



Inter-laboratory validation of the serial dilution culture—most probable number method for enumerating viable phytoplankton

Hugh L. MacIntyre¹  · John J. Cullen¹ · Shannah Rastin¹ · Magdalena Waclawik¹ · Kimberly J. Franklin¹ · Nicole Poulton² · Laura Lubelczyk² · Kate McPhee² · Tammi L. Richardson³ · Elise Van Meerssche³ · Brian Petri⁴

Received: 14 April 2018 / Revised and accepted: 4 June 2018 / Published online: 21 July 2018
© The Author(s) 2018

Abstract

The most probable number (MPN) assay is used to enumerate viable phytoplankton in the 10–50 μm size range for Type Approval of ballast water management systems. However, its validation has been questioned by the United States Coast Guard, which does not accept it as an alternative to its required method. To address key elements of method validation, three research laboratories used the same experimental protocols to enumerate viable organisms from uniformly viable and heat-killed cultures of three species of phytoplankton. Statistical analysis of the ratio of MPN estimates of viable organisms to measurements from flow cytometry was consistent with predictions based on perfect adherence to the method's assumptions. These include reliable detection of all viable organisms and no false-positive results. The validation confirmed established estimates of the MPN method's precision and demonstrated reproducibility, having been transferred to and implemented by two of the laboratories for the first time.

Keywords Most probable number · Bias · Precision · Viability · Phytoplankton · Validation

Introduction

Recognizing the threat of invasive species transported by shipping, the International Maritime Organization (IMO) adopted the International Convention for the Control and Management of Ships' Ballast Water and Sediments (BWM: IMO 2004). In force since September 2017, the BWM requires the installation of ballast water management systems (BWMS) that have been granted a Type Approval Certificate by the responsible

Administration. The type-approval process includes land-based and shipboard testing, taking into account detailed guidelines (IMO 2016a). These include, among other requirements of the Regulation D-2 Performance Standard, testing to demonstrate the discharge of fewer than 10 viable organisms per milliliter in the size category of $\geq 10 \mu\text{m}$ and $< 50 \mu\text{m}$ in minimum dimension (hereafter 10–50 μm), defining viable organisms as “organisms that have the ability to successfully generate new individuals in order to reproduce the species” (IMO 2016a). The United States Coast Guard (USCG) has its own standards for ships' ballast water discharged in US waters (USCG 2012). The numerical limit for 10–50 μm organisms is the same, but the USCG regulates the discharge of *living* instead of *viable* organisms.

As discussed extensively elsewhere (First and Drake 2013; Wright and Welschmeyer 2015, and references therein; Cullen and MacIntyre 2016), two fundamentally different methodologies are used to enumerate viable organisms (those capable of reproducing) vs. living organisms (those showing signs of life) in BWMS Type Approval testing. The serial dilution culture-most probable number method (SDC-MPN) can be used to enumerate viable 10–50 μm photosynthetic organisms in ballast water discharge (Wright and Welschmeyer 2015; Cullen and MacIntyre 2016). For Type Approval testing, SDC-MPN enumerations are augmented by microscopic

Hugh L. MacIntyre and John J. Cullen contributed equally to this work.

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10811-018-1541-z>) contains supplementary material, which is available to authorized users.

✉ Hugh L. MacIntyre
hugh.macintyre@dal.ca

¹ Department of Oceanography, Dalhousie University, PO Box 15000, Halifax, NS B3H 4R2, Canada

² Bigelow Laboratory for Ocean Sciences, 60 Bigelow Drive, P.O. Box 380, East Boothbay, ME 04544, USA

³ Department of Biological Sciences, University of South Carolina, Columbia, SC 29208, USA

⁴ Trojan Technologies, 3020 Gore Rd, London, ON N5V 4T7, Canada

counts of motile, non-photosynthetic organisms; this is the MPN Dilution Culture+Motility method (IMO 2016b). For the enumeration of living organisms, the USCG incorporated by reference a fluorescence-based live/dead test described in the Environmental Technology Verification (ETV) Protocol (USEPA 2010). This method employs the fluorescent probes fluorescein diacetate (FDA) and 5-chloromethylfluorescein diacetate (CMFDA)—commonly referred to as vital stains—to detect membrane integrity, assumed to be a characteristic of living organisms (see MacIntyre and Cullen 2016). Motile organisms that do not fluoresce are also classified as live, so the method can be called FDA/CMFDA+Motility or Stain-Motility (Cullen 2018). In part because their standards are for the discharge of living, not viable organisms, the USCG rejected a proposed MPN Dilution Culture+Motility method as an alternate to their required FDA/CMFDA+Motility method in 2016 (USCG 2016).

In its official guidance, the Marine Environment Protection Committee (MEPC) of the IMO (2017) states that of the two methodologies that may be used to enumerate viable organisms for Type Approval of BWMS—FDA/CMFDA+Motility and MPN Dilution Culture+Motility—only the SDC-MPN-based method is suitable for treatment technologies that are designed to destroy the ability to reproduce rather than to kill. Disinfection with ultraviolet radiation (UV) is one of these technologies (Cullen and MacIntyre 2016). Consequently, systems using UV-disinfection as an endpoint cannot be suitably tested by the USCG or any other administration that does not accept SDC-MPN-based enumerations of viable organisms for BWMS Type Approval. This is in contrast to US regulations for drinking water, which explicitly recognize that inactivation—i.e., destruction of the ability to reproduce, as measured with growth-based assays comparable to SDC-MPN—is the appropriate measure of efficacy for samples treated with UV (USEPA 2006). The IMO Marine Environment Protection Committee (MEPC) guidance remains open for addition of new or revised methodologies as they become available but, at time of writing, none besides MPN Dilution Culture+Motility is recognized as being applicable to all ballast water treatment technologies (IMO 2017). As recognized for both the FDA/CMFDA+Motility method (USEPA 2010) and SDC-MPN (Cullen and MacIntyre 2016), accurate enumeration of viable but dormant organisms in the 10–50 μm size-class, (e.g., dinoflagellate cysts) is challenging and may require different methods.

The USCG is working with the IMO to harmonize the IMO G8 Guidelines with the US type-approval process (USCG 2016), and this includes renewed consideration of SDC-MPN, based on additional documentation of method verification, either provided by present users of the method, or from an on-going project initiated by the ETV in the USA (see IMO 2016c). However, relevant performance metrics and validation criteria are not specified by IMO. Rather, the MEPC

guidance states that “Analytical methodologies should be validated to the satisfaction of the Administration” (IMO 2017).

Detailed and rigorous method validation guidelines exist for widely used applications such as microbiological analysis of foods or environmental testing (Feldsine et al. 2002; Parshionikar et al. 2009), but no comparably structured guidance exists for validating methods that may be used for enumerating viable organisms in the 10–50 μm size range for BWMS Type Approval. A quantitative framework for method validation has been proposed, however (Cullen 2018). The focus is on performance metrics: precision, bias, and limits of detection. Notably, a key source of uncertainty in the enumeration of viable or living organisms—false-positive and false-negative error in the classification of viable or living organisms—cannot be assessed directly in ballast water testing because there is no accepted reference for the concentration of viable or living organisms in natural waters. Although systematic observations can be used to constrain rates of false-positive and false-negative error for samples from nature (e.g., Steinberg et al. 2011; Adams et al. 2014), it is recognized that experiments on actively growing and killed laboratory cultures are required for validation of enumeration methods (see IMO 2016d; MacIntyre and Cullen 2016).

Focusing on *Tetraselmis* as a target organism for use in validating BWMS, Sun and Blatchley (2017) showed that for actively growing cultures, SDC-MPN results were statistically indistinguishable from hemocytometer counts, suggesting that SDC-MPN returned unbiased enumerations of viable organisms. In turn, estimates for cultures exposed to UV-C radiation in excess of the quantified dose-response relationships were consistently near the method’s limit of detection, well below the regulatory limit of 10 viable organisms mL^{-1} ; false-positive results were therefore insignificant. We have estimated concentrations of viable organisms by SDC-MPN on independently replicated cultures of 12 species of phytoplankton in active growth (MacIntyre and Cullen 2016; MacIntyre et al. 2018): The estimates did not differ significantly from counts of total organisms from flow cytometry. Five of these species were also subjected to killing heat treatment: SDC-MPN indicated 98.4–99.998% loss of viability (MacIntyre and Cullen 2016). Key assumptions of the SDC-MPN method have therefore been tested and shown to be valid for cultures of phytoplankton. To date, comparable laboratory studies on the FDA/CMFDA+Motility method have not been published.

Besides demonstrating that a method is suitable for its intended use, it is important to document the reproducibility of results among laboratories as part of method validation (Mishalanie et al. 2016). As these authors note, “A method that proved rugged for use in one laboratory may lose that characteristic when tried in several other laboratories”: Transferability is therefore an important component of validation. For direct inter-laboratory comparisons, this requires

testing the same material—an impossibility for natural samples of plankton, but achievable with the use of phytoplankton cultures grown under carefully controlled conditions (MacIntyre and Cullen 2005). Studies on laboratory cultures can thus contribute significantly to the validation of the SDC-MPN base method in support of its evaluation for use in BWMS Type Approval testing.

The objectives of this study were the following:

1. Determine if SDC-MPN can classify actively growing phytoplankton as viable, and heat-killed phytoplankton as not, without bias.
2. Characterize the precision of SDC-MPN results and compare it to the theoretical expectation for perfect adherence to method assumptions.
3. Test for reproducibility of SDC-MPN by comparing results from three laboratories—two of which are new to the method.

This account includes a description of methodological best practices and a comparison of SDC-MPN calculations, including uncertainty estimates and limits of detection. The scope of this study is restricted to validation of the SDC-MPN methodology when used on phytoplankton: These first results would complement the formidable body of research on MPN as applied to other target organisms using distinctly different culturing conditions, e.g., with coliform or marine bacteria (Haas and Heller 1988; Button et al. 1993) or the pathogenic protist *Cryptosporidium parvum* (Slifko et al. 1999).

Background

The SDC-MPN method

Rooted in methods for enumerating bacteria in sanitary analyses of liquids (McCrary 1915; Cochran 1950), the SDC-MPN method is an established approach for use with multi-species communities of phytoplankton (Thronsen 1978; Furuya and Marumo 1983; Thronsen and Kristiansen 1991; Andersen and Thronsen 2003; Cerino and Zingone 2006) that has more recently been used in the testing of ballast water discharge (Madsen and Petersen 2015; Wright and Welschmeyer 2015; Cullen and MacIntyre 2016) and for quantifying the efficacy of UV radiation as a potential ballast water treatment technology (Sun and Blatchley 2017; MacIntyre et al. 2018). The length of the assay (typically weeks even under optimized conditions, MacIntyre et al. 2018) effectively precludes its use for monitoring installed BWMS for regulatory compliance. Its practical application is in Type-Approval testing.

The principles of the SDC-MPN method (Cochran 1950; Jarvis et al. 2010) are the same, whether for bacteria or

phytoplankton. A sample is dispensed into a series of replicate culture tubes in tiers of dilution such that some tubes contain no viable organisms and thus can show no growth. After a suitable incubation (see “Materials and methods,” “Growth assays”), tubes are scored positive for the presence of one or more viable organisms, as demonstrated by detectable growth; otherwise, they are scored negative. The most probable number of viable organisms in the undiluted parent sample is calculated from the number of tubes scoring positive in each tier of the dilutions. The method assumes the following:

1. Organisms are randomly distributed in each tube and evenly distributed between subsamples.
2. Growth will be reliably detected in any tube containing one or more viable organisms.

The first assumption allows the calculation of MPN and its uncertainties based on the expectation of organisms-per-culture-tube from the Poisson distribution (Jarvis et al. 2010). If the second assumption is not satisfied, SDC-MPN will underestimate the concentration of viable organisms in direct proportion to the relative abundance of viable organisms that do not grow to detection under SDC-MPN culture conditions (Cullen 2018).

Metrics of method performance

Accuracy is the closeness of agreement between a test result and the accepted reference value (ISO 1994). It is determined by a combination of random error affecting precision and systematic error causing bias. Precision refers to the closeness of agreement between independent test results obtained under stipulated conditions. It is usually expressed in terms of imprecision, computed, for example, as the standard error of test results. Trueness is defined as the closeness of agreement between the suitably averaged value obtained from a large series of test results and the accepted reference. It is usually expressed in terms of bias. For SDC-MPN, the test result is C_{MPN} (organisms mL^{-1}) and the reference value would be the true concentration of viable organisms, μ_{viable} (organisms mL^{-1}). Here, μ_{viable} is estimated by measurements of the concentration of actively growing organisms (C_{viable} , organisms mL^{-1}), and bias is represented by the agreement between multiple measurements of C_{MPN} and C_{viable} (see “Background,” “Factor of agreement”).

Precision and confidence intervals

The SDC-MPN calculation is a maximum-likelihood estimator of $\ln(\mu_{viable})$. The estimates, $\ln(C_{MPN})$, follow an approximately normal distribution described by its standard error, $\sigma_{\ln(C_{MPN})}$. The 95% confidence interval on an estimate is

$\ln(C_{\text{MPN}}) - 2\sigma_{\ln(C_{\text{MPN}})}$, $\ln(C_{\text{MPN}}) + 2\sigma_{\ln(C_{\text{MPN}})}$ (Cochran 1950; Jarvis et al. 2010). Correspondingly, the lower and upper 95% confidence limits on an SDC-MPN estimate, LL_{MPN} and UL_{MPN} , respectively, are

$$LL_{\text{MPN}} = C_{\text{MPN}}/\exp(2\sigma_{\ln(C_{\text{MPN}})}), UL_{\text{MPN}} = C_{\text{MPN}} \cdot \exp(2\sigma_{\ln(C_{\text{MPN}})}) \quad (1)$$

As Cochran (1950) explained, the multiplicative factor for the confidence interval, which is represented as $FCI = \exp(2\sigma_{\ln(C_{\text{MPN}})})$ (Cullen 2018), is an appropriate substitute for the standard error of an estimate.

Factor of agreement

Systematic error in the SDC-MPN estimate, for example from failure to detect viable organisms (false-negative results), leads to bias that can be detected in the distribution of the factor of agreement (FOA), $C_{\text{MPN}}/C_{\text{viable}}$. False-negative results lead to $\text{FOA} < 1.0$ and false-positive results lead to $\text{FOA} > 1.0$. For any μ_{viable} , estimates of C_{MPN} follow a log-normal distribution with a variance that greatly exceeds that for C_{viable} (see “Results,” “Precision”). Consequently, the distribution of $\ln(\text{FOA})$, i.e., $\ln(C_{\text{MPN}}/C_{\text{viable}})$, is suitable for testing the hypothesis that SDC-MPN results are biased, i.e., $\ln(\text{FOA}) \neq 0$ (Cullen 2018).

Limit of detection

The upper confidence limit for a uniformly negative SDC-MPN result can also be considered the lower limit of detection (LLD_{MPN}) for that analytical configuration (Cullen 2018). (The analytical configuration refers to the tube volume, number and ratio of tiers of dilution, dilution in the upper tier of tubes, and the number of tubes per tier.) Jarvis et al. (2010) calculate the limit using a two-tailed, 0.025 probability for the highest concentration of target organisms that could return zero detects—no tubes in any dilution testing positive for growth—reporting the SDC-MPN estimate as zero with the upper confidence limit of UL_{MPN} . In contrast, the MPN table from the US Food and Drug Administration, USFDA, reports the MPN for a uniformly-negative result as less than the lowest MPN for an outcome with at least one positive tube (Blodgett 2005, 2010); the value is not called a limit of detection. Because their treatment has an explicitly described mathematical foundation, SDC-MPN estimates for uniformly negative results will be reported following Jarvis et al. (2010), and the upper bound will also be considered the LLD_{MPN} .

Materials and methods

Organization and communication

The experiments were performed at Dalhousie University (DAL), where the base methodology was developed (see MacIntyre et al. 2018), the Bigelow Laboratory for Ocean Sciences (BLOS), and the University of South Carolina (U-SC). Standard operating procedures for maintaining and harvesting cultures and assessing viability by SDC-MPN were developed in a workshop held at DAL prior to the experimental phase. The research laboratories at BLOS and U-SC were experienced in the culture of phytoplankton for experimentation, but new to SDC-MPN. Communications were archived on an electronic bulletin board, and all data were loaded to a common archive.

Culture maintenance

Cultures of the prasinophyte *Pyramimonas parkeae* Norris et Pearson (Strain CCMP725), the diatom *Thalassiosira weissflogii* (Grun.) Fryxell et Hasle (Strain CCMP1050), and the dinoflagellate *Amphidinium carterae* Hulbert (Strain CCMP1314) were obtained from the National Center for Marine Algae and Microbiota (NCMA; East Boothbay, ME, USA). Each species is unicellular, so organisms and cells are assumed to be equivalent measures. The cultures were grown in each laboratory on *f/2* + Si growth medium (Guillard 1973), using coastal seawater as a base. The medium was sterilized by filtration through a 0.2- μm Whatman Polycap AS filter cartridge (BLOS) or a 0.2- μm Pall Gelman Culture Capsule filter (DAL and U-SC). All glassware and tubing with which the cultures made contact were cleaned by soaking in 2–5% HCl, followed by copious rinsing with E-Pure water (Barnstead Nanopure/APS Water Services Corporation, Lake Balboa, CA, USA) or an equivalent, and were sterilized by autoclaving. All transfers were done in a laminar flow hood with 0.2- μm filtered air. All surfaces were wiped with 95% ethanol prior to transferring cultures, and the tubes were flamed on opening and closing to minimize the risk of contamination.

The cultures were maintained in semi-continuous growth (MacIntyre and Cullen 2005) in 50-mL Pyrex tubes closed with either unlined HDPE caps or PTFE-lined phenolic caps, on a 12:12 light/dark cycle at 20 ± 0.5 °C and at a growth irradiance of 100 $\mu\text{mol photons m}^{-2} \text{ s}^{-1}$ PAR supplied by fluorescent bulbs (General Electric Ecolux F34CW-RS-WM-ECO at BLOS; Osram FL40SS-W/37 at DAL; Sylvania Octron/Eco 32 W at U-SC), as monitored by thermistors and light meters in each laboratory.

To monitor growth of the cultures noninvasively, i.e., without compromising the sterility of the culture (Brand et al. 1981), dark-acclimated chlorophyll fluorescence (F , arbitrary

fluorometer units) was measured daily, 3–5 h into the dark period, by inserting the culture tube directly into the sample compartments of a Turner 10-AU fluorometer (Turner Designs, USA).

The cultures were diluted every one to five generations with fresh medium to maintain them in exponential growth at low optical density. The specific growth rate, μ_F (day^{-1}), was calculated from the between-day increase in dark-acclimated fluorescence (Parkhill et al. 2001; Wood et al. 2005):

$$\mu_F = \frac{1}{t_2 - t_1} \cdot \ln \left(\frac{F_2}{[1-d] \cdot F_1} \right) \quad (2)$$

where t_1 and t_2 are the times (day) of successive observations, F_1 and F_2 (Arb.) are the corresponding blank-corrected fluorescence obtained from the Turner 10-AU, and d (dimensionless) is the dilution (proportion of culture replaced with fresh medium) between readings.

Physiological status of the cultures was monitored by measuring variable chlorophyll *a* fluorescence using a FASTOcean Sensor FRRF (Chelsea Technologies Group, UK) at BLOS, a benchtop FIRE (Fluorescence Induction and Relaxation system, Satlantic, Canada) at DAL, and the Turner 10-AU with/without 3-(3,4-dichlorophenyl)-1,1-dimethylurea (DCMU) at a final concentration of 40 $\mu\text{mol L}^{-1}$ (Parkhill et al. 2001) at U-SC. Estimates were made non-destructively by inserting the culture tube into a custom-built cuvette holder (DAL) or by destructive analysis of sub-samples removed from the culture tubes (BLOS and U-SC). Blanks (culture medium) and standards were measured daily at DAL and U-SC. The FIRE was standardized with a 200 $\mu\text{mol L}^{-1}$ solution of rhodamine *b* (R6625, Sigma-Aldrich) in E-Pure water. The FASTOcean Sensor FRRF (BLOS) is calibrated at the manufacturer annually; No daily standard was used.

The maximum quantum efficiency of Photosystem II transport, F_v/F_m , based on multiple-turnover flashes, was calculated using the manufacturers' software, FastPro8, for data collected on the FRRF (BLOS) and as $[F_{\text{DCMU}} - F]/F_{\text{DCMU}}$, where F and F_{DCMU} are blank-corrected fluorescence measured before and after addition of DCMU (U-SC). For the FIRE (DAL), the minimum, maximum, and variable fluorescence (F_0 , F_m , F_v) and the dimensionless ratio F_v/F_m were obtained by non-linear fitting of the single-turnover fluorescence induction curve using Fireworx software (Audrey Ciochetto, née Barnett, <http://sourceforge.net/projects/fireworx/>). Curve fits were performed with the curvature parameter in the fit, p , set to zero.

Harvesting and heat treatment

The cultures were assumed to be in balanced exponential growth and ready for experimental use when the coefficients

of variation in μ_F and F_v/F_m were less than 10% over 10 generations (MacIntyre et al. 2018). Cultures were harvested 4–6 h after the start of the dark period. Each culture was divided into clean 7-mL glass tubes prior to setting up the growth assays, and the subsamples were either left untreated or immersed in a water bath at 50 °C for 10 min. The heat-treated organisms are assumed to be dead (MacIntyre and Cullen 2016). A sensitivity analysis supporting the assumption is presented in the [Supplementary Materials](#), “The 10-generation criterion for balanced growth.”

Growth assays

The concentrations of organisms in the untreated samples were measured by flow cytometry using a FACScan (Becton Dickinson, USA) at BLOS, an Accuri C6 (Becton Dickinson, Canada) at DAL, and Guava easyCyte (Millipore Sigma, USA) at U-SC. Optical alignment in the FACScan (BLOS) was validated weekly using 3 μm Rainbow Calibration Particles (Spherotech, Inc., Lake Forest, IL, USA). The Accuri C6 and Guava easyCyte were calibrated with validation beads prior to each use (Becton Dickinson Spherotech 8-peak at DAL; GuavaCheck Bead Reagent at 51,000 beads mL^{-1} at U-SC). Organism concentrations were estimated by counting events in a clearly defined cluster in a log-log bi-plot of chlorophyll *a* fluorescence (Ex. 488 nm; Em. > 670 nm) vs. side scatter after gating to exclude the background levels observed in distilled water or filtered seawater. For each sample, 150–300 μL of media were analyzed, counting c. 1000–15,000 organisms. At BLOS, sample volume was determined gravimetrically using a microbalance (Denver Instrument, USA). At DAL and U-SC, the instrument settings were assumed to deliver accurate sample volumes. Due to the mathematical certainty that non-reproductive organisms are eliminated during successive culture dilutions (see [Supplementary Materials](#), “The 10-generation criterion for balanced growth”), organisms in these actively growing cultures are assumed to be uniformly viable and their measured concentrations are represented as C_{viable} (organisms mL^{-1}).

The concentration of viable organisms in both untreated and heat-killed cultures was estimated using the SDC-MPN growth assay (McCrary 1915; Cochran 1950; Thronsen 1978), with growth assessed from changes in dark-adapted chlorophyll *a* fluorescence using “Best Practice” criteria described by MacIntyre et al. (2018). The cultures were diluted in 5-mL volumes in sterile Pyrex tubes (five per dilution) in three tiers separated by an order of magnitude (10^{-4} , 10^{-5} , 10^{-6} for untreated cultures; 10^0 , 10^{-1} , 10^{-2} for heat-killed cultures). The dilution ranges for untreated cultures were based on predicted concentrations of viable organisms (MacIntyre et al. 2018), with the starting assumption that all organisms in the untreated culture were viable. Use of undiluted culture for the

top tier of the assay with heat-killed cultures gives the maximum sensitivity in detecting viable cells.

The dilution-series tubes were incubated under the same conditions of irradiance, temperature, and day length as the preliminary cultures. Chlorophyll fluorescence (F) was measured with the Turner 10-AU fluorometers every 48 h. Tubes were scored as positive if fluorescence increased $8\times$ above the fluorescence corresponding to the lower limit of detection (LLD_F) of the fluorometer or $8\times$ above the initial fluorescence, whichever was higher (MacIntyre et al. 2018). The lower limit of detection was determined for each species on each fluorometer using a dilution series of the exponentially growing culture (Anderson 1989; Miller and Miller 2005, see Supplementary Materials).

To minimize the risk of false negatives—tubes that should exhibit growth but in which growth was not observed because they were not monitored for long enough—tubes were only scored as negative for growth if no increase in fluorescence had been observed at t_{end} , an interval based on the growth responses of the least-diluted tier of the exponentially growing

culture. This is illustrated in Fig. 1. The duration (day) of t_{end} is defined (Eq. 6a in MacIntyre et al. 2018) as

$$t_{end} = \left(3 \cdot \ln(2) - \ln \left(\frac{10^{X-x} \cdot F_{init}}{LLD_F} \right) \right) \cdot \mu_F^{-1} \quad (3)$$

where F_{init} (Arb.) and μ_F (day) are the intercept and slope, respectively, of a linear regression of \ln -transformed fluorescence on time for tubes exhibiting growth in the least-diluted tier of the exponentially growing culture (Fig. 1a); and $x = X$, $X-1$, or $X-2$ for the dilution series 10^X , 10^{X-1} , and 10^{X-2} . The underlying assumption is that the time for fluorescence to reach LLD_F in a tube exhibiting growth can be predicted by assuming that the initial fluorescence, F_{init} , varies with dilution and that the growth rate (μ_F) is constant across dilutions and treatments. The first term in the expression in Eq. 3 is an additional three-generation buffer added to the estimate to account for any lag in growth (MacIntyre et al. 2018).

The lower limit of the initial fluorescence, $F_{init,min}$, corresponds to the signal from a single organism in the tube. This is calculated as the ratio of F_{init} to the mean initial number of

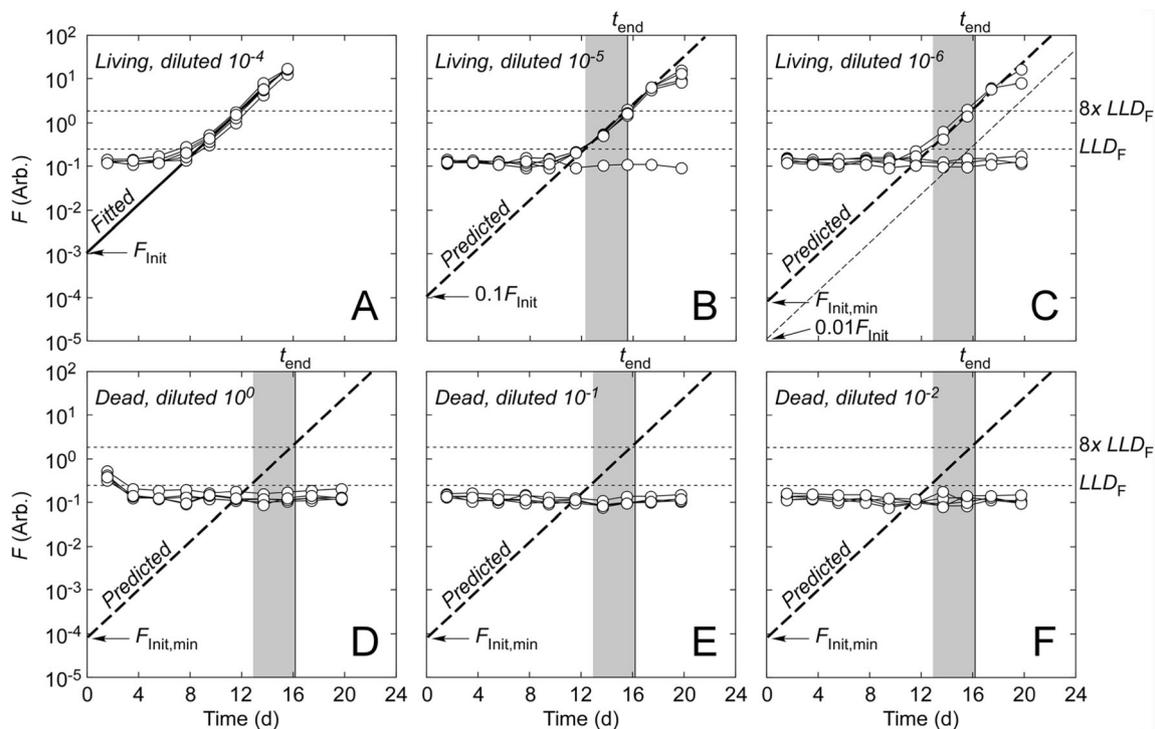


Fig. 1 Time courses of in vivo fluorescence in three serial dilutions of an exponentially growing (a–c) and heat-killed (d–f) treatments of a culture of *Amphidinium carterae*. There were five replicates per dilution. Note the log scales. Tubes in which fluorescence increases above $8\times$ the lower limit of detection (LLD_F) are scored as positive for growth. **a** The dashed line (“Fitted”) is an extrapolation of the exponential growth phase, determined by regression of \ln -transformed fluorescence on time for points above LLD_F (see text for details). The intercept is F_{init} and the slope is μ_F . **b–f** The dashed lines (“predicted”) are the expected time courses of fluorescence for tubes that are sequential dilutions of (a) or are sequential

dilutions (e, f) of a separate heat-killed subsample (d). The prediction assumes that the viable fraction of the culture, determined by the ratio of dilution relative to (a), would grow at the same rate. The lower limit for the prediction (c–f) is calculated for the case where a tube contains a single cell (see text for details). The shaded bars indicate a three-generation period beyond the time at which fluorescence is predicted to exceed LLD_F . Tubes without an increase in fluorescence at the end of this period (t_{end} , right margin of the gray bar) are scored as negatives. The final MPN score in the exponentially growing sample is 5–4–2, and the score in the heat-killed sample is 0–0–0

viable organisms in the tubes from which it was determined. In practice, $F_{init,min}$ was determined from the intercept of a regression for the least-diluted tier of tubes from the untreated culture (Fig. 1a) normalized to the total cell number in the tubes. Total cell number was calculated as the product of C_{viable} , the volume of the tubes, and their dilution.

$$F_{init,min} = \frac{F_{init}}{C_{variable} \cdot V \cdot 10^X} \quad (4)$$

where C_{viable} (organisms mL^{-1}) is the concentration of organisms in the parent culture, V is the volume of culture in the assay (here, 5 mL), and 10^X (dimensionless) is the dilution in the tubes from which F_{init} was estimated (here, 10^{-4}).

The starting value of fluorescence for the prediction of growth was based either on the product of F_{init} and the dilution ratio between the tier being observed and the tier on which F_{init} was estimated (e.g., Fig. 1b) or on $F_{init,min}$, whichever was higher (e.g., Fig. 1c). Predicted growth for all dilutions of the heat-treated cultures was based on $F_{init,min}$ (Fig. 1d–f), as it defines the limit of t_{end} .

Calculations

At t_{end} , results for the SDC-MPN assays were scored to indicate the number of positive scores out of five replicates in each of the three tiers of the assay (e.g., 5–4–2, Fig. 1a–c). Most Probable Number estimates of the concentration of viable organisms (C_{MPN} , organisms mL^{-1}), the standard deviation of $\log_{10}(C_{MPN})$, the 95% confidence intervals on C_{MPN} , and rarity values—all from Jarvis et al. (2010)—were generated using a spreadsheet (Wilrich 2017). Results were compared with those from a USDA table (Blodgett 2010), calculated as described by MacIntyre et al. (2018). Comparisons with the

results from other calculators are presented in Table S1 (Supplementary Materials).

Confidence limits from each calculation method were converted to the standard error of $\ln(C_{MPN})$ using the relationship, $\sigma_{\ln(C_{MPN})} = \ln(UL_{MPN}/C_{MPN}) / 2$ (see Eq. 1). The factor of agreement for each MPN determination is C_{MPN}/C_{viable} . Due to equipment failure, the three measurements of C_{viable} for *Pyramimonas parkeae* could not be obtained at U-SC, so they are excluded from the comparisons of C_{MPN} with C_{viable} .

Results

Enumeration of viable and heat-killed organisms

The concentration of organisms obtained from flow cytometry, C_{viable} , was within the 95% confidence intervals for the SDC-MPN estimate, C_{MPN} , for all tests on actively growing cultures from each of the three laboratories (Fig. 2, Table 1). This is consistent with the requirement for a regulatory test that SDC-MPN reliably enumerates viable organisms. More categorically, in every case, no growth was detected in tubes of the heat-treated samples (e.g., Fig. 1d–f): For scores of 0–0–0 (dilutions of 1, 10^{-1} , and 10^{-2}), $C_{MPN} = 0$, and $UL_{MPN} = 0.13$ organisms mL^{-1} . These results correspond to relative viabilities (cf. MacIntyre et al. 2018) of 2.9×10^{-6} to 2.1×10^{-5} , or less, at the cell concentrations of the cultures (Table 1).

MPN estimates from different calculators

The 27 assays of untreated cultures (Table 1) yielded 11 unique estimates of C_{MPN} (Table S1): For 9, the estimated MPNs (Jarvis et al. 2010) were identical to calculations from the USDA table (Blodgett 2010), rounded to two

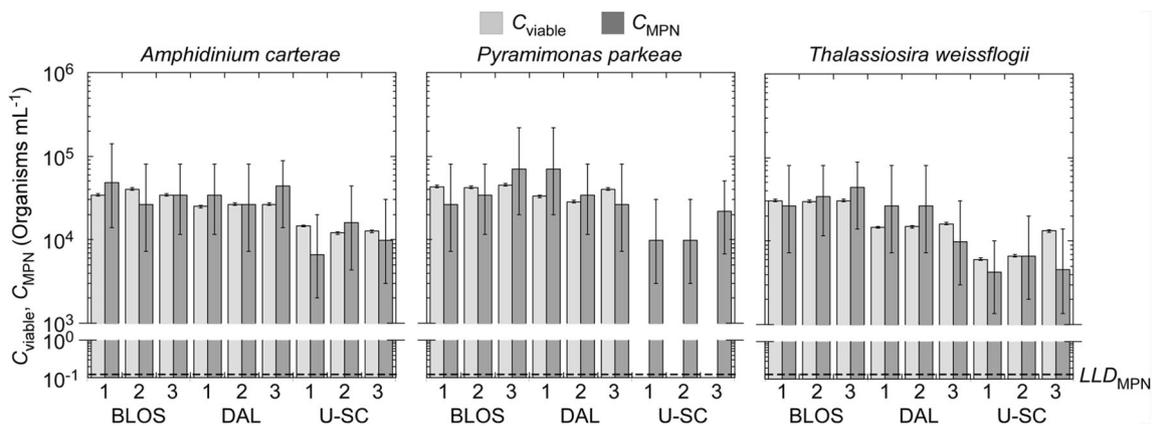


Fig. 2 Estimates of the concentrations of viable organisms by flow cytometry (C_{viable}) and by SDC-MPN (C_{MPN}). Independent replicates (1–3) were assayed for three species: the dinoflagellate *Amphidinium carterae*, the prasinophyte *Pyramimonas parkeae*, and the diatom *Thalassiosira weissflogii*. Assessments were made at Bigelow

Laboratory for Ocean Sciences (BLOS), Dalhousie University (DAL), and the University of South Carolina (U-SC). The dashed line (LLD_{MPN}) is the limit of detection for this configuration of the test, 0.13 organisms mL^{-1} (see text for details). Error bars are 95% confidence intervals

Table 1 Results of SDC-MPN for cultures of three species of unicellular phytoplankton maintained in exponential growth, ensuring that essentially all organisms measured with flow cytometry (C_{viable} , organisms mL^{-1}) are viable. The MPN estimates of viable organisms (C_{MPN} , organisms mL^{-1}) and their 95% confidence limits (LL_{MPN} , UL_{MPN}) are from Jarvis et al. (2010). Scores are the number of positive

tubes from five replicates each at dilutions of 10^{-4} , 10^{-5} , and 10^{-6} , respectively. The standard error of the logarithm of the estimate, $\sigma_{\ln(C_{\text{MPN}})} = \ln(UL_{\text{MPN}}/C_{\text{MPN}}) / 2$. The factor of agreement, $\text{FOA} = C_{\text{MPN}}/C_{\text{viable}}$. Corresponding SDC-MPN estimates for heat-killed cultures were all the same: scores = 0–0–0, $C_{\text{MPN}} = 0$ and $UL_{\text{MPN}} = 0.13$ organisms mL^{-1}

Species	Laboratory	C_{viable} (organisms $\text{mL}^{-1} \times 10^{-3}$)	C_{MPN} (95% CI) (organisms $\text{mL}^{-1} \times 10^{-3}$)	Scores	$\sigma_{\ln(C_{\text{MPN}})}$	FOA
<i>Amphidinium carterae</i>	BLOS	34.1	48 (16–150)	(5–5–0)	0.570	1.41
		40.1	26 (8.9–76)	(5–4–0)	0.536	0.65
		34.3	22 (7.7–61)	(5–3–1)	0.510	0.64
	DAL	24.9	34 (13–92)	(5–4–1)	0.498	1.37
		26.4	26 (8.9–76)	(5–4–0)	0.536	0.98
		26.4	44 (18–110)	(5–4–2)	0.458	1.66
	U-SC	14.7	6.6 (2.1–20)	(5–1–0)	0.554	0.45
		12.1	16 (5–51)	(5–3–0)	0.580	1.33
		12.7	9.9 (3–32)	(5–2–0)	0.587	0.78
<i>Pyramimonas parkeae</i>	BLOS	43.0	26 (8.9–76)	(5–4–0)	0.536	0.60
		42.3	34 (13–92)	(5–4–1)	0.498	0.80
		45.1	70 (22–220)	(5–5–1)	0.573	1.55
	DAL	33.0	70 (22–220)	(5–5–1)	0.573	2.12
		28.3	34 (13–92)	(5–4–1)	0.498	1.20
		40.2	26 (8.9–76)	(5–4–0)	0.536	0.65
	U-SC		9.9 (3–32)	(5–2–0)	0.587	
			9.9 (3–32)	(5–2–0)	0.587	
			22 (7.7–61)	(5–3–1)	0.510	
<i>Thalassiosira weissflogii</i>	BLOS	30.5	26 (8.9–76)	(5–4–0)	0.536	0.85
		29.9	34 (13–92)	(5–4–1)	0.498	1.14
		30.4	44 (18–110)	(5–4–2)	0.458	1.45
	DAL	14.7	26 (8.9–76)	(5–4–0)	0.536	1.77
		15.0	26 (8.9–76)	(5–4–0)	0.536	1.74
		16.2	9.9 (3–32)	(5–2–0)	0.587	0.61
	U-SC	6.05	4.3 (1.7–11)	(4–2–0)	0.470	0.71
		6.66	6.6 (2.1–20)	(5–1–0)	0.554	0.99
		13.23	4.6 (1.5–14)	(5–0–0)	0.557	0.35

significant digits for comparison. In the two remaining instances, the differences were small (-1.0 and $+2.3\%$, see Supplementary Materials, Table S1) and within rounding error. Further comparisons using an additional three calculators revealed only two more discrepancies, of $+4$ and -1.4% (see Supplementary Materials, Table S1). These deviations are minuscule compared to the confidence intervals of the estimates (cf. Table 1).

Estimates of uncertainty in the MPN calculations

The mean of 27 estimates of $\sigma_{\ln(C_{\text{MPN}})}$ in Table 1 is 0.535 (median 0.536; first and third quartiles 0.498 and 0.571). This corresponds to a multiplicative 95% confidence interval, $\text{FCI} = 2.9$. Using the USFDA table, the mean is

0.507 (median 0.557; first and third quartiles 0.428 and 0.562), which is not significantly lower according to a pooled t test (t -statistic = 1.698, $df = 52$, $p = 0.094$). The calculated $\text{FCI} = 2.76$ is 6% smaller than the Jarvis et al. (2010) estimate. Among four other calculators, deviations, represented as $(\text{FCI}_{\text{method } i} - \text{FCI}_{\text{Jarvis}}) / \text{FCI}_{\text{Jarvis}}$, ranged from -26 to $+27\%$ for individual estimates. The median differences ranged from 0 to -6.7% , (see Supplementary Materials, Table S1).

Differences between uncertainty estimates have been examined extensively—see Jarvis et al. (2010) as a point of entry to the literature. We report the Jarvis et al. (2010) confidence intervals in Table 1, but note that they have no bearing on our analyses of method precision and bias in the inter-laboratory comparison (see “Results,” “Precision” and “Bias”). In turn,

only the enumerations of viable organisms, not their confidence intervals, are used in BWMS Type Approval testing (see Cullen 2018).

Limits of detection

For the sample configuration used here to assay heat-killed cultures, $LLD_{MPN} = 0.13$ organisms mL^{-1} , the upper confidence limit for a negative result, based on a two-tailed, 0.025 probability (Jarvis et al. 2010). Applying a one-tailed test ($p = 0.05$), the LLD is 0.11 organisms mL^{-1} (Cullen 2018). The USFDA table (Blodgett 2010) reports the lowest MPN for an outcome with at least one positive tube (0–0–1); the result, $C_{MPN} < 0.36$ organisms mL^{-1} , is not called a limit of detection. For this study, the mathematically defined UL_{MPN} of 0.13 organisms mL^{-1} from Jarvis et al. (2010) represents the LLD for the SDC-MPN tests on heat-killed cultures. This corresponds to < 1 viable cell per 5-mL tube, confirming that the method should detect any survivor with a probability of $> 97.5\%$.

The dilution ranges for untreated cultures were based on predicted concentrations of viable organisms (MacIntyre et al. 2018) so that the upper limit of detection, a score of 5–5–5, was not reached. If growth had been detected in all tubes, as can happen with sampling of natural phytoplankton communities (see Cullen 2018), Jarvis et al. (2010) would report the result as $C_{MPN} = \infty$ and $LL_{MPN} = 130,000$ viable organisms mL^{-1} for this test configuration. For the same result, the USFDA table (Blodgett 2010) returns $C_{MPN} > 320,000$ (the result for 5–5–4) and $LL_{MPN} = 140,000$ viable organisms mL^{-1} . These lower confidence limits for a result of all positive tubes are useful even though the test is saturated because in BWMS Type Approval, testing is only considered valid if densities of viable organisms in uptake water and control discharge water exceed minimum standards. Consequently, a result of LL_{MPN} exceeding the minimum confirms compliance with the requirement at $p = 0.975$.

Precision

If the base assumptions of the method are satisfied (see Introduction), the precision of the MPN estimate is primarily a function of number of tubes per dilution tier and the dilution ratio, described by functions with a range of complexity (Cochran 1950; Hurley and Roscoe 1983; MacIntyre et al. 2018). For the configuration used in this study—five tubes per tier with a $10\times$ dilution ratio—the estimated $\sigma_{\ln(C_{MPN})}$ is 0.55, from a generalized equation simplifying the results of Hurley and Roscoe (1983), as presented by MacIntyre et al. (2018). The result differs little from the mean error in Table 1 (0.535) or Cochran's (1950) estimate of 0.57.

Results from this laboratory validation were used to test the hypothesis that the ideal, theoretical precision of MPN was degraded in practice by the introduction of additional random measurement error during the tests. Following Cullen (2018), the metric for comparison is the logarithm of FOA, i.e., $\ln(C_{MPN}/C_{viable})$. Independent of measurement bias that would lead to deviations from the expected mean of $\ln(FOA) = 0$, the standard error of $\ln(FOA)$ reflects the contributions of random error in each measurement. By propagation of uncertainty, for $\ln(FOA) = \ln(C_{MPN}) - \ln(C_{viable})$:

$$\sigma_{\ln(FOA)} = \sqrt{\sigma_{\ln(C_{MPN})}^2 + \sigma_{\ln(C_{viable})}^2} \quad (5)$$

In this study, the expected $\sigma_{\ln(C_{MPN})}$ is taken as 0.55. This corresponds to a multiplicative factor for 95% confidence intervals of $FCI = 3.0$. By propagation of error (Bevington 1969), the standard error of $\ln(C_{viable})$ is estimated as the coefficient of variation (CV) of C_{viable} , which we assume to be 3% based on triplicate assessments (i.e., analytical replicates) of samples of *Thalassiosira weissflogii*. The precision of MPN determinations in this study is thus predicted to be $\sqrt{0.55^2 + 0.03^2} = 0.55$. The observed precision would be worse ($\sigma_{\ln(FOA)} > 0.55$) if experimental procedures introduced additional random measurement error.

The eight sets of triplicate determinations of FOA in Table 1 provided unbiased estimates (s_{n-1} , $n = 3$) of the standard error of the estimates of $\ln(C_{MPN}/C_{viable})$. The mean of the standard deviations was 0.47 (range = 0.26–0.61, $n = 8$), and the Shapiro-Wilk statistic for normality was 0.86 ($p = 0.12$, from `scipy.stats.shapiro` in Python): Therefore, the null hypothesis that the observations came from a normal distribution could not be rejected. With caution, we calculated a mean s.d. of $\ln(FOA)$ of 0.47, with a 95% confidence interval of 0.38–0.56 (`scipy.stats.bayes_mvs`). The upper confidence limit corresponds to $FCI = 3.06$. This result indicates that implementation of SDC-MPN in the inter-laboratory validation introduced little or no additional measurement uncertainty beyond the expected $FCI = 3.0$ inherent to the method as applied.

Although *Pyramimonas* was tested at U-SC, breakdown of the flow cytometer prevented calculation of relative viability. Results for the number of viable organisms in exponentially growing cultures were well within the range of those for the other laboratories. As with all other tests on heat-killed organisms, the U-SC scores for *Pyramimonas* were uniformly 0–0–0.

Bias

The distribution of the logarithm of the FOA (Fig. 3) represents both random measurement error and systematic error (bias) in the estimation of the concentration of viable organisms (Cullen 2018). For the 24 paired determinations of C_{MPN}

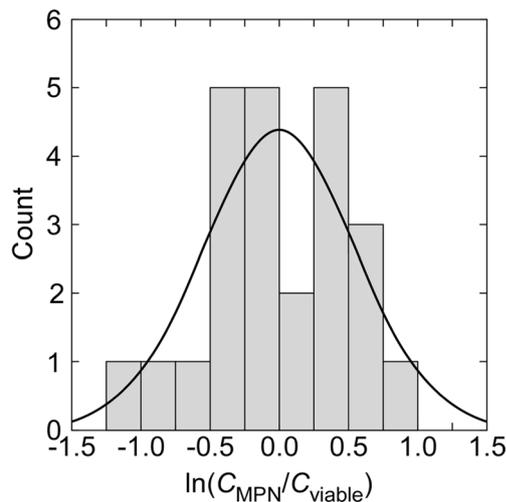


Fig. 3 Factor of agreement (log-transformed) between MPN estimates of the concentration of viable organisms (C_{MPN}) and the concentration of viable organisms measured with flow cytometry on exponentially growing cultures (C_{viable}). The mean is -0.030 (s.e., 0.097 , $n = 24$), not significantly different from 0, the expectation for unbiased estimation of viable organisms with MPN. The s.d. is 0.47 . The expectation for an ideal implementation of SDC-MPN in this configuration is mean = 0.0 with an s.d. of 0.55 (see “Results,” “Precision”), shown with the solid line. The line was scaled by binning a Gaussian curve with s.d. = 0.55 , centered on zero, in intervals of 0.25 (the bin size in the histogram) between -2.5 and 2.5 . A weighting factor was then calculated as 24 (the number of observations) normalized to the sum of bin values for the Gaussian curve, to weight the curve appropriately for the sample size

and C_{viable} from the validation, the mean $\ln(C_{MPN}/C_{viable})$ was -0.030 , not significantly different from 0. A two-way analysis of variance showed no significant influence of laboratory or species at $p = 0.05$, indicating that the pooled result was not compromised by counteracting biases (see [Supplementary Materials](#), “Test for bias in enumeration of viable organisms”). As expected if there were no significant influences of species or laboratory on the mean $\ln(\text{FOA})$, the standard deviation for the 24 determinations of $\ln(C_{MPN}/C_{viable})$ combined was 0.47 , the same as the estimate for method precision alone (see “Results,” “Precision”). This is not significantly different from the theoretical expectation of $\sigma_{\ln(\text{FOA})} = 0.55$ (chi-square test for variances statistic = 17.13 with 23 df , $p > 0.05$). These analyses provide no basis to reject the hypothesis that MPN determinations were unbiased estimators of the concentration of viable organisms and that method precision was equivalent to the theoretical ideal.

Sensitivity to the detection threshold

The criterion for scoring positive growth was set as an increase in fluorescence $8\times$ above the lower limit of detection, LLD_F (Fig. 1). This is equivalent to three generations for cultures in balanced growth (i.e., where the relationship between F and organism concentration is constant). Where there is no reason to suspect that the relationship varies, lower thresholds

for scoring positive growth are appropriate. The SDC-MPN assays were re-scored using thresholds that corresponded to $2\times$ and $4\times$ LLD_F . In all cases, there was no change in the scoring (not shown).

Discussion

In its guidance on methodologies that may be used for enumerating viable organisms for Type Approval of ballast water management systems, the Marine Environment Protection Committee of the IMO states that the methodologies “should be validated to the satisfaction of the [type-approval granting] Administration” (IMO 2017). But, criteria for validating enumeration methodologies are not well established, and gaps can be identified in published documentation of the efficacy of both the FDA/CMFDA+Motility and MPN Dilution Culture+Motility methodologies (Cullen 2018). Mindful that the USA has questioned the validation of SDC-MPN, the underpinning of the MPN Dilution Culture+Motility methodology, and recognizing that additional scientific evidence is required for any Administration to evaluate the method, we embarked on this inter-laboratory validation study. The objective was to test whether the assay provided unbiased estimates of viable organisms, whether the precision of the method conformed to the theoretical value, and to test whether or not it was reproducible between facilities and operators when based on a common set of SOPs.

Unbiased enumeration of viable organisms

As with vital stains methods, the accuracy of SDC-MPN cannot be measured directly on natural samples of plankton because there is no accepted reference for μ_{viable} , the true concentration of viable organisms. In particular, it is not possible to assume that all organisms in an untreated natural sample are viable. Consequently, experiments on actively growing and killed cultures are an integral part of BWMS-testing method validation (IMO 2016d; MacIntyre and Cullen 2016). Stringent procedures ensured that cultures were in balanced, exponential growth, guaranteeing that essentially, all organisms in untreated cultures were viable (see [Supplementary Materials](#), “The 10-generation criterion for balanced growth”). The results of this study are fully consistent with the hypothesis that the SDC-MPN method enumerated all viable organisms, as has been demonstrated for 13 species of phytoplankton from 7 divisions (MacIntyre and Cullen 2016; Sun and Blatchley 2017; MacIntyre et al. 2018). Consistent, negative results for heat-killed cultures confirmed that the method is not subject to false-positive error, as previously demonstrated for six species in culture (MacIntyre and Cullen 2016; Sun and Blatchley 2017).

The FOA analysis (see “Results,” “Bias” and Fig. 3) confirmed that there was no detectable bias in the estimation of viable organisms with SDC-MPN. Analysis of variance failed to detect counteracting biases among species or laboratories that might not be detectable in pooled data. Addressing our first objective, it can be concluded that SDC-MPN, the base method of the MPN Dilution Culture+Motility method, classified actively growing phytoplankton as viable, and heat-killed phytoplankton as non-viable, without bias.

In contrast, an examination of the base assumption of the FDA/CMFDA+Motility method showed that fewer than half of the 24 species studied were accurately classified as live vs. dead when quantitative measurements of stain fluorescence of exponentially growing and heat-killed cultures of phytoplankton were compared (MacIntyre and Cullen 2016). Notably, no comparable culture-based studies were included in the validation of the FDA/CMFDA+Motility method, as described by the USA (Steinberg et al. 2011; IMO 2016d), and no other directly relevant studies have been published.

Expected vs. observed precision

The imprecision inherent to the SDC-MPN method is well-known (Cochran 1950)—multiplicative confidence intervals of $FCI = 2$ to 3 are typical for configurations that are used in BWMS testing (MacIntyre et al. 2018). Considering the outcome of single tests, this uncertainty from random measurement error compares poorly to FDA/CMFDA-Motility, especially if the samples are concentrated prior to microscopic examination, which improves the counting precision (e.g., 95% CI = $\pm 15\%$, Cullen 2018). But, as documented by Cullen (2018), the BWMS type-approval regime, which has a “one strike and you’re out” requirement for five consecutive successful test results, reverses conventional interpretations of precision and testing efficacy: The method with wider confidence limits is more protective because five consecutive false passes are highly unlikely and the greater risk of a single false failure imposes a margin of safety on the outcome five tests.

In the context of this study, theoretical estimates of method precision (Cochran 1950; Hurley and Roscoe 1983; Jarvis et al. 2010) provide useful benchmarks against which to compare results for three species of phytoplankton enumerated in each of three laboratories. An appropriate metric is the FOA, C_{MPN}/C_{viable} , log-transformed to accommodate the log-normal distribution of estimates of C_{MPN} . If experimental procedures introduced additional random measurement error into the MPN estimates, the average of the standard deviations of $\ln(\text{FOA})$ would be greater than the theoretical $\sigma_{\ln(\text{FOA})} = 0.55$. The observed standard deviation of $\ln(\text{FOA})$, 0.47, was not; the upper confidence limit of this estimate of $\sigma_{\ln(C_{MPN})}$ was only marginally higher at 0.56 (see “Results,” “Precision”). Addressing our second

objective, we conclude that as implemented in this study, the precision of the SDC-MPN method was not detectably degraded from the theoretical ideal.

Reproducibility

The “Best Practices” SDC-MPN method was developed at DAL, and an SOP appropriate for inter-laboratory comparison was prepared during a workshop with BLOS and U-SC, which implemented MPN for the first time during this study. Addressing the third objective of this study, we found that in the hands of experimentalists from these three laboratories with established competence in phytoplankton culture, the SDC-MPN method enumerated viable phytoplankton with no evidence of systematic error (bias) and with no degradation of precision compared to the theoretical ideal.

Test criteria

The SOP developed for this comparison was based on stringent criteria designed to prevent false-negative and false-positive results in cultures that had been treated with UVC radiation and which were assayed in conditions that differed from those to which they had been acclimated (MacIntyre et al. 2018). As such, they had to account for both variability in the relationship between chlorophyll *a* fluorescence and cell concentration and, potentially for a lag in which organisms recovered from treatment prior to the resumption of growth in the assay (cf. Hull et al. 2017). The “Best Practice” criteria required an 8× increase in fluorescence to account for acclimation to different conditions of growth and temperature and to allow for recovery from any bleaching resulting from UVC exposure. The criteria also specified a five-generation buffer in defining the duration of monitoring, t_{end} (MacIntyre et al. 2018).

These criteria are unnecessarily stringent for the experiments described here, in which exponentially growing cultures were acclimated to and then assayed under the same growth conditions. Lack of acclimative change to growth conditions or recovery from an imposed stress translated to equally accurate scoring of positives or negatives with relaxed stringency. There was no difference in the scores if the criterion for a positive score was set at the equivalent of one- or two-generation increases (2× and 4× increases in fluorescence) rather than the original three-generation increase (8× increases in fluorescence). As discussed previously (MacIntyre et al. 2018), accepting a lower threshold for defining a positive score and/or reducing the duration of the buffer in calculating t_{end} (Eq. 3) can reduce the commitment of personnel time for the assay.

Conclusions

The MPN Dilution Culture+Motility methodology is one of two that may be used to enumerate viable organisms in the 10–50- μm size range for BWMS Type Approval, but it is the only methodology suitable for treatment technologies that are designed to destroy the ability to reproduce rather than to kill. The base method, SDC-MPN, has been used for decades to enumerate phytoplankton in natural communities; only more recently has it been used to quantify a reduction in the number of viable organisms after exposure to environmental stressors, disinfection, or treatment with BWMS (Lehr et al. 2007; Wright and Welschmeyer 2015; MacIntyre et al. 2018). The last requires validation for use in Type-Approval testing. Inter-laboratory comparison is an important component of method validation, and for enumeration of viable or living phytoplankton, this requires testing on laboratory cultures grown under carefully controlled conditions. Three research laboratories implemented the same experimental procedures to enumerate viable organisms using appropriate reference material—uniformly viable and heat-killed cultures of three species of unicellular microalgae—and including explicit evaluation of precision, bias, and detection limits. Results were not significantly different from the expectation of unbiased enumeration of viable organisms with no degradation of precision from the theoretical ideal. Reproducibility was demonstrated even though the SDC-MPN protocol was implemented by two of the laboratories for the first time in this study. These results might not surprise practitioners, but they should not be taken for granted in the process of method validation. This study is, to our knowledge, the first for a method used to enumerate viable organisms in the 10–50- μm size range for BWMS Type Approval.

Acknowledgements This work was funded by Trojan Marinex, Hyde Marine, Alfa Laval, Cathelco, Desmi Ocean Guard, Panasia, and Hanovia. All are manufacturers of UV water-treatment systems. Supplementary funding was provided by the Strategic Cooperative Education Incentive Program of Nova Scotia.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Adams JK, Briski E, Ram JL, Bailey SA (2014) Evaluating the response of freshwater organisms to vital staining. *Manag Biol Invasion* 5: 197–208
- Andersen P, Thronsen J (2003) Estimating cell numbers. In: Hallegraeff GM, Anderson DM, Cembella AD (eds) *Manual on harmful marine microalgae*. UNESCO, Paris, pp 99–129
- Anderson DJ (1989) Determination of the lower limit of detection. *Clin Chem* 35:2152–2153
- Bevington PR (1969) *Data reduction and error analysis for the physical sciences*. McGraw Hill, Hightstown
- Blodgett RJ (2005) Upper and lower bounds for a serial dilution test. *J AOAC Int* 88:1227–1230
- Blodgett RJ (2010) Appendix 2: Most probable number from serial dilutions. In: *Bacteriological analytical manual*. 8th Edition. US Food and Drug Administration
- Brand LE, Guillard RRL, Murphy LS (1981) A method for the rapid and precise determination of acclimated phytoplankton reproduction rates. *J Plankton Res* 3:193–201
- Button DK, Schut F, Quang P, Martin R, Robertson BR (1993) Viability and isolation of marine-bacteria by dilution culture—theory, procedures, and initial results. *Appl Environ Microbiol* 59:881–891
- Cerino F, Zingone A (2006) A survey of cryptomonad diversity and seasonality at a coastal Mediterranean site. *Eur J Phycol* 41:363–378
- Cochran WG (1950) Estimation of bacterial densities by means of the most probable number. *Biometrics* 6:105–116
- Cullen JJ (2018) Quantitative framework for validating two methodologies that are used to enumerate viable organisms for type approval of ballast water management systems. *Sci Total Environ* 627:1602–1626
- Cullen JJ, MacIntyre HL (2016) On the use of the serial dilution culture method to enumerate viable phytoplankton in natural communities of plankton subjected to ballast water treatment. *J Appl Phycol* 28: 279–298
- Feldsine P, Abeyta C, Andrews WH (2002) AOAC INTERNATIONAL methods committee guidelines for validation of qualitative and quantitative food microbiological official methods of analysis. *J AOAC Int* 85:1187–1200
- First MR, Drake LA (2013) Approaches for determining the effects of UV radiation on microorganisms in ballast water. *Manag Biol Invasion* 4:87–99
- Furuya K, Marumo R (1983) The structure of the phytoplankton community in the subsurface chlorophyll maxima in the western North Pacific Ocean. *J Plankton Res* 5:393–406
- Guillard RRL (1973) Methods for microflagellates and nannoplankton. In: Stein JR (ed) *Handbook of Phycological Methods - Culture Methods and Growth Measurements*. Cambridge University Press, Cambridge
- Haas CN, Heller B (1988) Test of the validity of the Poisson assumption for analysis of most-probable-number results. *Appl Environ Microbiol* 54:2996–3002
- Hull NM, Isola MR, Petri B, Chan P-S, Linden KG (2017) Algal DNA repair kinetics support culture-based enumeration for validation of ultraviolet disinfection ballast water treatment systems. *Environ Sci Technol Lett* 4:192–196
- Hurley MA, Roscoe ME (1983) Automated statistical-analysis of microbial enumeration by dilution series. *J Appl Bacteriol* 55:159–164
- IMO (2004) Convention BWM/CONF/36 International convention for the control and Management of Ships' ballast water and sediments. International Maritime Organization <http://www.imo.org/home.asp>,
- IMO (2016a) International Maritime Organization Marine Environment Protection Committee, Resolution MEPC.279(70). 2016 Guidelines for Approval of Ballast Water Management Systems (G8). <http://www.imo.org/en/KnowledgeCentre/IndexofIMOResolutions/Marine-Environment-Protection-Committee-%28MEPC%29/Documents/MEPC.279%2870%29.pdf>
- IMO (2016b) Review of the guidelines for approval of ballast water management systems (G8): analysis methods for determining the viability of organisms in the 10 to 50 μm size class. Originator: Denmark and Norway IMO PPR 4/7, 12 October 2016 IMO PPR 4/7, 12 October 2016
- IMO (2016c) Review of the guidelines for approval of ballast water management systems (G8): Comments on document PPR 4/7 on

- analysis methods for determining the viability of organisms in the 10 to 50 μm size class. Originator: United States. IMO PPR 4/7/1, 23 November 2016
- IMO (2016d) Review of the guidelines for approval of ballast water management systems (G8): Description of the FDA/CMFDA + Motility Method. Originator: United States. IMO PPR 4/INF.10, 23 November 2016
- IMO (2017) International convention for the control and management of ships' ballast water and sediments, 2004: Guidance on methodologies that may be used for enumerating viable organisms for type approval of ballast water management systems. BWM.2-Circ.61, 26 July 2017
- ISO (1994) Accuracy (trueness and precision) of measurement methods and results—part 1: General principles and definitions. <https://www.iso.org/obp/ui/#iso:std:iso:5725:-1:ed-1:v1:en>. Accessed May 12, 2017
- Jarvis B, Wilrich C, Wilrich PT (2010) Reconsideration of the derivation of most probable numbers, their standard deviations, confidence bounds and rarity values. *J Appl Microbiol* 109:1660–1667
- Lehr CR, Frank SD, Norris TB, D'Imperio S, Kalinin AV, Toplin JA, Castenholz RW, McDermott TR (2007) Cyanidia (cyanidiales) population diversity and dynamics in an acid-sulfate-chloride spring in Yellowstone National Park. *J Phycol* 43:3–14
- MacIntyre HL, Cullen JJ (2005) Using cultures to investigate the physiological ecology of microalgae. In: Andersen RA (ed) *Algal Culture Techniques*. Academic Press, New York, pp 287–326
- MacIntyre HL, Cullen JJ (2016) Classification of phytoplankton cells as live or dead using the vital stains fluorescein diacetate and 5-chloromethylfluorescein diacetate. *J Phycol* 52:572–589
- MacIntyre HL, Cullen JJ, Whitsitt TJ, Petri B (2018) Enumerating viable phytoplankton using a culture-based most probable number assay following ultraviolet-C treatment. *J Appl Phycol* 30:1073–1094
- Madsen T, Petersen GI (2015) MPN Dilution-Culture Method as an alternative method for measuring treatment of organisms in the 10–50 μm size category. DHI Ref: 11097400 (Viewed 19 November 2017). DHI, <http://mpnballastwaterfacts.com/wp-content/uploads/2016/01/DHI-Support-final.pdf>
- McCrary MH (1915) The numerical interpretation of fermentation-tube results. *J Infect Dis* 17:183–212
- Miller JN, Miller JC (2005) *Statistics and chemometrics for analytical chemistry*. 5th edn. Pearson Education Limited, Harlow
- Mishalanie E, Lesnik B, Araki R, Segall R (2016) Validation and peer review of U.S. Environmental Protection Agency chemical methods of analysis. The EPA Forum on Environmental Measurements (FEM)
- Parkhill JP, Maillet G, Cullen JJ (2001) Fluorescence-based maximal quantum yield for PSII as a diagnostic of nutrient stress. *J Phycol* 37:517–529
- Parshionikar S, Hunt M, Genthner F, Lincoff A, Haugland R, Cottrill M, Harris S, Pfaller S, Shanks O, Grimm A, Oshiro R, Sivaganesan M (2009) Method validation of US Environmental Protection Agency (EPA) microbiological methods of analysis, FEM 2009-01. vol Revision. U.S. Environmental Protection Agency, Fort Meade
- Slifko TR, Huffinan DE, Rose JB (1999) A most-probable-number assay for enumeration of infectious *Cryptosporidium parvum* oocysts. *Appl Environ Microbiol* 65:3936–3941
- Steinberg MK, Lemieux EJ, Drake LA (2011) Determining the viability of marine protists using a combination of vital, fluorescent stains. *Mar Biol* 158:1431–1437
- Sun Z, Blatchley ER (2017) *Tetraselmis* as a challenge organism for validation of ballast water UV systems. *Water Res* 121:311–319
- Thronsdon J (1978) The dilution-culture method. In: Sournia A (ed) *Phytoplankton manual*. UNESCO, Paris, pp 218–224
- Thronsdon J, Kristiansen S (1991) *Micromonas pusilla* (Prasinophyceae) as part of picoplankton and nanoplankton communities of the Barents Sea. *Polar Res* 10:201–207
- USCG (2012) Standards for living organisms in ships' ballast water discharged in US waters. Federal Register 77:17254–17320. US Coast Guard
- USCG (2016) 7/12/2016: Final action on ballast water management system appeals. Coast Guard Maritime Commons
- USEPA (2006) Ultraviolet Disinfection Guidance Manual For The Final Long Term 2 Enhanced Surface Water Treatment Rule (EPA 815-R-06-007). vol EPA 815-R-06-007. Office of Water (4601), Washington DC
- USEPA (2010) Generic protocol for the verification of ballast water treatment technology., vol Version 5.1. EPA/600/R-10/146. US Environmental Protection Agency Environmental Technology Verification Program, Washington, DC
- Wilrich PT (2017) EXCEL program for the determination of the Most Probable Numbers (MPN), their standard deviations, confidence bounds and rarity values, version 5, http://www.wiwiss.fu-berlin.de/fachbereich/vwl/iso/ehemalige/professoren/wilrich/MPN_ver5.xls, Accessed 23 May 2017
- Wood AM, Everrood RC, Wingard LM (2005) Measuring growth rates in microalgal cultures. In: Andersen RA (ed) *Algal culture techniques*. Academic Press, New York, pp 269–285
- Wright DA, Welschmeyer NA (2015) Establishing benchmarks in compliance assessment for the ballast water management convention by port state control. *J Mar Eng Technol* 14:9–18