View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Scholarship at UWindsor

University of Windsor

Scholarship at UWindsor

Biological Sciences Publications

Department of Biological Sciences

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

Hernandez, M. R.; Johansson, M. L.; Xiao, Y.; Lewis, M. A.; and MacIsaac, Hugh J., "Modeling sampling strategies for determination of zooplankton abundance in ballast water" (2017). *Marine Pollution Bulletin*, 115, 2, 80-85.
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.

1 # Modeling sampling strategies for determination of zooplankton abundance in ballast
2 water

3 Marco R. Hernandez[1*], Mattias L. Johansson[1*], Yanyu Xiao[2,3], Mark A. Lewis[2],
4 and Hugh J. MacIsaac[1]
5 [1]Great Lakes Institute for Environmental Research, University of Windsor, Windsor,
6 Ontario, Canada;
7 [2] Department of Mathematical and Statistical Sciences and Department of Biological
8 Sciences, University of Alberta, Edmonton, Alberta
9 [3] Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio
10 [*]Lead authors
11
12 **Corresponding author:** Marco Hernandez, GLIER, University of Windsor, Windsor,
13 Ontario N9B 3P4 email: hernand2@uwindsor.ca
14

15 **Running title:** Sampling strategy for zooplankton in ballast
16 **Word count:** Abstract: 200, main text: 3976, acknowledgements: 66, references: 725,
17 tables and figures: 321
18 **Keywords**: ballast water, modeling, International Maritime Organization, probability
19 density function, sampling effort
20

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 (α, β) 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 - 3.00m^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.

40 **Introduction**

41 Ballast water is one of the world's largest vectors for non-indigenous species
42 (NIS) transfer [1]. Efforts to control this vector in the Great Lakes began in 1989 with
43 voluntary mid-ocean ballast water exchange (BWE) for vessels entering with filled
44 ballast-water tanks, which was followed by mandatory regulations in 1993. Regulations
45 were extended to vessels with 'empty' ballast-water tanks in 2006 and 2008 in Canada
46 and the USA, respectively. Ballast water management (BWM) has become a standard
47 procedure worldwide, and is overseen by the International Maritime Organization (IMO).
48 Current IMO best management practises request vessels with full ballast tanks conduct
49 exchange on the open ocean to ensure that 95% of the ballast volume has been
50 exchanged, to achieve an in-tank salinity of at least 30‰ [2]. While this procedure is
51 effective in preventing the movement of NIS between freshwater ports that are
52 connected by transoceanic routes [3], it is less effective when both origin and
53 destination ports are marine [4]. In 2004 the IMO proposed new performance standards
54 (IMO D-2) [5]. This agreement sets numerical limits on the density of two plankton size
55 groups (< 10 viable organisms m^{-3} for minimum dimension > 50 μm and < 10 viable
56 cells mL^{-1} for organisms between 10 and 50 μm) as well as for three bacteria indicators
57 [5]. The IMO D-2 convention has yet to be ratified and implemented [5].

58 Many companies and research groups are testing technology devices and
59 processes to ensure compliance with IMO D-2 standards. Initial steps for approval
60 include testing of devices by an independent third party at verification facilities designed
61 to provide bench-scale estimations, usually referred to as land-based testing.
62 Verification centers also must replicate treatment trials as part of the bench-scale

63 evaluation. Sampling strategies and sampling effort are intended to be easily replicable
64 [6]. Model ballast tanks must be ≥ 200 m^3. For shipboard sampling, control and treated
65 samples need to be collected in triplicate, that uptake and final densities be determined
66 for control tanks, and that viable organism density be assessed before discharge of
67 treated ballast water [7]. However, current guidelines provide no guidance on sample
68 volumes or how they are collected.

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

79 Zooplankton sampling in ballast tanks may be done using plankton nets via
80 hatches [14, 15] or, less commonly, by pumping a known volume from the tank into a
81 plankton net [16, 17, 8]. Sampling a ballast tank is complicated as access is limited
82 while in port and very difficult while en route [18]. Samples must be representative of the
83 entire population, easy to replicate, and unbiased. Another consideration is inherent
84 stochasticity associated with low population densities, with concerns regarding both

85 accuracy and precision [19]. In addition, the sampling strategy must allow inferences to
86 be made regarding densities of viable zooplankton in treated water.
87 A number of studies have addressed the effects of low organism density and
88 sample volume on estimating the true density of zooplankton, using both Poisson and
89 negative binomial distributions [9, 10, 11, 20]. The validity of this theoretical approach
90 has not yet been affirmed empirically. The Poisson distribution is suitable under the
91 assumption of a centralized outflow that can be sampled entirely or in equal time
92 intervals [13]. A key challenge is access to the entire water column of a tank. Net tows
93 likely introduce bias as only the upper portion of the tank is typically sampled.
94 In this study, we tested different sampling volumes using three in-tank sampling
95 points to sample the full depth of a ballast tank on a working cargo vessel. Our goal was
96 to identify the sampling efforts that will provide accurate density estimations of
97 zooplankton at the very low abundances that the IMO D-2 standard requires for
98 compliance. We also designed a simple model to contrast common distributions that
99 have been examined theoretically to provide a sample volume that managers can utilize
100 to verify compliance with the IMO D-2 standard.
101
102 **Methods**
103 Ballast samples were collected during voyages by the Federal Venture, between
104 2012 and 2013 [see 21]. The vessel transited from three ports (Saguenay, Trois
105 Rivières, and Bécancour) in Quebec, Canada to two ports (Vila do Conde and Sao Luis)
106 in Brazil. A single trial was conducted during each voyage where samples were taken
107 and analyzed. Samples were collected from the largest ballast tank (Tank 2) on the

108 starboard side, with 25 mm diameter inlet pipes (Alfagomma 266GL Water S&D PVC
109 Standard Duty) installed at three depths (4.5, 14.5 and 16.0 m below top deck level) to
110 account for vertical variation in organism distribution (Fig. 1). We selected those depths
111 based on the geometry of the tank: 4.5 m is the middle section of the attached wing
112 tank, 14.5 m is the highest open space in the double-bottom tank, and 16.0 m is just
113 above the baffle line in the deepest portion of the tank. Each inlet pipe contributed one
114 third of the total sample volume. To assess sampling effort, triplicate samples totalling
115 0.10, 0.25, 0.50, 1.00 or 3.00 m^3 were collected. Samples were collected two days after
116 ballast-water exchange was performed in the North Atlantic region using a pneumatic,
117 self-priming diaphragm pump. Ballast water was transferred from the tank to the
118 forepeak of the vessel where it was filtered through a 35 μm plankton net. Water volume
119 sampled was measured with a Seametrics flowmeter (WMP-Series Plastic-Bodied
120 Magmeter). In-line valves were used to keep water flow rate to 40 L $minute^{-1}$ in order to
121 avoid mortality due to strong currents. Samples were then fixed in 95% ethanol for
122 microscope counting. We assumed that all intact individuals encountered when
123 processing under the microscope were alive at the time of capture. Each sample was
124 counted entirely to assess population density. The order in which sample volumes were
125 collected was randomized using a random number generator in Excel (Microsoft Inc.).
126 We conducted basic descriptive statistics (mean and standard deviation) for our
127 four trials. Variance was grouped for fall and spring as those samples were not
128 statistically different and mean densities were similar. Our first goal was to determine
129 the best volume for sampling. Since the true density of organisms in the ballast tank
130 was not known, we assumed that the mean density of organisms over all sample

131 volumes in each trial was an accurate estimate of true density. Preliminary analysis of
132 variance (ANOVA) revealed that volume sampled had a large impact on the density of
133 organisms in the tank ($p=0.0056$). We estimated density based on the data points
134 collected from the same volume. We assumed that if we sampled at the same volume
135 repeatedly inside the tank, the density of organisms would follow a given probability
136 distribution function (PDF). We performed the following analysis on each of five PDFs
137 (Poisson, Weibull, Negative binomial, Gamma, and Log-normal) with respect to each
138 volume individually. We estimated the parameters of each PDF by maximum likelihood
139 estimation (MLE). Then, we created random number generators based on the estimated
140 PDFs to sample more data points (i.e. one thousand data points) for the density of
141 organisms for each volume, and calculated the mean square error (MSE) based on our
142 assumption that the true density was the average of density estimates in all trials for
143 each volume [22].
144
145 *Modeling PDF for distribution of zooplankton*
146 Our second goal was to determine how altering the spatial distribution of
147 zooplankton would affect the sampling error rate. Specifically, our objective was to
148 identify the number of samples of a particular volume that would be required to
149 confidently state that a vessel was compliant with the IMO D-2 limit of < 10 viable
150 organisms m^{-3} for zooplankton-sized organisms while keeping the rate of Type I and II
151 errors below 5%. In other words, the cumulative sample number of each individual
152 density (from 1 to 20 organisms m^{-3}) required in each scenario was constrained to no
153 more than a 0.05 error rate for both false positives and false negatives.

154 We modeled sampling from the ballast tank using a three-dimensional array in R
155 (R Development Core Team, 2016). To simulate sampling from the tank, we defined
156 each cell of the array as 1 L of water and the total volume of the array as approximately
157 equal to the actual capacity of the tank used for our sampling (1,279,400 L in the actual
158 tank, 1,300,000 L in our model 100x100x130 cell array). For each of 1000 replicates, we
159 populated each cell in the array by drawing randomly from two commonly used PDFs
160 (Poisson and Gamma) with mean densities from 1 to 20 organisms m^{-3}. For each PDF,
161 we then sampled between 1 and 30 replicates using sampling points placed at particular
162 heights in the array (to model our field design) but with randomly assigned length and
163 width coordinates. In each case, we assessed the rate of false positives and false
164 negatives for all combinations of sample volume and replicate number and determined
165 the minimum replicate number required to achieve rates less than 5%.
166 For the Poisson distribution, we also tested the effect on error rates of having
167 organisms randomly but evenly distributed in the array (Even scenario) at the target
168 density versus organisms preferring the upper wing tank (Uneven scenario: organisms
169 randomly distributed in the 501,400 L upper section at a much higher density [up to
170 ~500X higher density] than the 778,000 L lower region while still achieving the same
171 overall density as the even distribution). In addition, we modeled the effect of sampling
172 only from the upper wing tank, as typically occurs in current working vessels. In an ideal
173 Poisson situation with evenly distributed organisms, there should be no difference
174 between sampling a given volume in a single large replicate versus a number of small
175 replicates. However, because our simulations sampled randomly from a distribution,
176 some variance between replicates occurred.

177 For the Gamma distribution, we simulated three different distribution shapes to
178 test the effect of variance on our ability to accurately estimate the true density with
179 different sample volumes and replicate numbers. In each simulation, we tested three
180 levels of dispersion by setting the rate to 0.5, 1.0, and 2.0 to correspond with wide,
181 medium, and narrow distributions, respectively, and then stepwise-adjusted the shape
182 to achieve the desired mean, from 1 to 20 organisms m^{-3}.
183

184 **Results**

185 Although the vessel traversed essentially the same route from Canada to Brazil
186 during all four trials, the geographic position of ballast-water exchange and subsequent
187 location of sampling varied slightly from one trial to the next. Mean plankton density
188 ranged from 285 to 1170 organisms m^{-3} (horizontal lines, Fig. 2), with a clear seasonal
189 pattern: trial 1 (July) was highest, trial 3 (November) the lowest, and trials 2 and 4
190 (September and March) were similar and had intermediate densities (Fig. 2). From our
191 field sampling, it was also evident that dispersion is larger in smaller volumes and that it
192 is generally low at volumes> 0.50 m^3 (Fig. 2).
193 We observed no significant difference fitting the five distribution functions in our
194 MLE for PDFs (Fig. 3), possibly owing to our small empirical dataset (12 data points
195 from each sample volume). We did however note that the 1.00m^3 sampling volume
196 exhibited the lowest MSE term relative to other volumes tested (Table 1).
197 When organisms were evenly Poisson distributed in the ballast tank, simulations
198 exhibited a clear relationship between sample volume, replicate number, and our ability
199 to confidently state whether the ballast tank was compliant or not. As mean density of

200 the sample approached the permissible limit of 10 organisms m^{-3}, the total volume of
201 samples required to assess compliance also increased (Fig. 4, upper panel).
202 Consequently, smaller sampling volumes reached our arbitrary limit of 30 replicates
203 earlier than did larger ones, leading to a larger window where sample sizes were
204 insufficient to confidently assess compliance. For example, a single 0.10 m^3 sample
205 (pink dotted line, Figure 6 upper panel) could be sufficient to identify the sample as
206 compliant (i.e. < 10 organisms m^{-3}) if the true density was below 3 organisms m^{-3},
207 though the number of replicates required at this volume exceeds 30 if true density was
208 >7 organisms m^{-3}. To avoid incorrectly declaring a sample compliant when the true
209 density is at or above 10 organisms m^{-3}, more than 30 samples of size 0.10 m^3 would
210 be required if the true density ranged between 10 and 14 organisms m^{-3} (i.e. just above
211 the permissible limit). Increasing the volume of samples improves our ability to
212 confidently assess compliance as the true density approaches the 10 organisms m^{-3}
213 limit (dotted vertical line, Fig. 4, upper panel).

214 In contrast to small volume samples, those of 3.00 m^3 required three or fewer
215 replicate samples to confidently determine compliance when the true density was below
216 8 organisms m^{-3} or above 12 organisms m^{-3} (red long dash line, Fig. 4 upper panel), and
217 compliance could be assessed with 11-12 replicates if true density was very close to the
218 maximum permissible limit (i.e. 9 or 11 organisms m^{-3}). Intermediate sample sizes could
219 be used to confidently assess compliance when the true density was <7 or >13
220 organisms m^{-3}, but as sample volume declined, the number of replicates required
221 increased (Fig. 4, upper panel). As expected, across the range of densities tested, total
222 sample volume seemed to be the key determinant of our ability to confidently assess

223 compliance when organisms were evenly Poisson distributed. For example, at a true
224 density of 7 organisms m^{-3}, compliance could be assessed with a minimum of 24, 9, 5, 3
225 or 1 sample(s) for volumes of 0.10, 0.25, 0.50, 1.00, or 3.00 m^3, respectively.

226 When organisms were unevenly distributed and were sampled from the full depth
227 of the ballast tank (all three sampling ports), we saw a very similar pattern, though it
228 moved the window of non-confidence (error rate >0.05) toward false negatives (Fig. 4,
229 lower panel). All volumes except for 0.10 m^3 could be used to assess compliance when
230 the true density of organisms was ≤ 9 organisms m^{-3} (pink dotted line, Fig. 4, lower
231 panel); however, when the sample volume was low (e.g. 0.25 m^3), a large (20) number
232 of replicates was required (green dashed line). The number of replicates required to
233 confidently assess compliance dropped progressively from 8 to 4 to 2 replicates at 0.50,
234 1.00 and 3.00 m^3 (blue dash dot dash, yellow solid, red long dash lines, respectively).
235 The lower total volume required for samples of 1.00 m^3 (4 m^3) versus 3.00 m^3 (6 m^3)
236 suggests that multiple 1.00 m^3 samples might be the most tractable sampling scheme,
237 given the time required to process samples under the microscope. The major difference
238 between “uneven” and “even” scenarios is that there were more true densities above
239 the compliance limit where we could not confidently assess compliance in the former
240 scenarios. At a density of 13 organisms m^{-3}, we could confidently assess compliance
241 with sample volumes of 1.00 m^3 (yellow solid line) and 3.00 m^3 (red long dash line), but
242 both required sampling impractically large volumes of water: 20 m^3 (20 samples) for
243 1.00 m^3 and 18 m^3 (6 samples) for 3.00 m^3.

244 In the uneven Poisson scenario, where organisms were concentrated in the top
245 section of the tank and only that region was sampled, (Fig. 4, lower panel) results were

246 quite different. As organism density in the upper portion of the tank was much higher
247 than the overall mean density, it was very easy to overestimate mean density;
248 consequently, large sample volumes from tanks with low overall density (i.e. <3
249 organisms m^{-3}) were required to achieve an acceptable rate of false positives. In
250 contrast, it took relatively small sample volumes (i.e. 1.00 m^3 total from any sample
251 volume/replicate combination) to avoid false negatives, as few samples estimated
252 densities lower than 10 organisms m^{-3}.

253 Similar to the Poisson results sampled from throughout the tank, all sampling
254 volumes with the Gamma PDF had a window of non-confidence for densities
255 approaching the IMO D-2 standard of 10 organisms m^{-3}. Overall, the relationships
256 between different sample sizes was similar to that seen in the Poisson model, above. In
257 all three dispersion scenarios, larger samples had narrower ranges where we failed to
258 confidently assign compliance with reasonable replicate numbers (i.e. <30 replicates;
259 Fig. 5). In the Gamma simulations, the key difference among the three different
260 dispersion scenarios is that as dispersion decreased (rate increased), the range where
261 we could not confidently assign compliance narrowed. This was most apparent in the
262 smallest sample size (0.10 m^3, Fig. 5, pink dotted line). In the highest dispersion
263 (rate=0.5) model, we failed to confidently assign compliance for true densities from 7 to
264 15 organisms m^{-3}, while for the intermediate dispersion (rate=1.0) model the range is 8
265 to 14 organisms m^{-3}, and for the more aggregated organisms (rate=2.0) model the
266 range is 9 to 12 organisms m^{-3}. The other sample volumes tested exhibited a similar, if
267 less pronounced, pattern. The other major difference was that the number of replicates
268 for a given volume decreased with decreasing statistical dispersion. This was very

269 pronounced in the 3.00 m^3 sample size, which maintained the same narrow range of
270 non-confidence throughout all three rate scenarios, but required >20 replicates for
271 confidence when dispersion was highest, 10-12 replicates at intermediate dispersion,
272 and 5-6 replicates when dispersion was low (Fig. 5, red long dash line). This pattern of a
273 narrowing of the non-confidence range with decreasing dispersion, and a decrease in
274 replicates required for confidence, was consistent across all five sample volumes.
275 Consistent with the Poisson model, the largest sample sizes again returned the
276 narrowest range of non-confidence for tractable sample numbers.
277

278 **Discussion**

279 Even at very low densities, sampling volumes of 1.00 and 3.00 m^3 were able to
280 accurately estimate zooplankton density in ballast tanks. However, the improvement in
281 accuracy by adding additional samples was more practical for 1.00 m^3 than for 3.00 m^3
282 samples. The1.00 m^3 samples had the lowest MSE scores in five out of six PDFs tested
283 (all except Log-normal), and were, therefore, the most accurate of all volumes tested
284 (Table 1 and Fig. 3).

285 Sampling across the water column addresses problems inherent in sampling
286 species with patchy distributions, and is required for testing IMO D-2 compliance [6, 12].
287 Individual zooplankton tend to aggregate in natural waters [13] and likely do so in ballast
288 tanks as well. Our multiport sampling design allowed us to sample the entire water
289 column, including the double-bottom portion, which is usually inaccessible. Thus,
290 multiple sampling ports provide more accurate estimates of organism density than
291 single ports or if researchers use deck-based plankton nets. Although we used an equal

292 number of ports as Murphy *et al*. [12], our design allowed us to collect water from the
293 lower portion of the tank, which is inaccessible to open hatch tow sampling. It also made
294 possible to take as many replicate samples as desired within a short period of time
295 without affecting vessel operations.

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

305 True zooplankton densities were not known in our trials, thus we relied on a
306 series of assumptions that justified using the mean of all sampling efforts per trial.
307 Under these assumptions, large volume samples had higher precision and lower
308 variability. Trials 1 and 3 also demonstrated that the largest volume (3.00 m^3) estimated
309 density better than smaller ones. However, in Trials 2 and 4 large volumes
310 underestimated densities. While larger volumes - such as 3.00 m^3 - provided- in
311 general- better estimates, they increased work load prohibitively and thus cannot be
312 recommended (see [11]). We observed that 1.00 m^3 samples had the lowest MSE and
313 provided a good estimation with a low rate of false positives when organism abundance
314 was ≤10 individuals m^{-3}, and a low false negative rate when density ≥10 individuals m^{-3}

315 for the two PDFs evaluated here. The error rate can be improved for estimates based
316 on 1.00 m^3 samples by increasing the number of replicates (Figs. 5 and 6). Because our
317 sampling technique was already an integration of three equal volumes, even a single
318 replicate enhanced accuracy of the density estimate, and replicates at this volume are
319 manageable.

320 There exists support for the argument that large volume samples offer better
321 estimations assuming Poisson-based models (e.g. see [9, 10]). However when the
322 dispersion of organisms in the tank is unknown, there is a possibility to overestimate
323 densities and wrongly conclude that vessels are not in compliance with the IMO D-2
324 standard (see Fig. 4). In our 'uneven' Poisson simulations, altering how animals are
325 distributed in the tank modified not only the proportion of false positives and negatives,
326 but the capability to accurately assess organism densities at all tested volumes. We
327 agree with the aforementioned authors that larger volumes (e.g 7.00 m^3) provide a
328 better estimator of density, though these volumes are impractical for organism
329 enumeration at anything other than, and possibly including, a land-based testing facility.
330 Our three sampling port design provides better opportunities to accurately quantify
331 plankton present at low density.

332 Our descriptive statistics highlighted that dispersion was larger on small sample
333 volumes and decreased as volume increased (Fig. 2). Despite the non-significant
334 difference among sampling volumes, we observed that sampling volumes below 0.50
335 m^3 are much more variable and thus less reliable (Fig. 2). Our comparison of MSE
336 scores for all trials and volumes demonstrated that 1.00 m^3 had the smallest MSE and
337 thus the best accuracy.

338 The two PDFs that we used to simulate sampling allow us to infer that when
339 zooplankton populations are present at low densities, both 1.00 and 3.00 m^3 sample
340 volumes provide good estimates of density with acceptable error rates (<0.05) versus
341 smaller volumes.

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

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

355

356 **Author Contributions**

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

359

360 **Acknowledgements**

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

367

368 **References**

369 1. Molnar, J. L., Gamboa, R. L., Revenga, C., & Spalding, M. D. (2008). Assessing the
370 global threat of invasive species to marine biodiversity. *Frontiers in Ecology and*
371 *the Environment*, **6**, 485-492.

372 2. IMO. (2008b) Guidelines for ballast water exchange (G6) [Internet]. London, United
373 Kingdom: International Maritime Organization [accessed 2015 December 7].
374 Available from http://globallast.imo.org/wp-content/uploads/2015/01/G6-
375 GUIDELINES-FOR-BALLAST-WATER-EXCHANGE.pdf

376 3. Bailey, S. A., Deneau, M. G., Jean, L., Wiley, C. J., Leung, B., & MacIsaac, H. J.
377 (2011). Evaluating efficacy of an environmental policy to prevent biological
378 invasions. *Environmental Science & Technology*, **45**, 2554-2561.

379 4. Wonham, M. J., Walton, W. C., Ruiz, G. M., Frese, A. M., & Galil, B. S. (2001). Going
380 to the source: role of the invasion pathway in determining potential invaders.
381 *Marine Ecology Progress Series*, **215**, 1-12.

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

384 United Kingdom: International Maritime Organization [accessed 2015 December
385 7]. Available from http://www.imo.org/en/About/Conventions/ListOfConventions/
386 Pages/International-Convention-for-the-Control-and-Management-of-Ships'-
387 Ballast-Water-and-Sediments-(BWM).aspx

388 6. IMO. (2008a) Guidelines for ballast water sampling (G2) [Internet]. London, United
389 Kingdom: International Maritime Organization [accessed 2015 December 7].
390 Available from http://globallast.imo.org/wp-content/uploads/2015/01/G2-
391 GUIDELINES-FOR-BALLAST-WATER-SAMPLING.pdf

392 7. IMO. (2008c) Guidelines for approval of ballast water management systems (G8).
393 London, United Kingdom: International Maritime Organization [accessed 2015
394 December 7]. Available from http://globallast.imo.org/wp-
395 content/uploads/2015/01/G8-GUIDELINES-FOR-APPROVAL-OF-BALLAST-
396 WATER-MANAGEMENT-SYSTEMS.pdf

397 8. Gollasch, S., & David, M. (2010). Testing sample representativeness of a ballast
398 water discharge and developing methods for indicative analysis. European
399 Maritime Safety Association (EMSA). Report No. 4.

400 9. Lee II, H., Reusser, D.A., Frazier, M., & Ruiz, G. (2010). Density Matters: Review of
401 Approaches to Setting Organism-Based Ballast Water Discharge Standards. U.S.
402 EPA, Office of Research and Development, National Health and Environmental
403 Effects Research Laboratory, Western Ecology Division. EPA/600/R-10/031.

404 10. Miller, A. W., Frazier, M., Smith, G. E., Perry, E. S., Ruiz, G. M., & Tamburri, M. N.
405 (2011). Enumerating sparse organisms in ships' ballast water: why counting to 10
406 is not so easy. *Environmental Science & Technology*, **45**, 3539-3546.

407 11. Frazier, M., Miller, A. W., Lee, H., & Reusser, D. A. (2013). Counting at low
408 concentrations: the statistical challenges of verifying ballast water discharge
409 standards. *Ecological Applications*, **23**, 339-351.
410 12. Murphy, K. R., Ritz, D., & Hewitt, C. L. (2002). Heterogeneous zooplankton
411 distribution in a ship's ballast tanks. *Journal of Plankton Research*, **24**, 729-734.
412 13. First, M. R., Robbins-Wamsley, S. H., Riley, S. C., Moser, C. S., Smith, G. E.,
413 Tamburri, M. N., & Drake, L. A. (2013). Stratification of living organisms in ballast
414 tanks: how do organism concentrations vary as ballast water is discharged?
415 *Environmental Science & Technology*, **47**, 4442-4448.
416 14. Briski, E., Bailey, S. A., Casas-Monroy, O., DiBacco, C., Kaczmarska, I., Lawrence,
417 Nasmith, L. E. (2013). Taxon-and vector-specific variation in species richness and
418 abundance during the transport stage of biological invasions. *Limnology and*
419 *Oceanography*, **58**, 1361-1372.
420 15. Simard, N., Plourde, S., Gilbert, M., & Gollasch, S. (2011). Net efficacy of open
421 ocean ballast water exchange on plankton communities. *Journal of Plankton*
422 *Research*, **33**, 1378-1395.
423 16. McCollin, T., Shanks, A. M., & Dunn, J. (2008). Changes in zooplankton abundance
424 and diversity after ballast water exchange in regional seas. *Marine Pollution*
425 *Bulletin*, **56**, 834-844.
426 17. Veldhuis, M. J., Fuhr, F., Boon, J. P., & Ten Hallers-Tjabbers, C. C. (2006).
427 Treatment of ballast water; how to test a system with a modular concept?
428 *Environmental Technology*, **27**, 909-921.

429 18. Wright, D. A., & Mackey, T. P. (2006). Shipboard and dockside trials of ballast water
430 treatment technology. *Naval Engineers Journal*, **118**, 37-43.
431 19. Lemieux, E. J., Robbins, S., Burns, K., Ratcliff, S., & Herring, P. (2008). Evaluation
432 of representative sampling for rare populations using microbeads (No. CG-D-03-
433 08). Coast Guard Washington DC Office of Research and Development.
434 20. Costa, E. G., Lopes, R. M., & Singer, J. M. (2015). Implications of heterogeneous
435 distributions of organisms on ballast water sampling. *Marine Pollution Bulletin*, **91**,
436 280-287.
437 21. Paolucci, E. M., Hernandez, M. R., Potapov, A., Lewis, M. A., & MacIsaac, H. J.
438 (2015). Hybrid system increases efficiency of ballast water treatment. *Journal of*
439 *Applied Ecology*, **52**, 348-357.
440 22. Walther, B. A., & Moore, J. L. (2005). The concepts of bias, precision and accuracy,
441 and their use in testing the performance of species richness estimators, with a
442 literature review of estimator performance. *Ecography*, **28**, 815-829.
443 23. Engen, S., & Lande, R. (1996). Population dynamic models generating species
444 abundance distributions of the gamma type. *Journal of Theoretical Biology*, **178**,
445 325-331.
446
447

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

Volume (m^3)	Poisson	Weibull	Negative Binomial	Gamma	Log-normal
0.10	1.2981	2.5946	2.5350	2.5364	2.7047
0.25	2.0119	3.9496	4.0674	4.0826	4.7422
0.50	1.6707	3.2963	4.0197	4.1046	6.3578
1.00	0.7853	1.5300	1.7222	1.7800	2.3707
3.00	1.4096	2.8947	3.2303	3.2271	5.5991

451

452

453 **List of Figures**

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

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

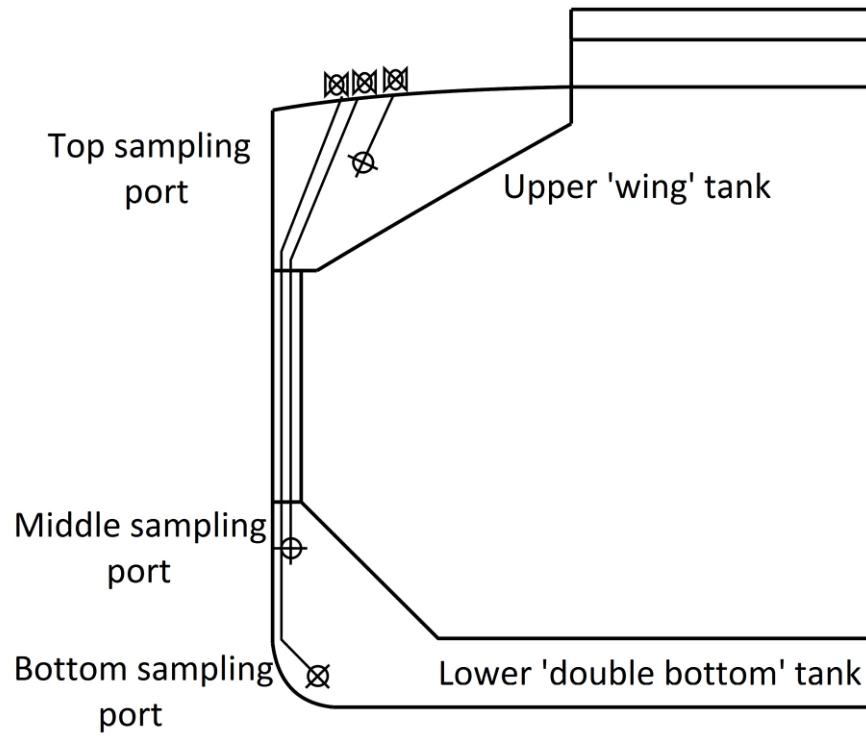
458 Figure 5: Box and whisker plot for maximum likelihood of six probability density function
459 testing 1.00 m^3 sample volumes.

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

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

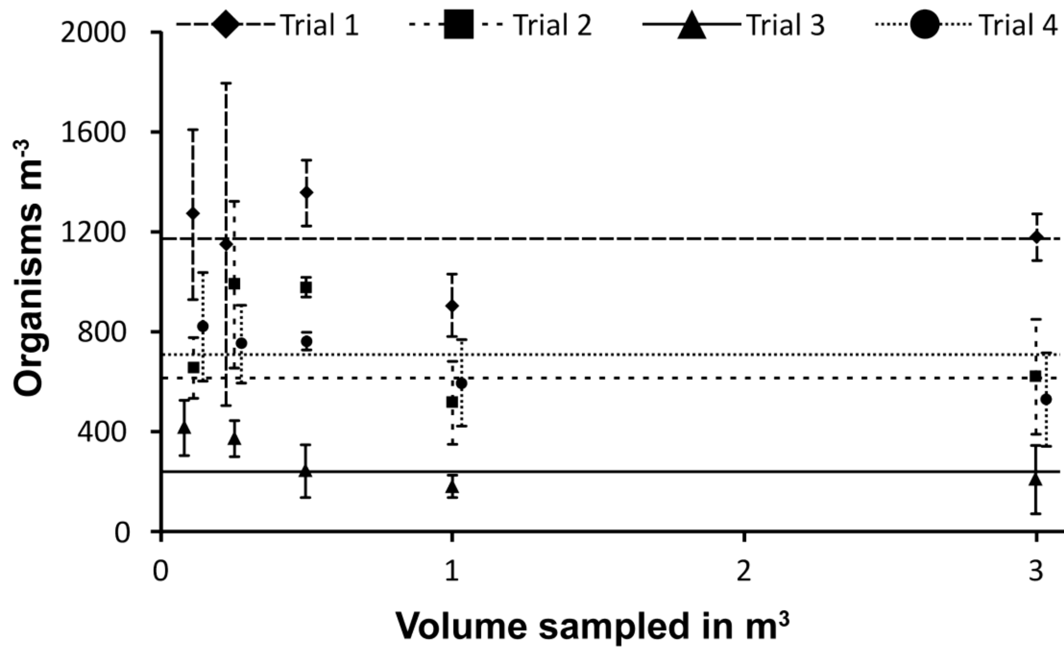
475

476
477
478



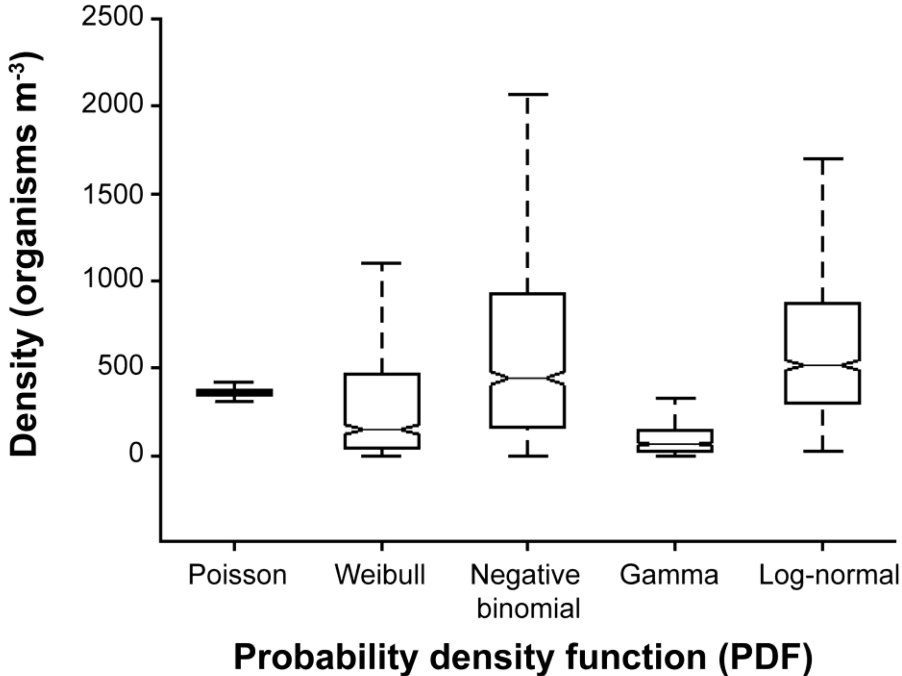
479
480
481 Figure 1.
482

483
484
485



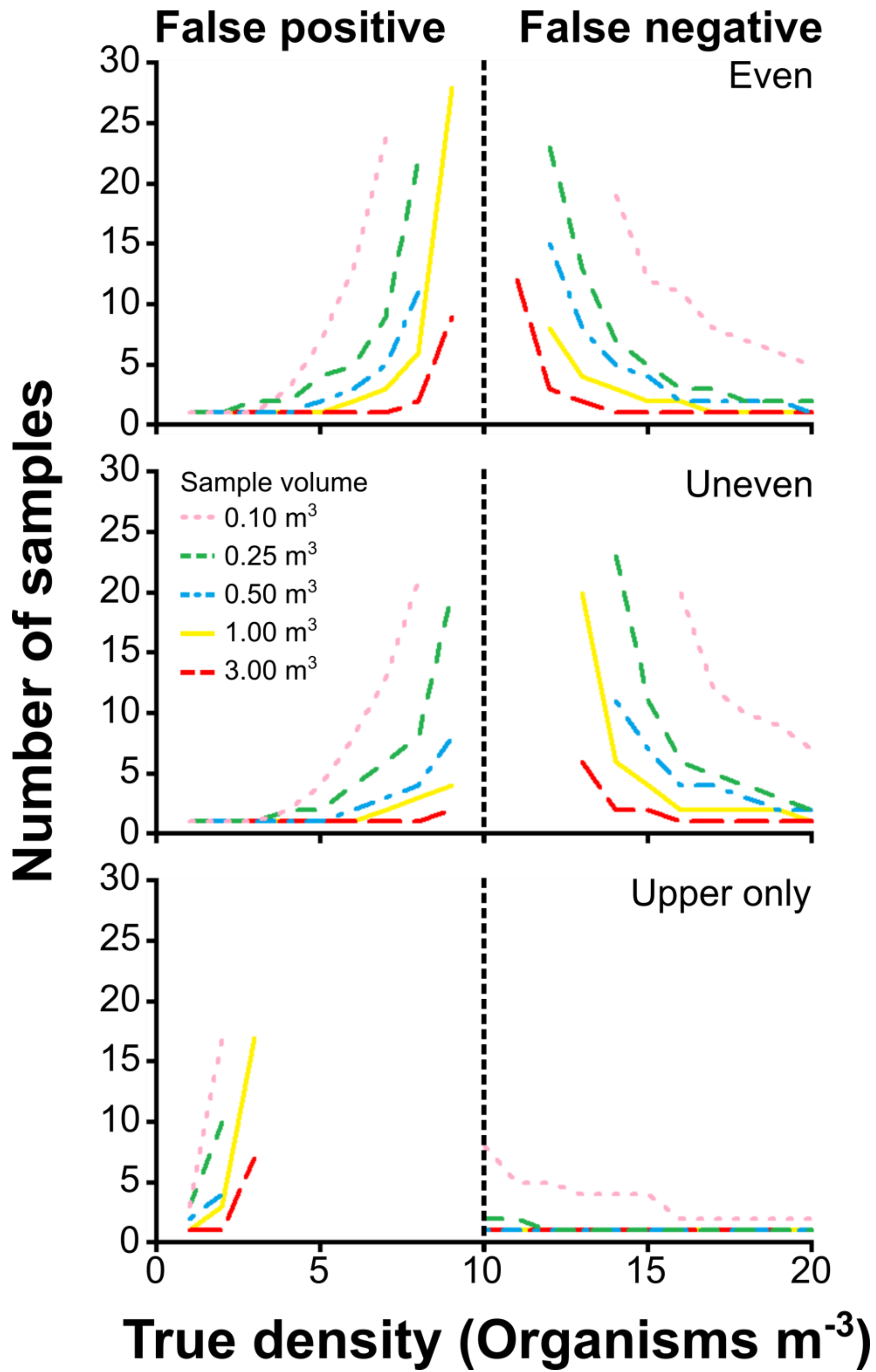
486
487
488 Figure 2.
489

490
491
492



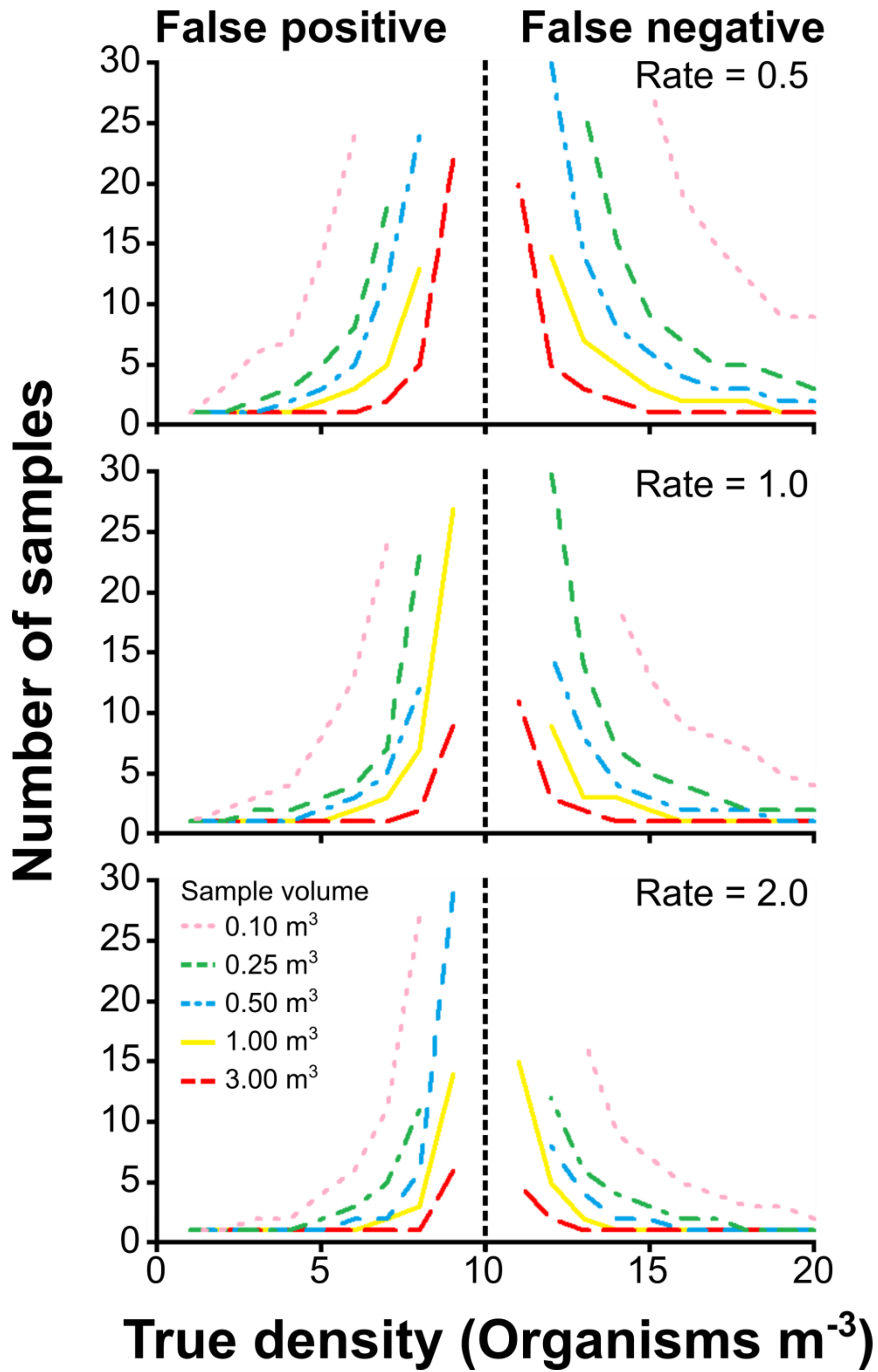
493
494
495 Figure 3.
496

497
498
499



500
501
502 Figure 4.
503

504
505
506



507
508
509 Figure 5.