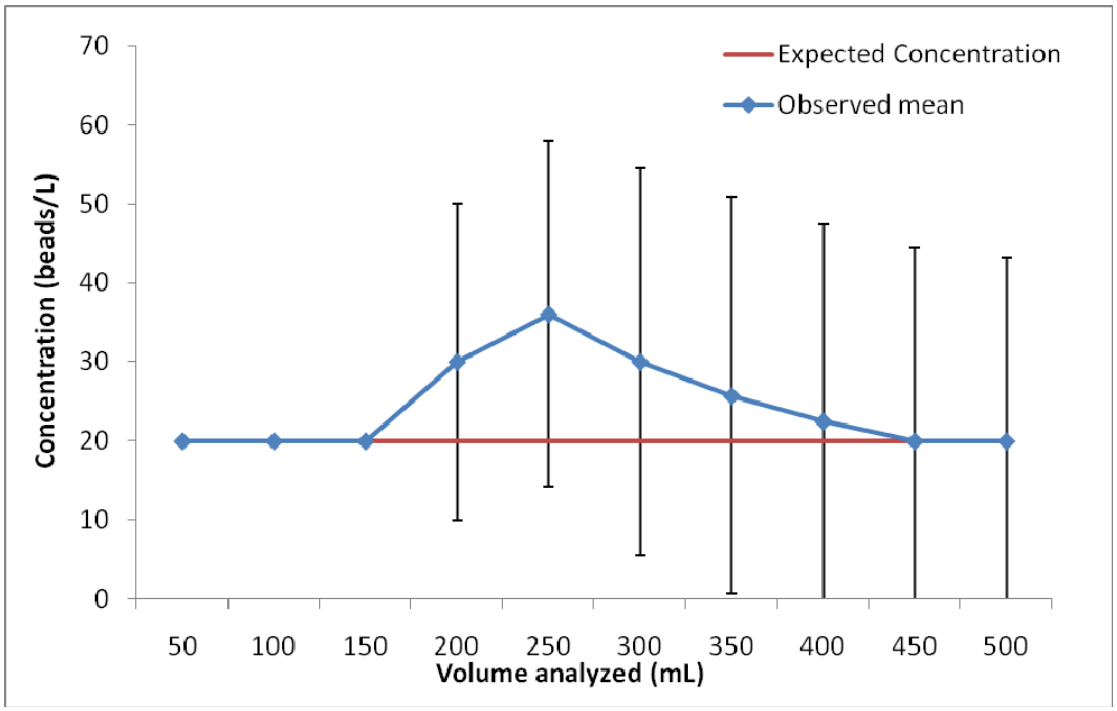

Figure 12. Zooplankton-size Bead Matrix (1 bead/50 mL). A comparison of means and standard deviations for measured counts of 1 bead per 50 mLs of sample volume measured.

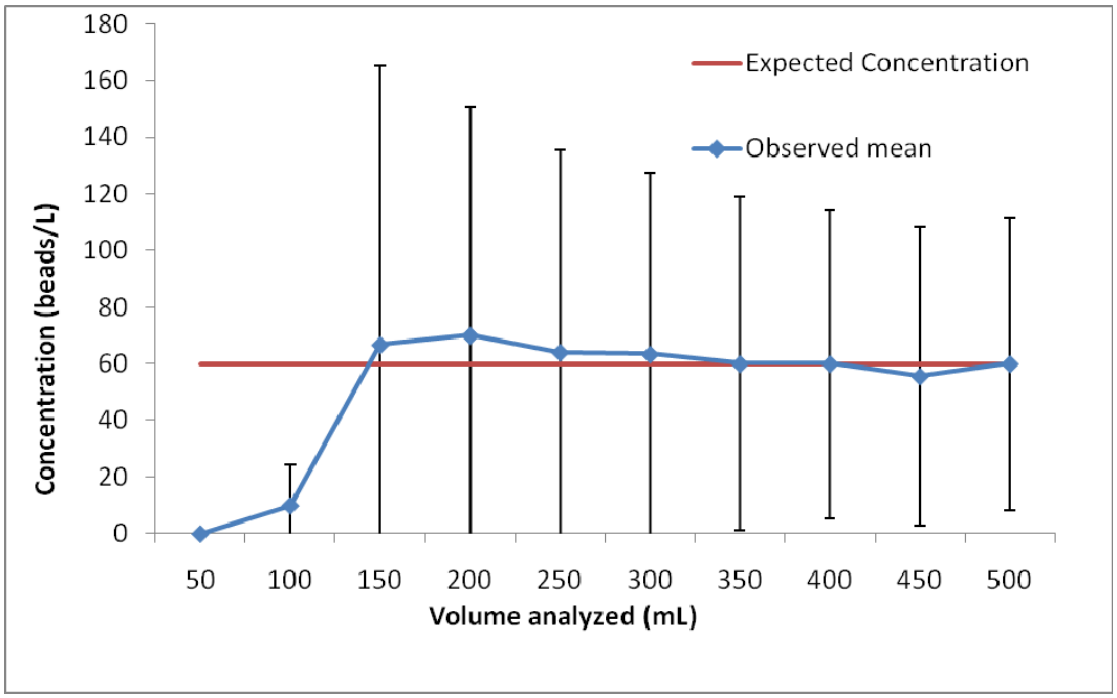

Figure 13. Zooplankton-size Bead Matrix (3 beads/50 mL). A comparison of means and standard deviations for measured counts of 3 beads per 50 mLs of sample volume measured.

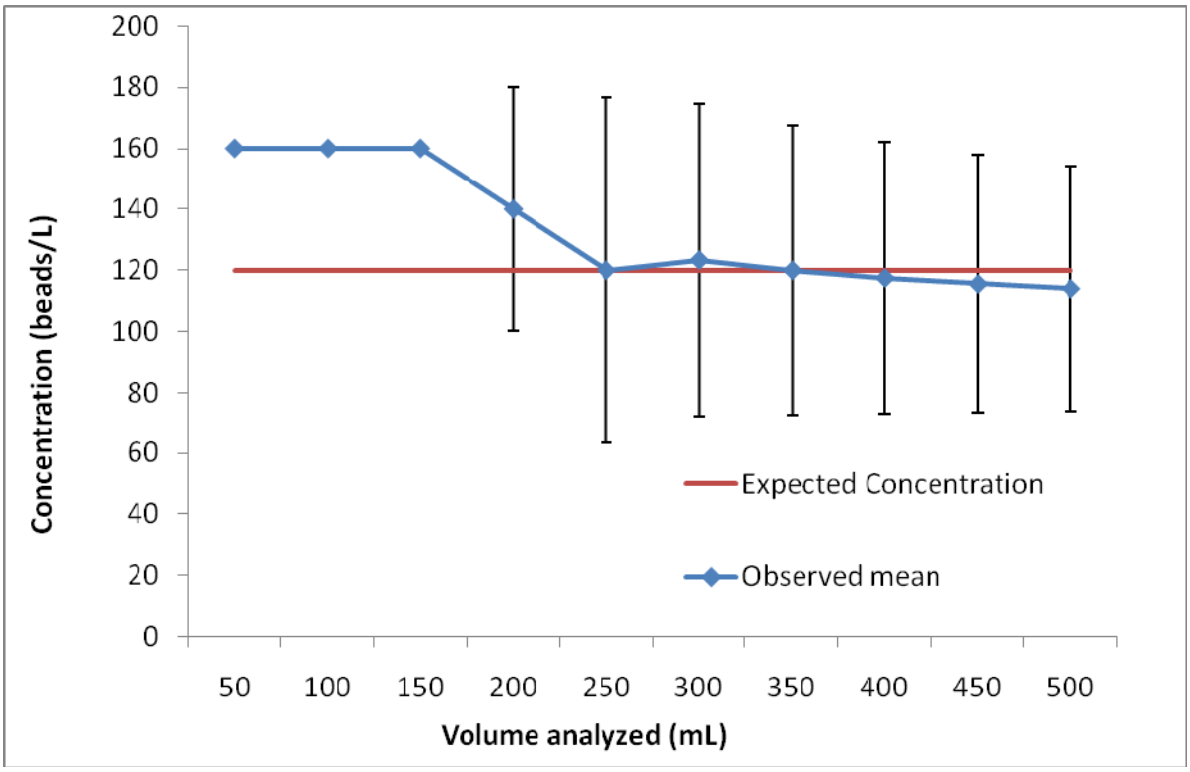

Figure 14. Zooplankton-size Bead Matrix (6 beads/50 mL). A comparison of means and standard deviations for measured counts of 6 beads per 50 mLs of sample volume measured.

4.2.3 Observed vs. Expected Counts

As an alternative way to visualize the data presented in 4.2.1, Figure 15 through 21 display the comparison of the observed bead counts to the expected bead counts for each of the bead concentrations in the simulated phytoplankton-sized test matrix. The observed data includes comparisons made for both whole chamber counts (20 rows (volume 1 mL)) and partial chamber counts (10 random rows (volume 0.5 mL)). It is evident from these data that increased bead or organism stock concentrations resulted in improved goodness of fit of the observed data relative to the expected concentration. However, with respect to sample volume processed, the goodness of fit did not definitively increase with an increase in volume sampled, though this may simply reflect systematic error in the stock concentrations or improper mixing of the sample. With the exception of the most dilute concentration, all concentrations exhibited relatively good linearity and followed the expected count line. Moreover, the observed counts were relatively unaffected by whether the whole chamber was counted or 10 random rows were used. This is a potentially significant finding which could reduce the overall time of sampling/analysis and thereby increase the number of sub-samples which may be examined.

The zooplankton data in Figure 22 through Figure 24 also showed an increase in the goodness of fit to the expected line with an increase in bead concentration. Furthermore, it also showed an increase in fit with volume sampled. This was expected given that the entire sample volume was processed and so there was a finite number of beads counted.

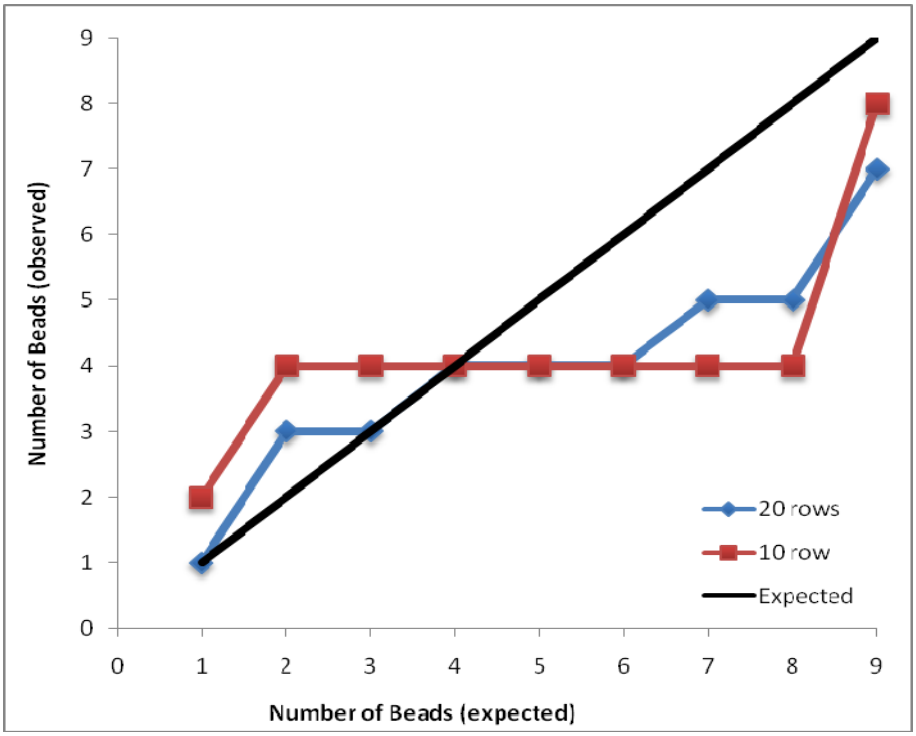

Figure 15. Phytoplankton-sized Matrix (1 bead/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

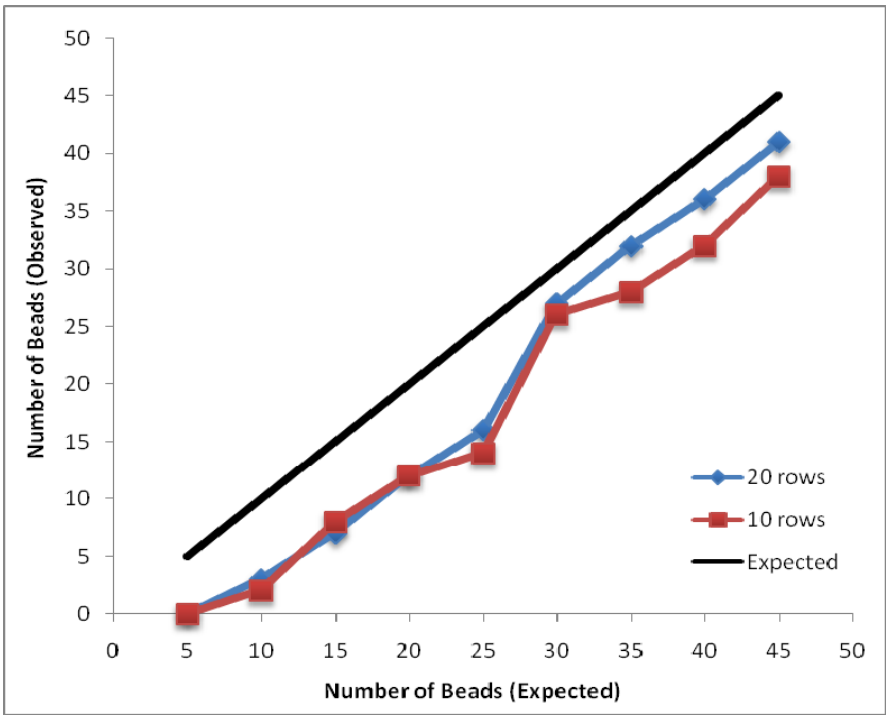

Figure 16. Phytoplankton-sized Matrix (5 beads/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

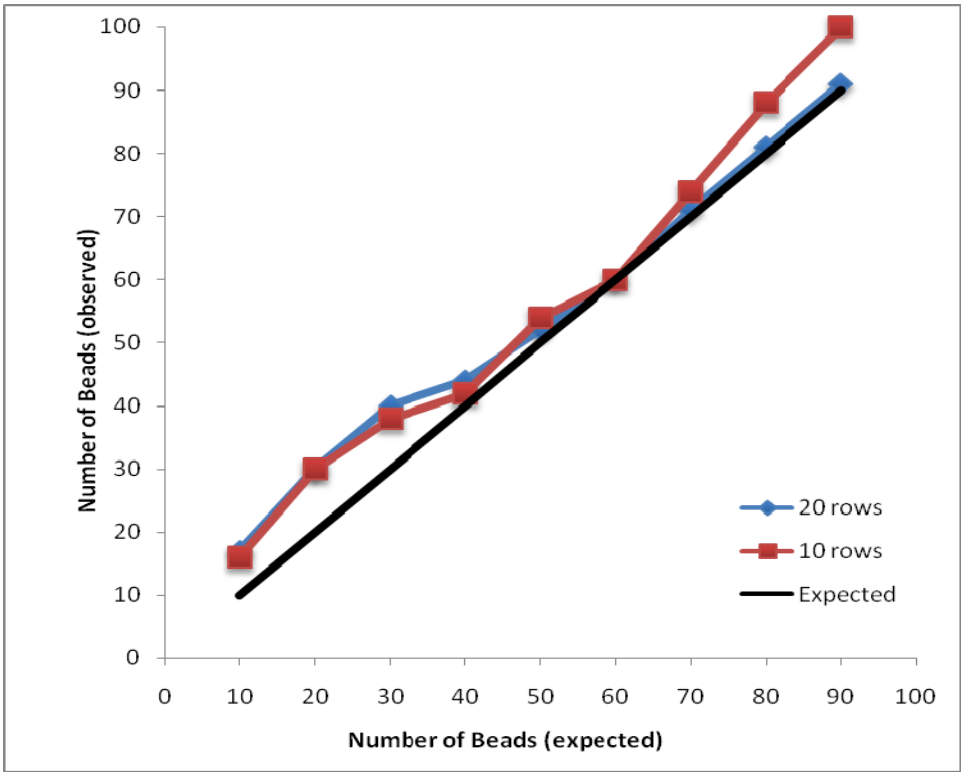

Figure 17. Phytoplankton-sized Matrix (10 beads/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

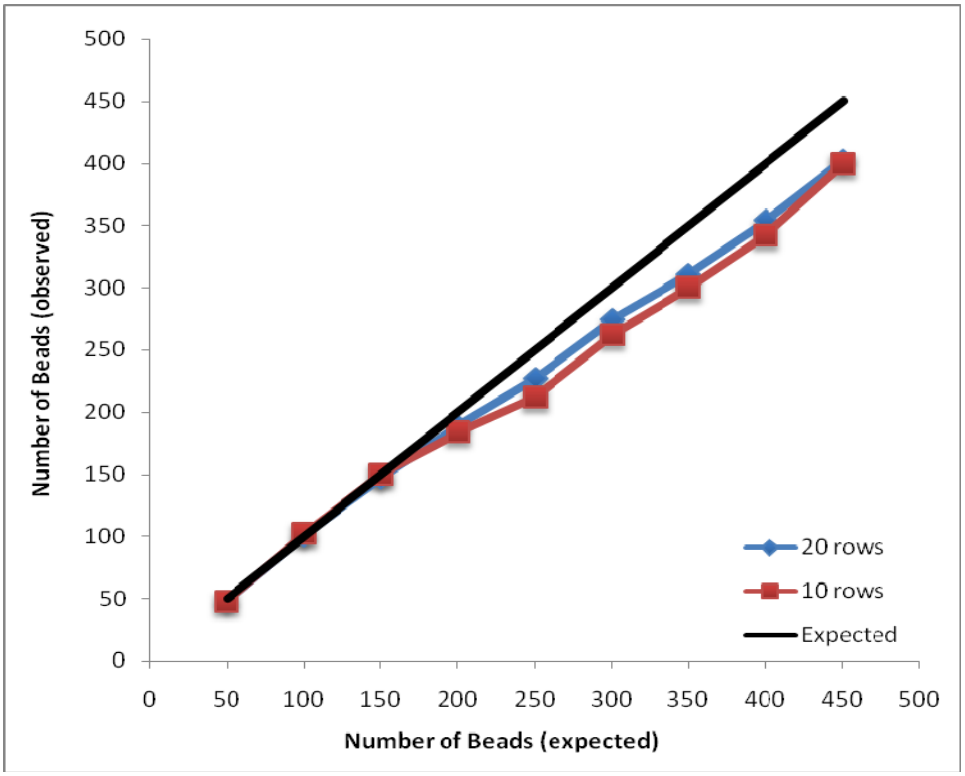

Figure 18. Phytoplankton-sized Matrix (50 beads/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

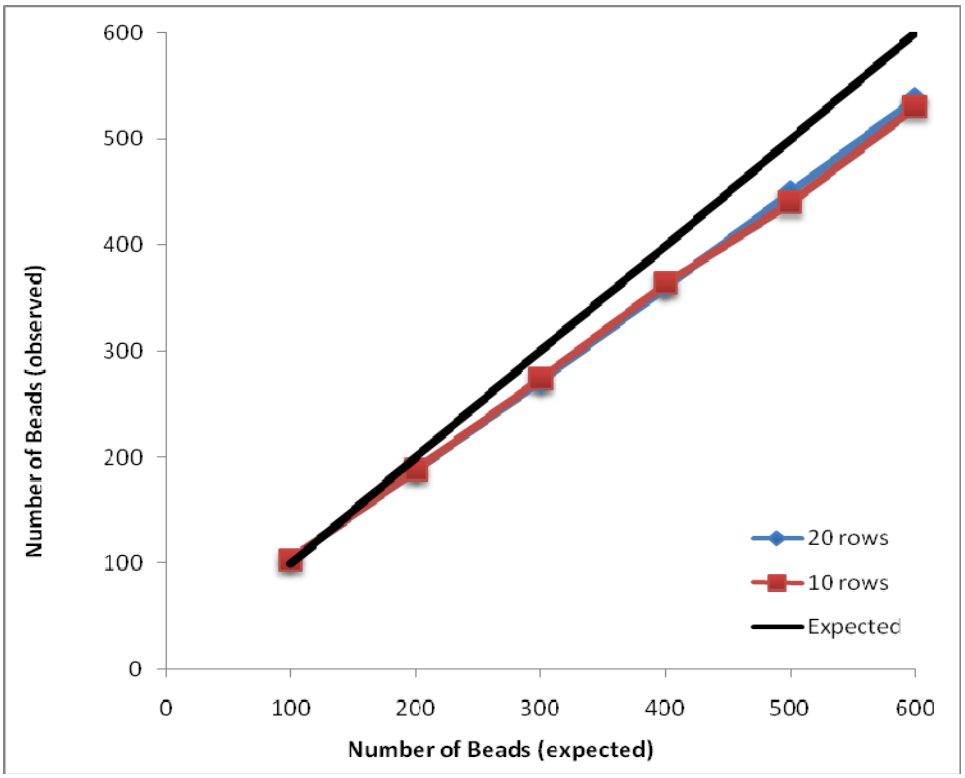

Figure 19. Phytoplankton-sized Matrix (100 beads/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

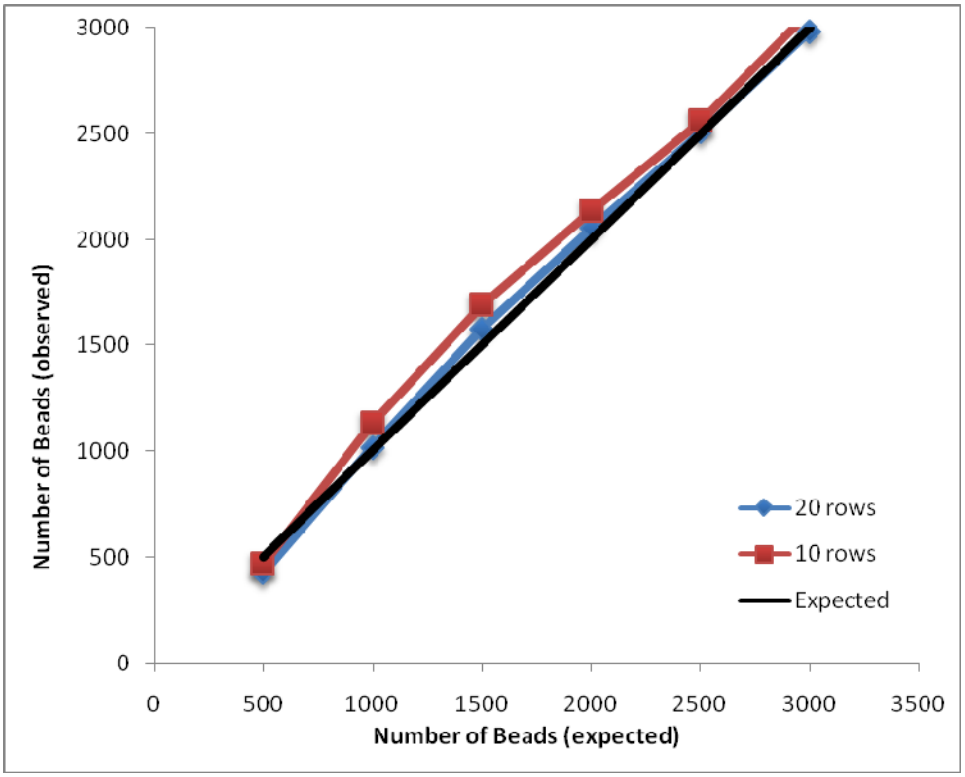

Figure 20. Phytoplankton-sized Matrix (500 beads/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

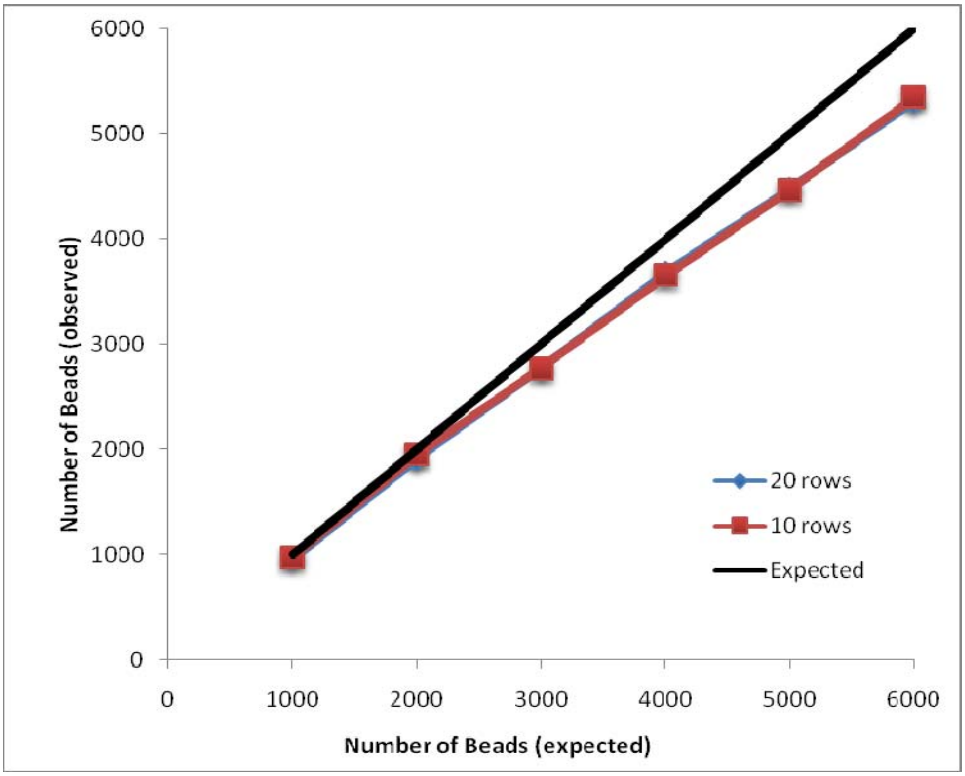

Figure 21. Phytoplankton-sized Matrix (1000 beads/mL, 10 and 20 rows counted). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

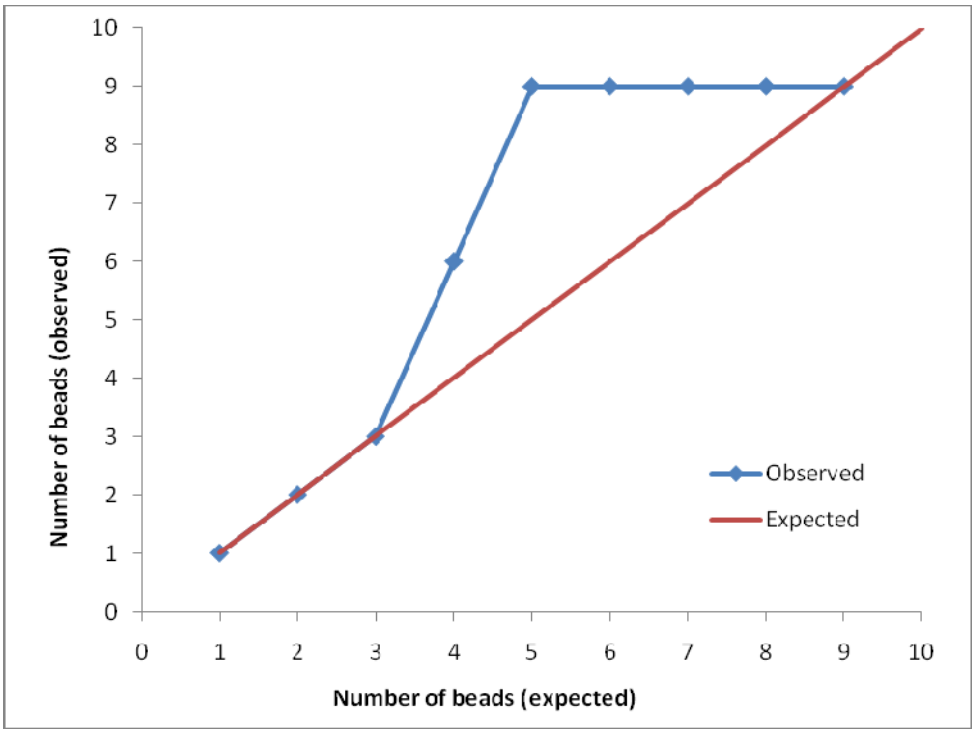

Figure 22. Zooplankton-Size Matrix: (10 bead per 50 mL). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

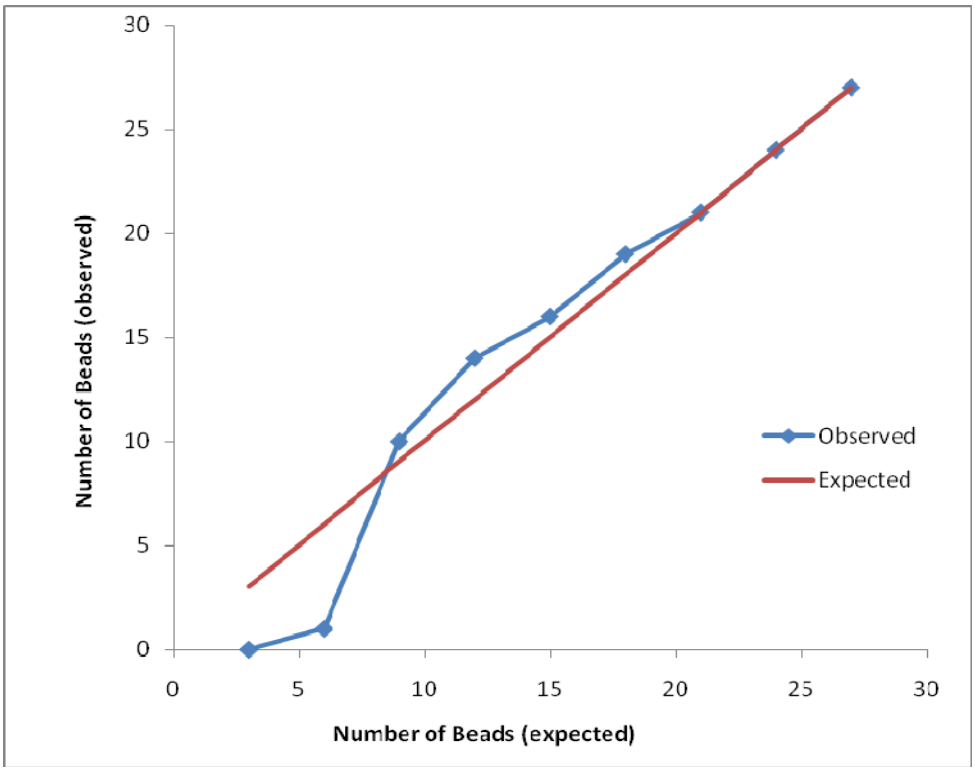

Figure 23. Zooplankton-Size Matrix: (30 bead per 50 mL). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

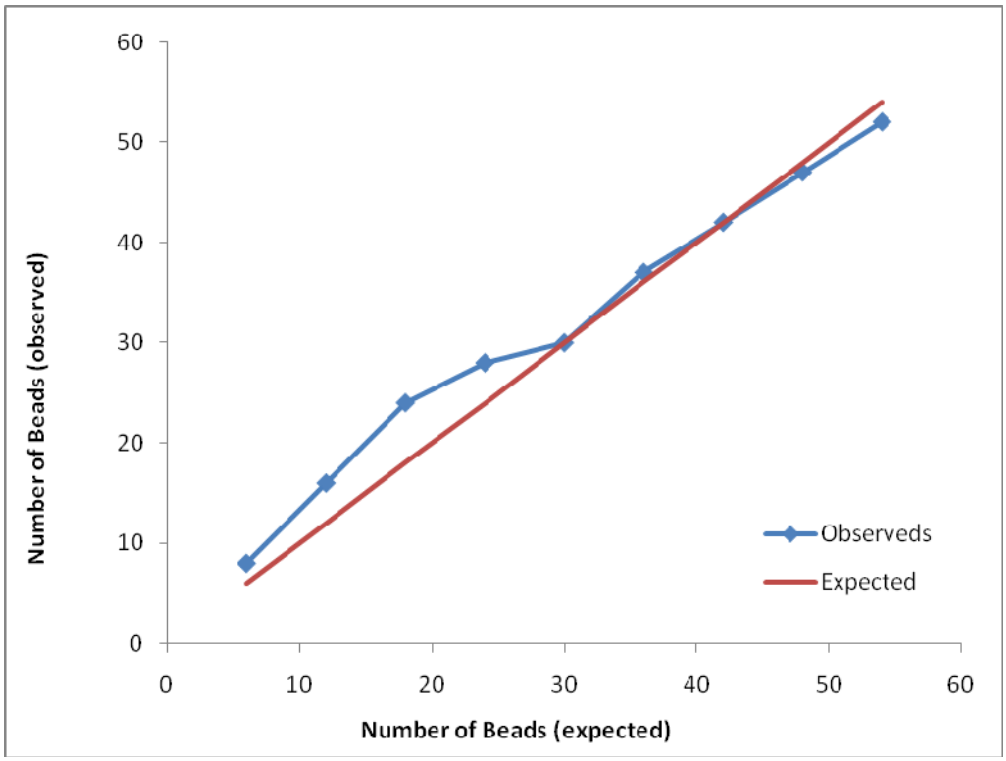

Figure 24. Zooplankton-Size Matrix: (60 bead per 50 mL). Comparison of Observed to Expected Number of Beads in Relation to Volume (mLs) Sampled.

4.2.4 Accuracy Evaluated as Percent Difference

Accuracy was evaluated using the percent difference, as explained previously, for enumerations with whole chamber counts. Table 4 presents the percent difference between the expected or known stock concentration and the observed concentration as a function of stock concentration and number or volume of sub-samples analyzed. Since the concentrations greater than 50 beads/mL did not require more than six replicates, subsequent data are listed as not applicable (na). At the most dilute concentration, the percent difference is highly variable and does not stabilize even after nine replicates. Some improvement is noted at 5 beads/mL, but only after a minimum of six replicates (6 mL). In the case of 50, 100 and 1000 beads/mL, percent difference increases with increasing replicates which is not anticipated as accuracy would be expected to improve with increased replication. However, this result is plausible since the percent difference among these at high sample quantities was a maximum of 11.5% which is not significantly different than the vendor's reported +/- 10% variation in stock concentrations. Thus, it is possible that the expected concentration is inaccurate. In general however, further improvements in the accuracy of the observed mean were not achieved for replication beyond 5 replicates for concentrations ≥ 50 beads/mL. A stable accuracy within the systematic error bounds of the stock concentration was not achieved with 1 or 5 beads/mL samples.

Table 4. Comparison of accuracy in terms of absolute percent difference of the observed from the expected concentration for simulated phytoplankton concentrations. (na = not applicable).

Calculating % Difference	*Volume Sampled (mL)*								
# Beads per mL	***1***	***2***	***3***	***4***	***5***	***6***	***7***	***8***	***9***
1	0.0	50.0	0.0	0.0	20.0	33.3	28.6	37.5	22.2
5	100.0	70.0	53.3	40.0	36.0	10.0	8.6	10.0	8.9
10	70.0	50.0	33.3	10.0	4.0	0.0	1.4	1.3	1.1
50	4.0	1.0	2.0	5.5	9.2	8.7	11.1	11.5	10.4
100	4.0	6.0	9.7	9.8	10.0	10.3	na	na	na
500	14.4	1.9	5.3	2.7	0.3	0.6	na	na	na
1000	4.1	4.6	7.7	8.0	10.6	11.6	na	na	na

4.2.5 Precision Evaluated as Coefficient of Variation

The coefficient of variation for all seven simulated phytoplankton concentrations are shown in Figure 25. The values shown are counts performed on the whole Sedgewick Rafter counting chamber (vice 10 random rows). At the most dilute concentrations of 1 and 5 beads/mL the CV, reported as a decimal number, was well over the 0.20 benchmark. Specifically the CV's were above 1.00 indicating 100% and showed that these measurements were imprecise and that no useful conclusions could be expected at these sample concentrations with similar enumeration methods.

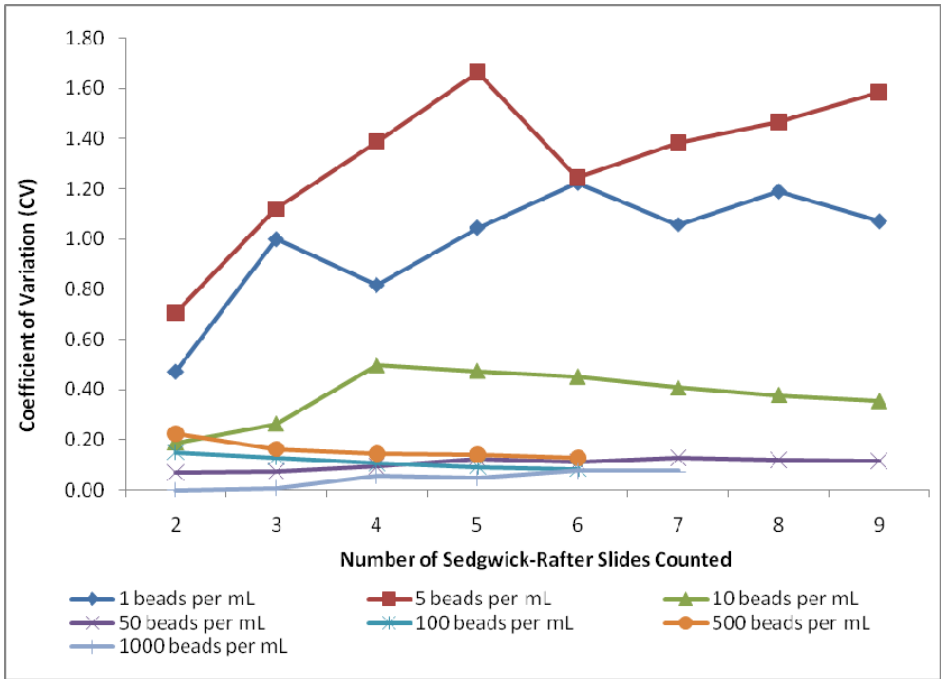

Figure 25. Comparison of the coefficient of variation (CV) at a range of discharge concentrations in relation to volume sampled (Number of Sedgwick-Rafter slides counted).

The CV's for the remaining four concentrations are much lower and are displayed separately in Figure 26 for a clearer comparison. These data indicate that a substantial improvement in precision, as defined by CV, was realized when discharge samples were concentrated sufficiently to provide sample concentrations greater than or equal to 50 organisms/mL. Moreover, if a CV of 0.2 (the standard deviation is 20% of the mean) is used as an arbitrary benchmark for acceptable precision, then it would be recommended that a target sample concentration of at least 50 organisms/mL should be utilized. These data also show that precision is not appreciably improved, at these sample concentrations, after analyzing approximately five full Sedgwick-Rafter counting chambers.

These results have several implications. First, treated ballast water discharges which are expected to meet the IMO D-2 standard for organisms in the size class of greater than 10 µm but less than or equal to 50 µm, should be concentrated at least 5 times to provide a precision of not less than 20%. In terms of sample replicates, when using the Sedgwick Rafter counting chamber method, at least four counting chambers should be analyzed to insure both adequate precision and accuracy (<10%). Additional replicates at this concentration range should not be expected to significantly increase precision or accuracy. Finally, future efforts should examine the practicality of concentrating discharge samples of living organisms to insure that the sample remains representative (no loss of organisms or effect on mortality) and to insure that adequate time is available to both conduct the necessary concentration step and the enumeration. Note that the latter may be further frustrated if larger concentration factors are required for improved accuracy or precision.

As a caution, these conclusions and analyses must be considered with the understanding that these data are the ideal scenario. In other words, all additional sources of error or variability have been eliminated by virtue of the experimental method; specifically, no sample degradation is anticipated; there is no need to assess viability or health; enumerated objects do not clump and are homogeneous in size and shape. With respect to actual sampling conditions, this is wholly unrealistic. As such, the results of this work should be considered conservatively.

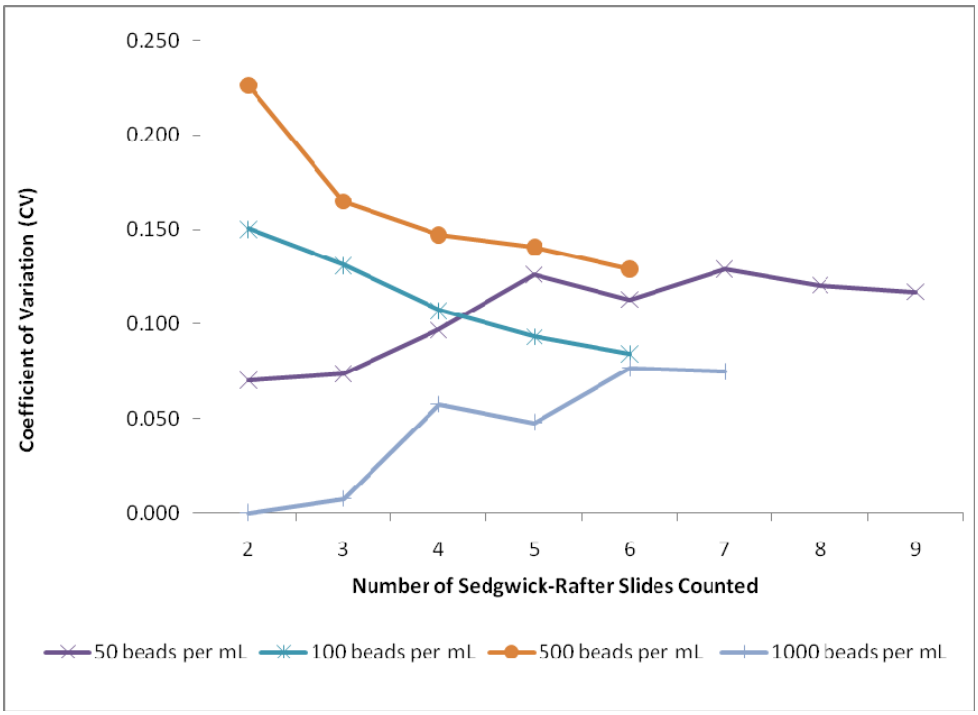

Figure 26. A closer comparison of the coefficient of variation (CV) for the highest simulated phytoplankton sample concentrations analyzed.

4.2.6 Statistical Tests for Sample Concentration and Volume Differences

4.2.6.1 Phytoplankton

The multivariate Model I, fixed factor, ANOVA was used with bead concentration as the between-subjects factor. The four concentrations tested were 50, 100, 500 and 1000. Table 5 shows that the Levene test of equal variances was not significantly different (α=0.05, p-value = 0.931) indicating that the variances between the four bead concentrations was not significantly different. As previously stated, the purpose of the Levene test is to determine whether the variance of the data used is essentially the same thereby supporting the use of parametric test such as ANOVA to perform statistical comparative analyses.

With respect to the ANOVA test comparing these four bead concentrations, Table 6 displays the statistical output which shows a p-value of 0.001. Thus the null hypothesis is rejected and the conclusion is that there is a significant difference in the CV between the 50, 100, 500 and 1000 bead concentrations. Furthermore, the eta-squared value of 0.649 indicated that CV is affected by bead concentration. Table 7 shows the multiple comparisons made using the Tukey Test which further clarifies where any potential significant differences might have occurred. It can be seen that the significant differences (marked with *) are between the 500 when compared to the 50 and 1000 bead concentrations at a p-value of 0.019 and 0.001.

Table 5. Phytoplankton matrix (20 rows counted): Levene test of equal variances. 95 percent confidence (α=0.05).

Levene test of Error Variances	F	p-value (significance)
Bead Concentration	0.145	0.931

Table 6. ANOVA test results comparing four bead concentrations (50, 100, 500, 1000) between sample volumes of 2 ml to 6 mls analyzed. Alpha set to 95% confidence (α =0.05). a p-value < 0.05 indicates a significant difference.

Between Subjects-Dependent Variable	F	p-value (significance)	ξ^2	n
Bead Concentration	9.865	0.001	0.649	5

Table 7. Phytoplankton matrix (20 rows counted): ANOVA tukey post hoc test of multiple comparisons. Comparisons were performed on three test bead concentrations (50, 100, 500, 1000) between sample volumes of 2 ml to 6 mls analyzed. alpha set to 95% confidence (α =0.05). a p-value less than 0.05 indicates a significant difference (indicated by *).

(A) Bead Concentration	(B) Bead Concentration	Mean Difference (A-B)	p-value (significance)
50	100	-1.60E-02	0.849
	500	-6.60E-02	0.019*
	1000	4.00E-02	0.220
100	50	1.60E-02	0.849
	500	-5.00E-02	0.093
	1000	5.60E-02	0.052
500	50	6.60E-02	0.019*
	100	5.00E-02	0.093
	1000	0.1060	0.001*
1000	50	-4.00E-02	0.220
	100	-5.60E-02	0.052
	500	-0.1060	0.001*

4.2.6.2 Zooplankton

Figure 27 shows coefficient of variation for the zooplankton matrix of all three bead concentrations. The 20 beads/mL concentration had a CV of 0 - 1.5 which appeared to be asymptotically increasing with volume analyzed. The 60 beads/mL had an initially high CV but decreased to a stable value of 2.5 after 0.45 liters (n = 450) were analyzed. The 120 beads/mL analytical solution presented a CV of 0.8 at the conclusion of all samples. As these sample concentrations represented the concentration of 1000, 3000 and 6000 for a total whole water sample volume of 6 m^3 at an assumed discharge concentration equivalent to the IMO D-2 standard (less than 10 organisms/m^3), they represented the most concentrated samples planned as shown in Table 2. On the basis that more dilute concentrations would only result in reduced precision, further analyses on lower concentrations were not performed. Furthermore, looking for significant differences between bead concentrations with high CVs would not provide useful information given that it would be hard to confidently say anything about the counted populations due to this high variability.

The relatively poor precision identified in these analyses result in several significant findings or recommendations. First, current test protocols designed to sample discharge concentrations at or below 10 organisms/m^3 with triplicate 1 m^3 samples or duplicate 2 m^3 samples as proposed in the most recent version of the ETV protocol will result in inadequate precision even when concentrated by 6000 times with this enumeration method at n = 450. On the basis of these data, concentration of 6 m^3 by 6000 times and 450 1-mL aliquots analyzed would provide improved precision but would still remain below the 20% (as CV) identified as the benchmark for phytoplankton. Because higher concentration factors are likely unrealistic, larger sample sizes and improved analytical methods should be considered.

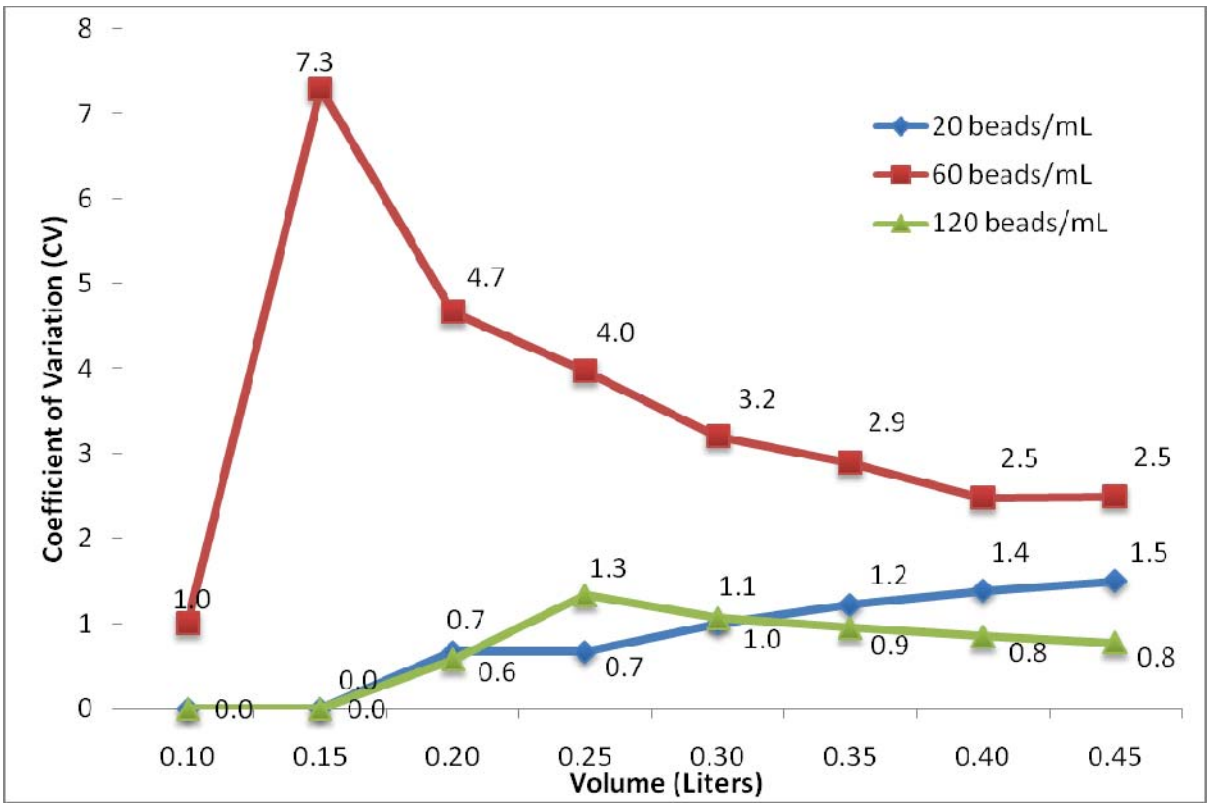

Figure 27. Zooplankton matrix: comparison of the coefficient of variation (CV) to volume sampled for three concentrations (20, 60 and 120 beads/mL).

4.3 Distribution Analyses for Simulated Phytoplankton

If counts follow a Poisson distribution, the distribution variance equals the distribution mean. Figure 28 is a plot of the sample variance versus the sample average for the phytoplankton size microbead data. The line y = x has been added to the plot. For clarity, the axes have been transformed to the logarithmic scale. At low concentrations, the plotted points are fairly close to the line, but the points corresponding to the two highest nominal concentrations (500, 1000 beads/mL) are far above the line.

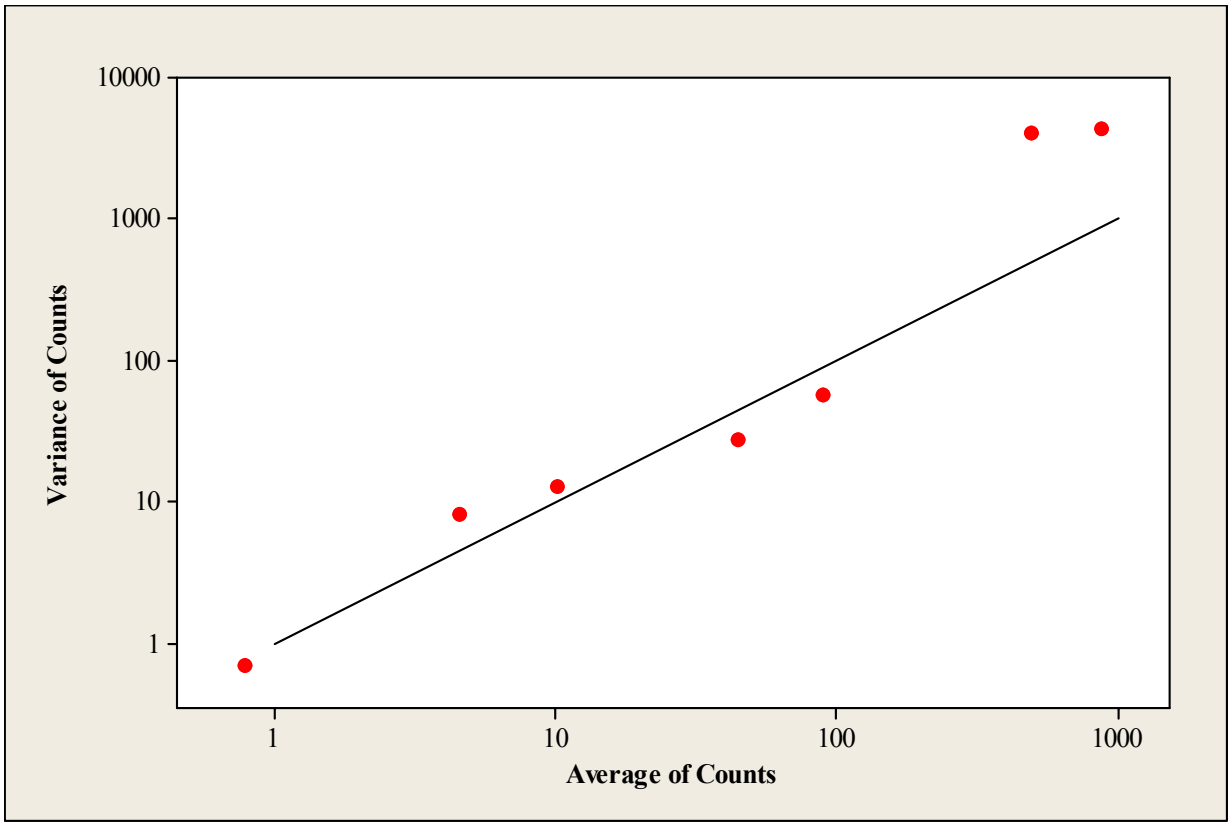

Figure 28. Variance versus mean for simulated phytoplankton enumerations.

Given the Poisson model, the statistic $\Sigma (X_i - X^*)^2/X^*$ (where X_i = value of the ith sample and X^* = sample mean) should have an approximate chi-square distribution with (n-1) degrees of freedom. Large values indicate that the variance is larger than the mean and lead to the rejection of the Poisson hypothesis. To be sure how this statistic would behave for small n, a simulated distribution was used though concentrations larger than 700 were not possible since the software in use does not allow this for the Poisson random variables. The 95th percentiles of the simulated distributions are compared to the corresponding chi-square percentiles in Table 8. While not exact, the simulated percentiles are close enough to the chi-square percentiles that it is reasonable to use the chi-square approximation to the distribution of the test statistic.

Table 8. Simulated Poisson distribution and corresponding chi-square percentiles.

Nominal Concentration	Volume Analyzed or Number of 1mL Slides	Simulated Percentile	Chi-Square Percentile
1	9	15.0	15.5
5	9	15.4	15.5
10	9	15.5	15.5
50	9	15.3	15.5
100	6	11.0	11.1
500	6	11.2	11.1

The test statistic values and chi-square p-values are shown in Table 9. The p-values are greater than .05 for nominal concentrations up through 100, so there is insufficient evidence to reject the Poisson hypothesis. The Poisson model does not fit well for the two highest concentrations.

Table 9. Comparison of the test statistic to the chi-square percentile.

Nominal Concentration	Volume Analyzed or Number of 1mL Slides	Test Statistic	p value
1	9	7.1	0.522
5	9	14.5	0.069
10	9	10.2	0.253
50	9	4.9	0.768
100	6	3.2	0.672
500	6	41.3	<0.001
1000	6	24.9	<0.001

Figure 29 is a scatterplot of Observer 2 versus Observer 1 counts for the two highest nominal concentrations with the line y = x overlaid on the plot. There is consistent and good agreement between the two observers, despite the fact that Observer 2 was not always able to count the complete slide. The departures from the Poisson model at high concentrations do not appear to result from a counting problem, so they must arise from the sub-sampling. There is no way to be certain, but more thorough mixing of the prepared volume may be required at high concentrations.

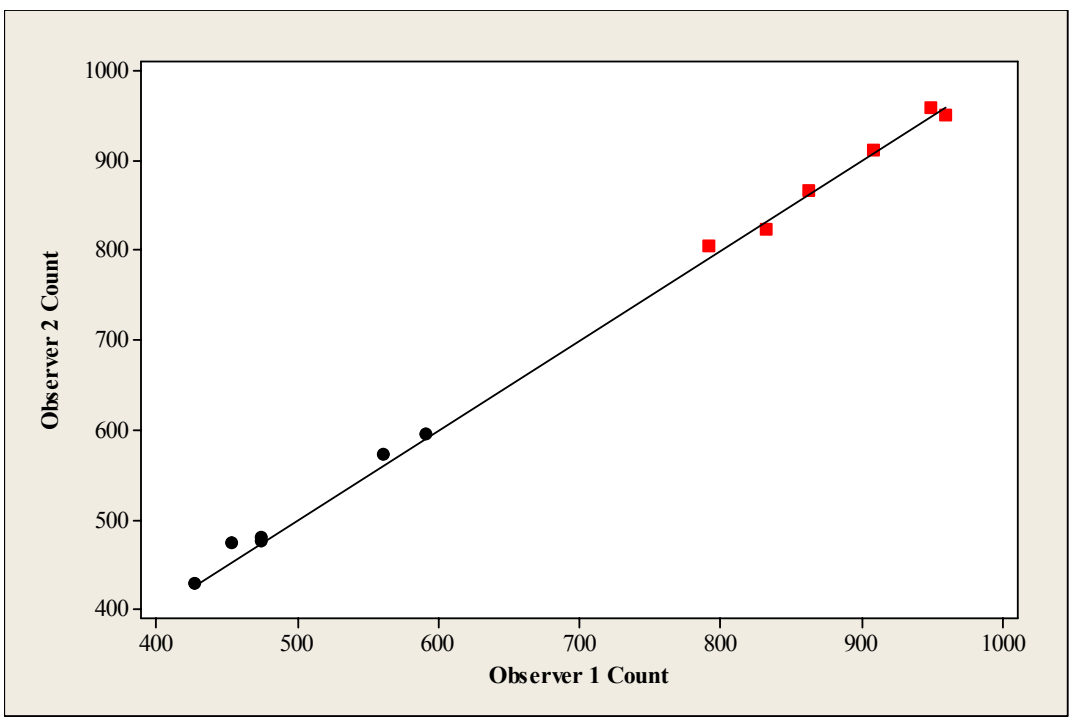

Figure 29. Observer comparison for high microbead concentrations.

4.3.1 Whole versus Partial Phytoplankton Counting Chamber Enumeration

The data presented thus far represents all 20 rows per counting slide. Calculations were then performed to determine how accuracy, precision and CV would be affected if 10 of the 20 rows were chosen randomly to simulate counting half of the slide being measured. Again the one and five bead concentrations had the highest percent difference as compared to the remaining five concentrations as can be seen in Table 10. The trends seen with the 20 rows counted were also evident in the 10 random row counts in that there was a decrease in absolute Percent Difference with an increase in bead concentration.

Table 10. Comparison of accuracy in terms of percent difference for simulated phytoplankton stocks with only 10 random rows counted in the counting chamber.

Calculating % Difference	*Volume Sampled (mLs)*								
# Beads per mL	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1	100.0	100	33.3	0.0	20.0	33.3	42.9	50.0	11.1
5	100.0	80.0	46.7	40.0	44.0	13.3	20.0	20.0	15.6
10	60.0	50.0	26.7	5.0	8.0	0.0	5.7	10.0	11.1
50	4.0	2.0	0.0	8.0	15.2	12.7	14.3	14.5	11.1
100	4.0	6.0	8.7	9.0	12.0	11.7	na	na	na
500	6.4	13.2	12.9	6.7	2.4	2.6	na	na	na
1000	2.8	2.8	7.7	8.8	10.7	10.9	na	na	na

The coefficients of variation for all seven bead concentrations are shown graphically in Figure 30. When multiplied by 100 to measure in percentages, the 1 and 5 bead concentrations displayed %CV's that ranged from 0.0% to 185.2% for 1 bead per mL and 81.4% to 141.4% for 5 beads per mL. The 10 bead concentration was between 0.0% and 52.5%. The %CV values for the remaining four concentrations were much lower and are displayed in Figure 31 for a clearer comparison. The 50 bead per mL concentration had a minimum %CV of 6.9% and a maximum of 25.7%. The 100 bead concentration had a minimum %CV of 9.9% and a maximum of 15.0%. The 500 bead concentration had a minimum %CV of 17.0% and a maximum of 24.0%. Finally, the 1000 bead concentration had a minimum %CV of 0.0% and a maximum of 9.3%. This data suggests that half of a Sedgwick-Rafter slide can be counted as long as the rows are chosen randomly and the precision will still be below 25% which is acceptable in many biological applications. Ultimately being able to analyze half of the slide would allow more time for processing and result in analyzing more slides in the allowable eight hour time frame.

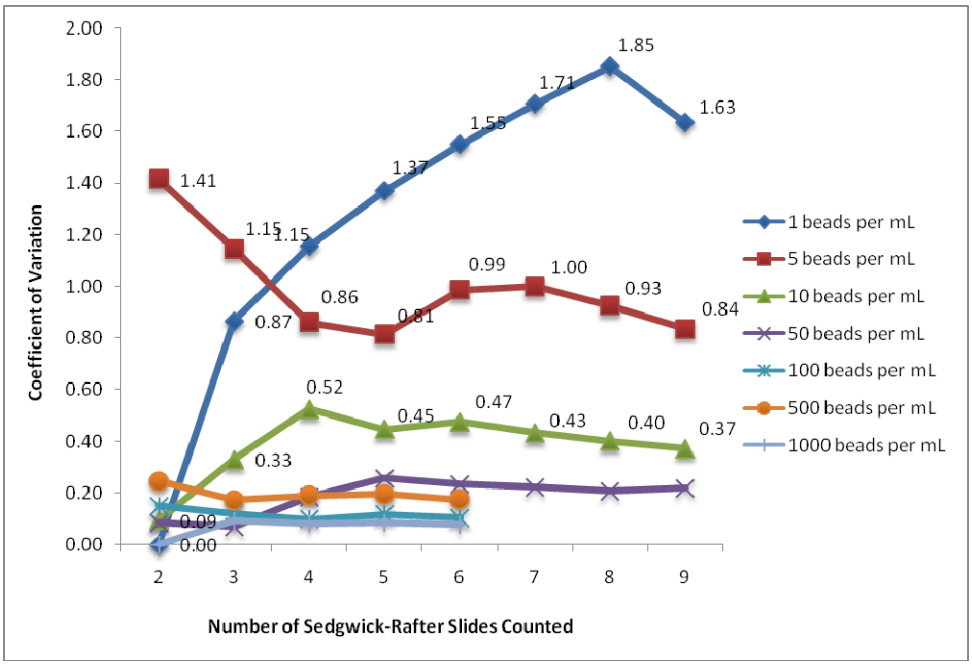

Figure 30. Comparison of the coefficient of variation (CV) relative to volume sampled. Comparisons made for seven bead concentrations (1, 5, 10, 50, 100, 500, 1000 beads/mL) in relation to volume sampled (number of Sedgwick-Rafter slides counted) when only partial chambers (i.e., 10 rows) were enumerated.

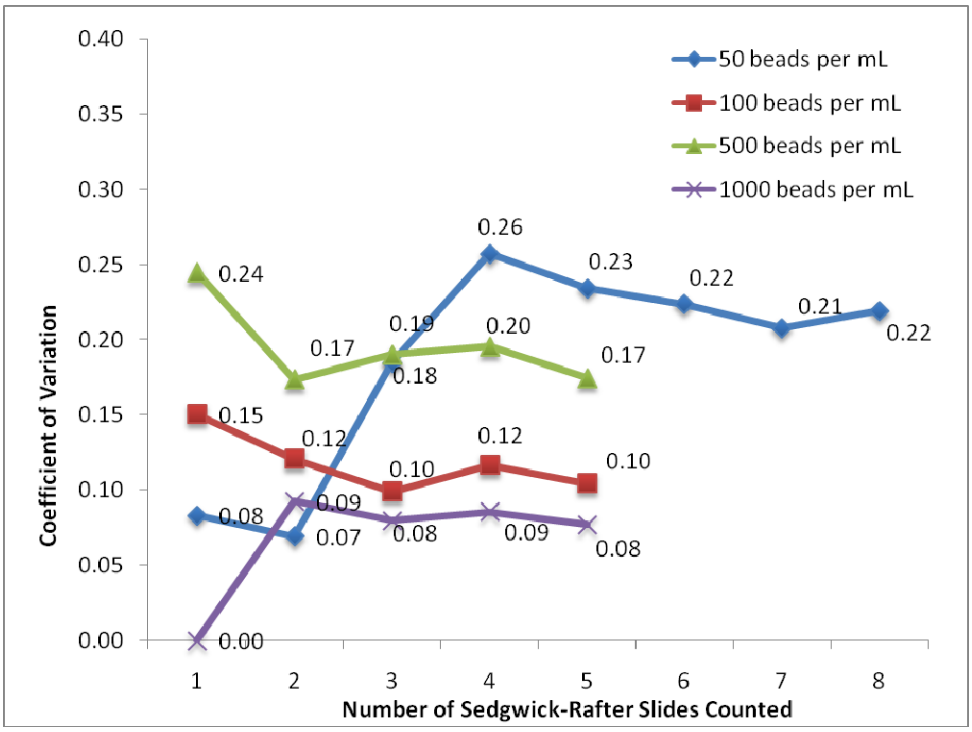

Figure 31. Comparison of the coefficient of variation (CV) relative to volume sampled for concentrated samples. Comparisons made for the most concentrated solutions (50, 100, 500, 1000 beads/mL) in relation to volume sampled (number of Sedgwick-Rafter slides counted) when only partial chambers (i.e., 10 rows) were enumerated.

The multivariate Model I, fixed factor, ANOVA was used with a between-subjects factor of bead concentration. The four concentrations tested were 50, 100, 500 and 1000. Table 11 shows that there was a significant difference in the variance between the bead concentrations (α=0.05, p-value = 0.005). This indicates that either transformation of the data or use of a nonparametric test may be called for in statistical comparative analyses of similar data sets. These additional statistical analytic considerations could partially negate the benefit of counting half of the slides. However, the t-test is relatively robust to departures from the assumption of equal variances. Further, as shown in

Table 12, there were no significant differences between the CV's of the 20 and 10 rows counted for each bead concentration. This would indicate that half of the Sedgwick-Rafter slide could be counted as long as the rows were chosen randomly. This is worthwhile given the potential decrease in analysis time and increase in volume sampled that would be produced by analyzing only half of a Sedgwick-Rafter slide.

Table 11. Phytoplankton matrix (10 rows counted): Levene test of equal variances with 95 Percent confidence (α=0.05).

Levene test of Error Variances	F	p-value (significance)
Bead Concentration	6.324	0.005

Table 12. Comparison of CV's between the 20 row and 10 row counts for the 50, 100, 500, and 1000 bead concentrations. For dependent samples T-Test, if absolute value of observed t (shown in table) is greater than the t-critical value of 2.78 then the difference between the CV's is statistically significant. T-critical value based on two-tailed test, degrees of freedom equals 4 and alpha set to 95% confidence (α =0.05).

Bead Concentration	Variance Assumption	p-value (Sig. (2-tailed)	t	n
50	Equal	0.130	-1.688	5
	Not Equal	0.156	-1.688	
100	Equal	0.694	-0.407	5
	Not Equal	0.696	-0.407	
500	Equal	0.163	-1.537	5
	Not Equal	0.166	-1.537	
1000	Equal	0.623	-0.511	5
	Not Equal	0.624	-0.511	

4.3.2 Differences Between Observers 1 & 2 for Sub-samples

Table 13 displays the results of the comparison of counts between two different observers for all 20 rows for all phytoplankton concentrations. The results of the independent T-Test showed that across all bead concentrations counted and regardless of equal variance, the p-value was greater than 0.05 which meant that the null hypothesis was accepted and that the bead counts did not differ among observers.

Table 14 displays the results of the comparison of counts between two different observers for the entire volume counted for each of the zooplankton-size bead concentrations. The results of the single factor

analysis of variance showed that across all bead concentrations, the p-value was greater than 0.05 which meant that the null hypothesis was accepted and that the bead counts did not differ among observers.

Table 13. Phytoplankton matrix (20 rows counted): Independent T-test test results comparing observer counts for seven bead concentrations. Alpha set to 95% confidence (α =0.05).

Bead Concentration	Variance Assumption	p-value (Sig. (2-tailed))	t	n
1	Equal	0.296	1.103	6
	Not Equal	0.296	1.103	
5	Equal	0.936	0.082	6
	Not Equal	0.936	0.082	
10	Equal	1.00	0.000	6
	Not Equal	1.00	0.000	
50	Equal	0.701	0.395	6
	Not Equal	0.701	0.395	
100	Equal	0.128	1.659	6
	Not Equal	0.128	1.659	
500	Equal	0.855	0.188	6
	Not Equal	0.855	0.188	
1000	Equal	0.966	0.044	6
	Not Equal	0.966	0.044	

Table 14. Zooplankton matrix: ANOVA test results comparing observer counts for three bead concentrations. Alpha set to 95% confidence (α =0.05).

Total Number of Beads per 500 mLs Volume	Average # Beads per mL(for each observer)	F	p-value (significance)	n = number of mLs analyzed	Total Number of Beads Counted
10	0.020	0.226	0.634	480	10
	0.017			480	8
30	0.060	0.017	0.895	500	30
	0.062			500	31
60	0.116	0.032	0.858	490	57
	0.112			490	55

5 CONCLUSIONS

The experiments conducted during this study provided increased insight in quantifying the bias that occurs in counting varying bead concentrations and the potential benefit of collecting and concentrating larger sample volumes or increased sub-sampling. Baseline accuracy and precision measurements were obtained to quantify variability involving "ideal" conditions which eliminated potential variables found in real samples that can compound variability measurements.

Phytoplankton Class (Bead Size = 10-15 microns)

- Microscopists were able to complete whole chamber (20 rows =1 mL) of nine Sedgwick-Rafter slides of concentrations at the most dilute concentrations (1, 5, 10 and 50 beads/mL) within an 8 hour time frame. This is considered the practical limit for sub-sampling under ideal conditions.
- Repeated, observed whole sample counts of 10,000 beads/mL stocks yielded accuracies within 10%, which was in agreement with the manufacturer certification. As a result, it was concluded that use of these stocks to prepare analytical solutions of varying concentrations was a valid and accomplishable approach.
- The percent difference of the observed mean from the expected mean was used as a measure of enumeration accuracy, while the coefficient of variation (ratio of the standard deviation to the mean) was related to precision. A percent difference of 10% and a CV of 0.2 were used as benchmarks for acceptable accuracy and precision, respectively.
- In general, an increasing bead concentration and increasing sub-samples analyzed resulted in a reduced standard deviation as expected, though similar trends in accuracy were not as apparent (i.e., the mean did not consistently approach the expected value with an increase in bead concentration or volume sampled).
- The goodness of fit, also reflective of accuracy, in observed versus expected counts did not definitively increase with an increase in volume sampled, though this may simply reflect systematic error in the stock concentrations or improper mixing of the sample.
- With the exception of the most dilute concentration, all observed versus expected trends exhibited relatively good linearity and goodness of fit. Moreover, the observed counts were relatively unaffected by whether the whole chamber or 10 random rows were counted. This is a potentially significant finding to reduce the overall time of sample analysis and thereby increase the number of sub-samples which may be examined.
- At the most dilute concentration, the percent difference was highly variable and did not stabilize even after nine replicates. Some improvement was noted at 5 beads/mL, but only after a minimum of 6 replicates (6 mL) were examined.
- In the case of 50, 100 and 1000 beads/mL, percent difference increases with increasing replicates which was not anticipated, as accuracy would be expected to improve with increased replication. However, this result is plausible since the percent difference amongst these at high replicates was a maximum of 11.5% which is not significantly different than the vendors reported +/- 10% variation in stock concentrations. Thus, it is possible that the expected concentration was inaccurate.
- In general, further improvements in the accuracy of the observed mean were not achieved by replication beyond 5 replicates for concentrations ≥ 50 beads/mL. A stable accuracy within the systematic error bounds of the stock concentration were not achieved with 1 or 5 beads/mL concentrations.

- At the most dilute concentrations of 1 and 5 beads/mL the CV was well over 1.00 and concentrations of 10 beads/mL resulted in a CV above 0.4. This indicates that these measurements were imprecise and that no useful conclusions could be expected at these sample concentrations with similar enumeration method. These sample concentrations represent all concentration factors tested for an anticipated discharge concentration of 1 organism/mL and a non-concentrated 10 organism/mL discharge. Hence actual test samples would be expected to be similarly imprecise.
- A substantial improvement in precision was realized when discharge samples were sufficiently concentrated to greater than or equal to 50 organisms/mL. It is recommended that a target sample concentration of at least 50 organisms/mL should be utilized. Precision was not appreciably improved at these sample concentrations after analyzing five full Sedgwick-Rafter counting chambers.
- The chi-square was shown to be a sufficient approximation to the sample test statistic and that for concentrations up to 100 organisms/mL, the Poisson distribution is valid, so there is insufficient evidence to reject the Poisson hypothesis. The Poisson model does not fit well for the two highest concentrations. Since there was consistent and good agreement between the two observers, the departures from the Poisson model at high concentrations do not appear to result from a counting problem, so they must arise from the sub-sampling. There is no way to be certain, but more thorough mixing of the prepared volume may be required at high concentrations
- There was no significant difference in the CV between the 20 row and 10 row counts for all of the bead concentrations. However significant differences were found among the variances of the bead concentrations for the 10 row counts themselves indicating that the data may need to be transformed or nonparametric tests used.
- There were no significant differences in the bead counts among observers for all of the bead concentrations.

Zooplankton Class (>50 microns) (Bead Size = 150 microns)

- The most concentrated solutions (20, 60 and 120 beads/mL) provided accuracy within 10% after sufficient 1-mL aliquots were examined to result in a stable mean.
- The comparison of the CV versus volume of sample analyzed revealed that the enumeration method utilized was imprecise at all concentrations (20, 60 and 120 beads/mL) regardless of sample volume analyzed. Only counts at 120 beads/mL resulted in a CV < 1.0.
- The sample concentrations tested represented concentration factors of 1000, 3000 and 6000 for a total whole water sample volume of 6 m^3 with an assumed discharge concentration equivalent to the IMO D-2 standard (< 10 organisms/m^3). Since these, the most concentrated samples planned, displayed high CVs, more dilute concentrations were not conducted.
- Current test protocols designed to sample discharge concentrations at or below 10 organisms/m^3 with triplicate 1 m^3 samples (or duplicate 2 m^3 samples as proposed in the most recent version of the ETV protocol) will result in inadequate precision even when concentrated by 6000 times with this enumeration method at n = 450. On the basis of these data, concentration of 6 m^3 by 6000 times and 450 1-mL aliquots analyzed would provide improved precision but below the 20% (as CV) identified as the benchmark. Because higher concentration factors are probably unrealistic and because analyzing 450 1-mL aliquots of natural samples before they degrade is not currently feasible, larger sample sizes and improved analytical methods should be considered.

5.1 Overall Conclusions/Recommendations

These microbead experiments provided ideal conditions for investigating analysis of sparse populations and determining concentration factors required. The results of these microbead experiments have several implications.

First, treated ballast water discharges which are expected to meet the IMO D-2 standard for organisms in the size class of greater than 10 µm but less than or equal to 50 µm should be concentrated at least 5 times to provide a precision of not less than 20%.

In terms of sample replicates, using the Sedgwick Rafter counting chamber method, at least four counting chambers should be analyzed to insure both adequate precision and accuracy (< 10%). Further replicates at this concentration range should not be expected to provide significantly increased precision or accuracy.

Due to the expected low concentrations for zooplankton (< 10 organisms/m^3 for the IMO D-2 standard), volumes larger than those currently stated in the ETV protocol are recommended. A 6 m^3 discharge sample would have to be concentrated by a factor of 6000 times and 450 1-mL sub-samples analyzed to provide sufficient precision (20%). It is therefore recommended that larger sample sizes and improved analytical methods be considered for actual sampling conditions.

Future efforts should examine the practicality of concentrating discharge samples of living organisms to insure that the sample remains representative (no loss of organisms or effect on mortality) and to insure that adequate time is available to both conduct the necessary concentration step and the enumeration. Note that the latter may be further frustrated if larger concentration factors are required for improved accuracy or precision.

As a caution, these conclusions and analyses must be considered with the understanding that these data are from an ideal scenario. In other words, all additional sources of error or variability have been eliminated by virtue of the experimental method: specifically, no sample degradation is anticipated, there is no need to assess viability or health, enumerated objects do not clump and are homogeneous in size and shape. With respect to actual sampling conditions, this is wholly unrealistic, and as such, the results of these experiments should be considered conservatively.

6 REFERENCES

Battelle, 2003. "Final Draft Generic Protocol for the Verification of Ballast Water Treatment Technologies", Version 2.3, Duxbury, MA.

Duke Scientific Corporation, 2463 Faber Place, P.O. Box 50005, Palo Alto California (www.dukescientific.com).
International Maritime Organization Convention BWM/CONF/36 (2004) International Convention for the Control and Management of Ships' Ballast Water and Sediments (D2). International Maritime Organization, London. UK.

International Maritime Organization G8 Resolution MEPC 125(53) (2005) Guidelines for Approval of Ballast Water Management Systems (G8). International Maritime Organization, London, UK.

Lemieux, E.J. et al (2007) Pilot Environmental Technology Verification (ETV) Test Report of the Severn Trent DeNora BalPure™ Ballast Water Treatment System. Naval Research Laboratory Center for Corrosion Engineering, Washington, DC.

(This page intentionally left blank.)