2017

Modeling sampling strategies for determination of zooplankton abundance in ballast water

M. R. Hernandez
M. L. Johansson
Y. Xiao
M. A. Lewis
Hugh J. MacIsaac

University of Windsor

Follow this and additional works at: https://scholar.uwindsor.ca/biologypub

Part of the Biology Commons

Recommended Citation
https://scholar.uwindsor.ca/biologypub/525

This Article is brought to you for free and open access by the Department of Biological Sciences at Scholarship at UWindsor. It has been accepted for inclusion in Biological Sciences Publications by an authorized administrator of Scholarship at UWindsor. For more information, please contact scholarship@uwindsor.ca.
Modeling sampling strategies for determination of zooplankton abundance in ballast water

Marco R. Hernandez¹*, Mattias L. Johansson¹*, Yanyu Xiao²,³, Mark A. Lewis², and Hugh J. MacIsaac¹

¹Great Lakes Institute for Environmental Research, University of Windsor, Windsor, Ontario, Canada;
²Department of Mathematical and Statistical Sciences and Department of Biological Sciences, University of Alberta, Edmonton, Alberta
³Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio

¹Lead authors

Corresponding author: Marco Hernandez, GLIER, University of Windsor, Windsor, Ontario N9B 3P4 email: hernand2@uwindsor.ca

Running title: Sampling strategy for zooplankton in ballast


Keywords: ballast water, modeling, International Maritime Organization, probability density function, sampling effort
Abstract

Ballast water has been one of the world’s largest sources of non-indigenous species introductions. The International Maritime Organization has proposed a performance standard that will establish a numerical limit of <10 viable individuals m\(^{-3}\) for zooplankton-sized organisms in discharged ballast. Here we test a variety of sampling efforts for zooplankton-sized organisms in post-exchange ballast water on a commercial vessel. We fit five widely-used probability density functions (PDF) to find the most representative PDF and evaluated sampling efforts necessary to achieve error rates (\(\alpha, \beta\)) of < 0.05. Our tests encompassed four seasonal trials and five sample volumes. To estimate error rates, our simulations drew from 1 to 30 replicates of each volume (0.10 - 3.00 m\(^3\)) for mean densities ranging between 1 and 20 organisms m\(^{-3}\). Field sampling revealed that >0.5 m\(^3\) samples had better accuracy and precision than other volumes tested, and that the Poisson distribution fit these communities best. Simulations of ballast sampling for all PDFs tested also revealed that the optimal and practical sample volume was >0.5 m\(^3\). This study provides the first field test of an alternative sampling strategy to assess compliance with the future IMO D-2 standard that will be applied to all large vessels.
Introduction

Ballast water is one of the world’s largest vectors for non-indigenous species (NIS) transfer [1]. Efforts to control this vector in the Great Lakes began in 1989 with voluntary mid-ocean ballast water exchange (BWE) for vessels entering with filled ballast-water tanks, which was followed by mandatory regulations in 1993. Regulations were extended to vessels with ‘empty’ ballast-water tanks in 2006 and 2008 in Canada and the USA, respectively. Ballast water management (BWM) has become a standard procedure worldwide, and is overseen by the International Maritime Organization (IMO). Current IMO best management practises request vessels with full ballast tanks conduct exchange on the open ocean to ensure that 95% of the ballast volume has been exchanged, to achieve an in-tank salinity of at least 30‰ [2]. While this procedure is effective in preventing the movement of NIS between freshwater ports that are connected by transoceanic routes [3], it is less effective when both origin and destination ports are marine [4]. In 2004 the IMO proposed new performance standards (IMO D-2) [5]. This agreement sets numerical limits on the density of two plankton size groups (< 10 viable organisms m\(^{-3}\) for minimum dimension > 50 μm and < 10 viable cells mL\(^{-1}\) for organisms between 10 and 50 μm) as well as for three bacteria indicators [5]. The IMO D-2 convention has yet to be ratified and implemented [5].

Many companies and research groups are testing technology devices and processes to ensure compliance with IMO D-2 standards. Initial steps for approval include testing of devices by an independent third party at verification facilities designed to provide bench-scale estimations, usually referred to as land-based testing. Verification centers also must replicate treatment trials as part of the bench-scale
evaluation. Sampling strategies and sampling effort are intended to be easily replicable [6]. Model ballast tanks must be ≥ 200 m³. For shipboard sampling, control and treated samples need to be collected in triplicate, that uptake and final densities be determined for control tanks, and that viable organism density be assessed before discharge of treated ballast water [7]. However, current guidelines provide no guidance on sample volumes or how they are collected.

Current technology devices have been tested primarily using land-based tests, though a subset have also used shipboard testing [8]. However, no clear method exists for sampling onboard vessels, particularly for sampling directly from ballast tanks. Thus, an imbalance exists in the prescribed sampling process for land-based versus shipboard testing. Onboard sampling poses a major challenge as the IMO D-2 standard requires very low densities of zooplankton, and estimating live density of organisms requires large sample volumes, even under the best case (and unrealistic) scenario that organisms are randomly distributed [9, 10, 11]. Moreover, random dispersion of zooplankton in ballast tanks cannot be assumed, as organisms may aggregate and thus may exhibit a patchy distribution [12, 13].

Zooplankton sampling in ballast tanks may be done using plankton nets via hatches [14, 15] or, less commonly, by pumping a known volume from the tank into a plankton net [16, 17, 8]. Sampling a ballast tank is complicated as access is limited while in port and very difficult while en route [18]. Samples must be representative of the entire population, easy to replicate, and unbiased. Another consideration is inherent stochasticity associated with low population densities, with concerns regarding both
accuracy and precision [19]. In addition, the sampling strategy must allow inferences to be made regarding densities of viable zooplankton in treated water.

A number of studies have addressed the effects of low organism density and sample volume on estimating the true density of zooplankton, using both Poisson and negative binomial distributions [9, 10, 11, 20]. The validity of this theoretical approach has not yet been affirmed empirically. The Poisson distribution is suitable under the assumption of a centralized outflow that can be sampled entirely or in equal time intervals [13]. A key challenge is access to the entire water column of a tank. Net tows likely introduce bias as only the upper portion of the tank is typically sampled.

In this study, we tested different sampling volumes using three in-tank sampling points to sample the full depth of a ballast tank on a working cargo vessel. Our goal was to identify the sampling efforts that will provide accurate density estimations of zooplankton at the very low abundances that the IMO D-2 standard requires for compliance. We also designed a simple model to contrast common distributions that have been examined theoretically to provide a sample volume that managers can utilize to verify compliance with the IMO D-2 standard.

Methods

Ballast samples were collected during voyages by the Federal Venture, between 2012 and 2013 [see 21]. The vessel transited from three ports (Saguenay, Trois Rivières, and Bécancour) in Quebec, Canada to two ports (Vila do Conde and Sao Luis) in Brazil. A single trial was conducted during each voyage where samples were taken and analyzed. Samples were collected from the largest ballast tank (Tank 2) on the
starboard side, with 25 mm diameter inlet pipes (Alfagomma 266GL Water S&D PVC Standard Duty) installed at three depths (4.5, 14.5 and 16.0 m below top deck level) to account for vertical variation in organism distribution (Fig. 1). We selected those depths based on the geometry of the tank: 4.5 m is the middle section of the attached wing tank, 14.5 m is the highest open space in the double-bottom tank, and 16.0 m is just above the baffle line in the deepest portion of the tank. Each inlet pipe contributed one third of the total sample volume. To assess sampling effort, triplicate samples totalling 0.10, 0.25, 0.50, 1.00 or 3.00 m³ were collected. Samples were collected two days after ballast-water exchange was performed in the North Atlantic region using a pneumatic, self-priming diaphragm pump. Ballast water was transferred from the tank to the forepeak of the vessel where it was filtered through a 35 μm plankton net. Water volume sampled was measured with a Seametrics flowmeter (WMP-Series Plastic-Bodied Magmeter). In-line valves were used to keep water flow rate to 40 L minute⁻¹ in order to avoid mortality due to strong currents. Samples were then fixed in 95% ethanol for microscope counting. We assumed that all intact individuals encountered when processing under the microscope were alive at the time of capture. Each sample was counted entirely to assess population density. The order in which sample volumes were collected was randomized using a random number generator in Excel (Microsoft Inc.). We conducted basic descriptive statistics (mean and standard deviation) for our four trials. Variance was grouped for fall and spring as those samples were not statistically different and mean densities were similar. Our first goal was to determine the best volume for sampling. Since the true density of organisms in the ballast tank was not known, we assumed that the mean density of organisms over all sample
volumes in each trial was an accurate estimate of true density. Preliminary analysis of variance (ANOVA) revealed that volume sampled had a large impact on the density of organisms in the tank (p=0.0056). We estimated density based on the data points collected from the same volume. We assumed that if we sampled at the same volume repeatedly inside the tank, the density of organisms would follow a given probability distribution function (PDF). We performed the following analysis on each of five PDFs (Poisson, Weibull, Negative binomial, Gamma, and Log-normal) with respect to each volume individually. We estimated the parameters of each PDF by maximum likelihood estimation (MLE). Then, we created random number generators based on the estimated PDFs to sample more data points (i.e. one thousand data points) for the density of organisms for each volume, and calculated the mean square error (MSE) based on our assumption that the true density was the average of density estimates in all trials for each volume [22].

**Modeling PDF for distribution of zooplankton**

Our second goal was to determine how altering the spatial distribution of zooplankton would affect the sampling error rate. Specifically, our objective was to identify the number of samples of a particular volume that would be required to confidently state that a vessel was compliant with the IMO D-2 limit of < 10 viable organisms m\(^{-3}\) for zooplankton-sized organisms while keeping the rate of Type I and II errors below 5%. In other words, the cumulative sample number of each individual density (from 1 to 20 organisms m\(^{-3}\)) required in each scenario was constrained to no more than a 0.05 error rate for both false positives and false negatives.
We modeled sampling from the ballast tank using a three-dimensional array in R (R Development Core Team, 2016). To simulate sampling from the tank, we defined each cell of the array as 1 L of water and the total volume of the array as approximately equal to the actual capacity of the tank used for our sampling (1,279,400 L in the actual tank, 1,300,000 L in our model 100x100x130 cell array). For each of 1000 replicates, we populated each cell in the array by drawing randomly from two commonly used PDFs (Poisson and Gamma) with mean densities from 1 to 20 organisms m$^{-3}$. For each PDF, we then sampled between 1 and 30 replicates using sampling points placed at particular heights in the array (to model our field design) but with randomly assigned length and width coordinates. In each case, we assessed the rate of false positives and false negatives for all combinations of sample volume and replicate number and determined the minimum replicate number required to achieve rates less than 5%.

For the Poisson distribution, we also tested the effect on error rates of having organisms randomly but evenly distributed in the array (Even scenario) at the target density versus organisms preferring the upper wing tank (Uneven scenario: organisms randomly distributed in the 501,400 L upper section at a much higher density [up to ~500X higher density] than the 778,000 L lower region while still achieving the same overall density as the even distribution). In addition, we modeled the effect of sampling only from the upper wing tank, as typically occurs in current working vessels. In an ideal Poisson situation with evenly distributed organisms, there should be no difference between sampling a given volume in a single large replicate versus a number of small replicates. However, because our simulations sampled randomly from a distribution, some variance between replicates occurred.
For the Gamma distribution, we simulated three different distribution shapes to test the effect of variance on our ability to accurately estimate the true density with different sample volumes and replicate numbers. In each simulation, we tested three levels of dispersion by setting the rate to 0.5, 1.0, and 2.0 to correspond with wide, medium, and narrow distributions, respectively, and then stepwise-adjusted the shape to achieve the desired mean, from 1 to 20 organisms m$^{-3}$.

**Results**

Although the vessel traversed essentially the same route from Canada to Brazil during all four trials, the geographic position of ballast-water exchange and subsequent location of sampling varied slightly from one trial to the next. Mean plankton density ranged from 285 to 1170 organisms m$^{-3}$ (horizontal lines, Fig. 2), with a clear seasonal pattern: trial 1 (July) was highest, trial 3 (November) the lowest, and trials 2 and 4 (September and March) were similar and had intermediate densities (Fig. 2). From our field sampling, it was also evident that dispersion is larger in smaller volumes and that it is generally low at volumes $> 0.50$ m$^3$ (Fig. 2).

We observed no significant difference fitting the five distribution functions in our MLE for PDFs (Fig. 3), possibly owing to our small empirical dataset (12 data points from each sample volume). We did however note that the 1.00 m$^3$ sampling volume exhibited the lowest MSE term relative to other volumes tested (Table 1).

When organisms were evenly Poisson distributed in the ballast tank, simulations exhibited a clear relationship between sample volume, replicate number, and our ability to confidently state whether the ballast tank was compliant or not. As mean density of
the sample approached the permissible limit of 10 organisms m\(^{-3}\), the total volume of samples required to assess compliance also increased (Fig. 4, upper panel).

Consequently, smaller sampling volumes reached our arbitrary limit of 30 replicates earlier than did larger ones, leading to a larger window where sample sizes were insufficient to confidently assess compliance. For example, a single 0.10 m\(^3\) sample (pink dotted line, Figure 6 upper panel) could be sufficient to identify the sample as compliant (i.e. < 10 organisms m\(^{-3}\)) if the true density was below 3 organisms m\(^{-3}\), though the number of replicates required at this volume exceeds 30 if true density was >7 organisms m\(^{-3}\). To avoid incorrectly declaring a sample compliant when the true density is at or above 10 organisms m\(^{-3}\), more than 30 samples of size 0.10 m\(^3\) would be required if the true density ranged between 10 and 14 organisms m\(^{-3}\) (i.e. just above the permissible limit). Increasing the volume of samples improves our ability to confidently assess compliance as the true density approaches the 10 organisms m\(^{-3}\) limit (dotted vertical line, Fig. 4, upper panel).

In contrast to small volume samples, those of 3.00 m\(^3\) required three or fewer replicate samples to confidently determine compliance when the true density was below 8 organisms m\(^{-3}\) or above 12 organisms m\(^{-3}\) (red long dash line, Fig. 4 upper panel), and compliance could be assessed with 11-12 replicates if true density was very close to the maximum permissible limit (i.e. 9 or 11 organisms m\(^{-3}\)). Intermediate sample sizes could be used to confidently assess compliance when the true density was <7 or >13 organisms m\(^{-3}\), but as sample volume declined, the number of replicates required increased (Fig. 4, upper panel). As expected, across the range of densities tested, total sample volume seemed to be the key determinant of our ability to confidently assess
compliance when organisms were evenly Poisson distributed. For example, at a true
density of 7 organisms m$^{-3}$, compliance could be assessed with a minimum of 24, 9, 5, 3
or 1 sample(s) for volumes of 0.10, 0.25, 0.50, 1.00, or 3.00 m$^3$, respectively.

When organisms were unevenly distributed and were sampled from the full depth
of the ballast tank (all three sampling ports), we saw a very similar pattern, though it
moved the window of non-confidence (error rate >0.05) toward false negatives (Fig. 4,
lower panel). All volumes except for 0.10 m$^3$ could be used to assess compliance when
the true density of organisms was ≤ 9 organisms m$^{-3}$ (pink dotted line, Fig. 4, lower
panel); however, when the sample volume was low (e.g. 0.25 m$^3$), a large (20) number
of replicates was required (green dashed line). The number of replicates required to
confidently assess compliance dropped progressively from 8 to 4 to 2 replicates at 0.50,
1.00 and 3.00 m$^3$ (blue dash dot dash, yellow solid, red long dash lines, respectively).
The lower total volume required for samples of 1.00 m$^3$ (4 m$^3$) versus 3.00 m$^3$ (6 m$^3$)
suggests that multiple 1.00 m$^3$ samples might be the most tractable sampling scheme,
given the time required to process samples under the microscope. The major difference
between “uneven” and “even” scenarios is that there were more true densities above
the compliance limit where we could not confidently assess compliance in the former
scenarios. At a density of 13 organisms m$^{-3}$, we could confidently assess compliance
with sample volumes of 1.00 m$^3$ (yellow solid line) and 3.00 m$^3$ (red long dash line), but
both required sampling impractically large volumes of water: 20 m$^3$ (20 samples) for
1.00 m$^3$ and 18 m$^3$ (6 samples) for 3.00 m$^3$.

In the uneven Poisson scenario, where organisms were concentrated in the top
section of the tank and only that region was sampled, (Fig. 4, lower panel) results were
quite different. As organism density in the upper portion of the tank was much higher than the overall mean density, it was very easy to overestimate mean density; consequently, large sample volumes from tanks with low overall density (i.e. <3 organisms m \(^{-3}\)) were required to achieve an acceptable rate of false positives. In contrast, it took relatively small sample volumes (i.e. 1.00 m \(^3\) total from any sample volume/replicate combination) to avoid false negatives, as few samples estimated densities lower than 10 organisms m \(^{-3}\).

Similar to the Poisson results sampled from throughout the tank, all sampling volumes with the Gamma PDF had a window of non-confidence for densities approaching the IMO D-2 standard of 10 organisms m \(^{-3}\). Overall, the relationships between different sample sizes was similar to that seen in the Poisson model, above. In all three dispersion scenarios, larger samples had narrower ranges where we failed to confidently assign compliance with reasonable replicate numbers (i.e. <30 replicates; Fig. 5). In the Gamma simulations, the key difference among the three different dispersion scenarios is that as dispersion decreased (rate increased), the range where we could not confidently assign compliance narrowed. This was most apparent in the smallest sample size (0.10 m \(^3\), Fig. 5, pink dotted line). In the highest dispersion (rate=0.5) model, we failed to confidently assign compliance for true densities from 7 to 15 organisms m \(^{-3}\), while for the intermediate dispersion (rate=1.0) model the range is 8 to 14 organisms m \(^{-3}\), and for the more aggregated organisms (rate=2.0) model the range is 9 to 12 organisms m \(^{-3}\). The other sample volumes tested exhibited a similar, if less pronounced, pattern. The other major difference was that the number of replicates for a given volume decreased with decreasing statistical dispersion. This was very
pronounced in the 3.00 m³ sample size, which maintained the same narrow range of non-confidence throughout all three rate scenarios, but required >20 replicates for confidence when dispersion was highest, 10-12 replicates at intermediate dispersion, and 5-6 replicates when dispersion was low (Fig. 5, red long dash line). This pattern of a narrowing of the non-confidence range with decreasing dispersion, and a decrease in replicates required for confidence, was consistent across all five sample volumes. Consistent with the Poisson model, the largest sample sizes again returned the narrowest range of non-confidence for tractable sample numbers.

Discussion

Even at very low densities, sampling volumes of 1.00 and 3.00 m³ were able to accurately estimate zooplankton density in ballast tanks. However, the improvement in accuracy by adding additional samples was more practical for 1.00 m³ than for 3.00 m³ samples. The 1.00 m³ samples had the lowest MSE scores in five out of six PDFs tested (all except Log-normal), and were, therefore, the most accurate of all volumes tested (Table 1 and Fig. 3).

Sampling across the water column addresses problems inherent in sampling species with patchy distributions, and is required for testing IMO D-2 compliance [6, 12]. Individual zooplankton tend to aggregate in natural waters [13] and likely do so in ballast tanks as well. Our multiport sampling design allowed us to sample the entire water column, including the double-bottom portion, which is usually inaccessible. Thus, multiple sampling ports provide more accurate estimates of organism density than single ports or if researchers use deck-based plankton nets. Although we used an equal
number of ports as Murphy et al. [12], our design allowed us to collect water from the lower portion of the tank, which is inaccessible to open hatch tow sampling. It also made possible to take as many replicate samples as desired within a short period of time without affecting vessel operations.

The Poisson distribution had the lowest MSE scores in all volumes (Table 1). The results we obtained were similar for Gamma distribution in deriving the likelihood of over dispersion due to clumping. The Poisson distribution is commonly used for modeling zooplankton distributions in ballast tanks [9, 10, 11, 20], however, the Gamma distribution also has been used as a Poisson approximation. Gamma distribution estimates abundance distributions [23] and has been suggested for zooplankton in ballast water [20]. A need exists to build data sets that allow identification of an appropriate PDF based on empirical data. Our attempt with a rather limited data set proved inconclusive.

True zooplankton densities were not known in our trials, thus we relied on a series of assumptions that justified using the mean of all sampling efforts per trial. Under these assumptions, large volume samples had higher precision and lower variability. Trials 1 and 3 also demonstrated that the largest volume (3.00 m$^3$) estimated density better than smaller ones. However, in Trials 2 and 4 large volumes underestimated densities. While larger volumes - such as 3.00 m$^3$ - provided- in general- better estimates, they increased work load prohibitively and thus cannot be recommended (see [11]). We observed that 1.00 m$^3$ samples had the lowest MSE and provided a good estimation with a low rate of false positives when organism abundance was ≤10 individuals m$^{-3}$, and a low false negative rate when density ≥10 individuals m$^{-3}$.
The error rate can be improved for estimates based on 1.00 m³ samples by increasing the number of replicates (Figs. 5 and 6). Because our sampling technique was already an integration of three equal volumes, even a single replicate enhanced accuracy of the density estimate, and replicates at this volume are manageable.

There exists support for the argument that large volume samples offer better estimations assuming Poisson-based models (e.g. see [9, 10]). However when the dispersion of organisms in the tank is unknown, there is a possibility to overestimate densities and wrongly conclude that vessels are not in compliance with the IMO D-2 standard (see Fig. 4). In our ‘uneven’ Poisson simulations, altering how animals are distributed in the tank modified not only the proportion of false positives and negatives, but the capability to accurately assess organism densities at all tested volumes. We agree with the aforementioned authors that larger volumes (e.g. 7.00 m³) provide a better estimator of density, though these volumes are impractical for organism enumeration at anything other than, and possibly including, a land-based testing facility.

Our three sampling port design provides better opportunities to accurately quantify plankton present at low density.

Our descriptive statistics highlighted that dispersion was larger on small sample volumes and decreased as volume increased (Fig. 2). Despite the non-significant difference among sampling volumes, we observed that sampling volumes below 0.50 m³ are much more variable and thus less reliable (Fig. 2). Our comparison of MSE scores for all trials and volumes demonstrated that 1.00 m³ had the smallest MSE and thus the best accuracy.
The two PDFs that we used to simulate sampling allow us to infer that when zooplankton populations are present at low densities, both 1.00 and 3.00 m$^3$ sample volumes provide good estimates of density with acceptable error rates (<0.05) versus smaller volumes.

Our study is limited by the number of trials and replicates within each sample volume, however it presents realistic working conditions and constraints likely to be encountered on ocean-going vessels. Validation procedures for IMO D-2 standard are in development. At present there exist no clear guidelines on sample volumes or sample number. We suggest 1.00 m$^3$ as a starting point and encourage collection of additional empirical data and assessment of sampling strategies.

Empirical data highlighted that integrative samples added precision to density estimations by reducing variance, and that large but practicable volumes - such as 1.00 m$^3$ - benefit from it. MSE scores for 1.00 m$^3$ were lowest regardless of which PDF was used to fit our data, suggesting that this volume most accurately estimated true density. Finally, our simulations revealed that increasing the size and number of samples improves confidence in compliance assessments, with the best tradeoff between accuracy and precision and work load seemingly optimized with 1.00 m$^3$ samples.

**Author Contributions**

MRH, MLJ and HJM designed the study and wrote the paper, MLJ, YX and MAL conducted simulations, and all authors edited the manuscript.

**Acknowledgements**
We are grateful to Fednav Inc. and the captains and crews of the Federal Venture for assistance in preparing and executing onboard experiments. Lab assistance was provided by Vishal Vara and Lucas Wilson. We acknowledge funding from Fednav Inc. and the NSERC CAISN network, and NSERC Discovery grants and Canada Research Chairs to M.A.L. and H.J.M. M.R.H. was supported by a scholarship from CONACyT Mexico.

References


5. International Maritime Organization (IMO). (2004). International convention for the control and management of ships’ ballast water and sediments [Internet]. London,


Table 1. Mean squared error (MSE*10^{-5}) computed for each probability density function and each volume (m^3). Lower values indicate less dispersion between data points and the distribution curve.

<table>
<thead>
<tr>
<th>Volume (m^3)</th>
<th>Poisson</th>
<th>Weibull</th>
<th>Negative Binomial</th>
<th>Gamma</th>
<th>Log-normal</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.10</td>
<td>1.2981</td>
<td>2.5946</td>
<td>2.5350</td>
<td>2.5364</td>
<td>2.7047</td>
</tr>
<tr>
<td>0.25</td>
<td>2.0119</td>
<td>3.9496</td>
<td>4.0674</td>
<td>4.0826</td>
<td>4.7422</td>
</tr>
<tr>
<td>0.50</td>
<td>1.6707</td>
<td>3.2963</td>
<td>4.0197</td>
<td>4.1046</td>
<td>6.3578</td>
</tr>
<tr>
<td>1.00</td>
<td>0.7853</td>
<td>1.5300</td>
<td>1.7222</td>
<td>1.7800</td>
<td>2.3707</td>
</tr>
<tr>
<td>3.00</td>
<td>1.4096</td>
<td>2.8947</td>
<td>3.2303</td>
<td>3.2271</td>
<td>5.5991</td>
</tr>
</tbody>
</table>
**List of Figures**

Figure 1: Location of sampling ports inside the ballast tank.

Figure 2: Densities estimated from all four trials and five sampling efforts. Markers (diamonds – Trial 1, squares – Trial 2, triangles – Trial 3, and circles – Trial 4) indicate mean volume (n=3) ± one standard deviation.

Figure 5: Box and whisker plot for maximum likelihood of six probability density function testing 1.00 m³ sample volumes.

Figure 6: Minimum sample numbers required at a given animal density and sample volume to achieve < 5% false positive/false negative rate for Poisson-distributed organisms. False positives are shown to the left of the midline, false negatives to the right. The central gap indicates that the minimum sample number required exceeds our arbitrary cutoff of 30 replicates at a given volume. The upper panel represents a case where organisms are evenly distributed throughout the tank. Middle panel shows the case where organisms favor the upper 1/3 of the tank and sampling is through three sampling ports (as in our field experiment). In the bottom panel, organisms are aggregated in the upper 1/3 of the tank and sampling is restricted to the upper portion of the tank.

Figure 7: Minimum sample numbers required at a given animal density and sample volume to achieve < 5% false positive/false negative rate for Gamma-distributed organisms. False positives are shown to the left of the midline, false negatives to the right. Panels represent high-dispersion (top, rate=0.5), moderate-dispersion (middle, rate=1), and low-dispersion (bottom, rate=2) scenarios.
Midship section

Figure 1.

Top sampling port

Upper 'wing' tank

Middle sampling port

Lower 'double bottom' tank

Bottom sampling port
Figure 2.

- Trial 1
- Trial 2
- Trial 3
- Trial 4

Organisms m$^{-3}$

Volume sampled in m$^3$
Figure 3.

Probability density function (PDF)
Figure 4.

False positive vs. False negative in different conditions:

- Even
- Uneven
- Upper only

Number of samples vs. True density (Organisms m$^{-3}$)

Sample volume:
- 0.10 m$^3$
- 0.25 m$^3$
- 0.50 m$^3$
- 1.00 m$^3$
- 3.00 m$^3$
Figure 5.