



## Comparative assessment of four ballast water compliance monitoring devices with natural UV-treated water using IMO's monitoring approaches

Leonardo Romero-Martínez<sup>a,\*</sup>, Cees van Slooten<sup>b</sup>, Michiel van Harten<sup>b</sup>, Enrique Nebot<sup>a</sup>, Louis Peperzak<sup>c</sup>

<sup>a</sup> Department of Environmental Technologies, Faculty of Marine and Environmental Sciences, INMAR – Marine Research Institute, CEIMAR – International Campus of Excellence of the Sea, University of Cadiz, Campus of Puerto Real, 11510 Puerto Real, Spain

<sup>b</sup> Control Union Vessel Performance Centre B.V., Boompjes 270, 3011 XZ Rotterdam, the Netherlands

<sup>c</sup> NIOZ Royal Netherlands Institute for Sea Research, P.O. Box 59, NL-1790 AB Texel, the Netherlands

### ARTICLE INFO

#### Keywords:

Ballast water  
UV-treatment  
CMD-validation  
Cohen's kappa

### ABSTRACT

Compliance Monitoring Devices (CMDs) are instruments indicating if the 10–50  $\mu\text{m}$  organism abundance in ballast water complies with the International Maritime Organizations' discharge standard. For that, they trade the presumed accuracy of detailed methods for speed and simplicity. In an experiment using UV-treated water, four CMDs were compared, using a Cohen's kappa analysis (ISO 3725), with two detailed methods: Fluorescence Microscopy (FM) and the Most Probable Number (MPN). Of two CMDs, BallastWISE and B-QUA ATP, the standard (non)-compliance limits were adapted. Correlated to FM (non)-compliance results, ranking resulted in: 1) BallastWISE, 2) MPN, 3) Ballast-Check-2 and 4) WALZ-PAM. The ATP method had too few data for this ranking. MPN proved problematic: the abundance of 10–50  $\mu\text{m}$  organism decreased while that <10  $\mu\text{m}$  organisms increased. It is concluded that in case of UV-treatment, the outcome of CMD-validations will depend on the detailed method, the CMD instrument, and the (non)-compliance limits chosen.

### 1. Introduction

Maritime transport is the major means for moving goods worldwide. Merchant vessels use ballast water to improve their stability and maneuverability, especially when they have to navigate without cargo or below their capacity. Port water, containing organisms and sediments, is pumped into the ballast tanks inside the ship; after the voyage and the reception of new cargo, the ballast water is discharged into the environment. Once released, transported species can develop and then colonize the recipient ecosystem, causing environmental, economic and health concerns (Khandeparker and Anil, 2017; Pimentel et al., 2005; Ruiz et al., 2000).

Ship owners currently implement international regulations to minimize the threats derived from the use of ballast water. The International Maritime Organization (IMO) adopted the International Convention for the Control and Management of Ships' Ballast Water and Sediments (BWMC) in 2004, which entered into force in 2017. Regulation D-2 of the BWMC stipulates the standards that must be fulfilled in the ballast water discharge: <10 viable organisms  $\text{m}^{-3}$  greater than or equal to 50

$\mu\text{m}$  in minimum dimension (>50  $\mu\text{m}$ ), and <10 viable organisms  $\text{mL}^{-1}$  < 50  $\mu\text{m}$  in minimum dimension and greater than or equal to 10  $\mu\text{m}$  in minimum dimension (10–50  $\mu\text{m}$ ) (IMO, 2004). Additionally, IMO regulates three bacterial indicators. BWMC D-2 understands the term “viable organisms” as those able to reproduce successfully (IMO, 2018). In turn, the United States Coast Guard (USCG) adopted its own regulations, in which the maximum allowable concentrations are similar as the BWMC, but referring to “living” instead “viable” organisms, where living >50  $\mu\text{m}$  organisms should move and 10–50  $\mu\text{m}$  organisms should move or be fluorescent green after treatment with an FDA (FluoresceinDiAcetate) - CMFDA (5-Chloromethyl FluoresceinDiAcetate) stain (USCG, 2012). The theoretical difference is that viable cells are living cells but that live organisms that show green fluorescence after staining may not be viable and able to reproduce (Peperzak and Brussaard, 2011).

To be compliant with the BWMC and USCG regulations ships are being equipped with a ballast water management system (BWMS). BWMSs need to be tested successfully in land-based facilities and on-board trials to obtain their Type Approval (IMO, 2018). Once the

\* Corresponding author.

E-mail address: [leonardo.romero@uca.es](mailto:leonardo.romero@uca.es) (L. Romero-Martínez).

<https://doi.org/10.1016/j.marpolbul.2024.117193>

Received 12 September 2024; Received in revised form 19 October 2024; Accepted 21 October 2024

Available online 7 November 2024

0025-326X/© 2024 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

BWMS is installed, a commissioning test is required to demonstrate that all processes are compliant with regulations (IMO, 2020a). Additionally, Port State Control (PSC) authorities may require compliance tests during the commercial activity of the ships. While so-called detailed microscopic tests are used in type approval testing, faster indicative tests can be applied in IMO commissioning and PSC compliance tests. Commissioning testing is not required by the USCG while compliance testing for the US Vessel General Permit (VGP), regulated by the U.S. Environmental Protection Agency (EPA), is limited to the indicator microbes.

The typical configuration of a BWMS is an initial filtration followed by a physical and/or chemical inactivation technology. Almost half of the approved BWMSs use ultraviolet (UV) radiation to inactivate the target organisms (Gerhard et al., 2019; IMO, 2020b). The major non-lethal effect from UV radiation is the formation of DNA dimers, which prevent the organisms from replication, i.e. they are rendered non-viable (Li et al., 2020; Sinha and Häder, 2002). However, the esterases that convert FDA and CMFDA to a green fluorescent dye may still be active, which means that a non-viable cell is scored “live” in the microscopic analysis. Nevertheless, UV radiation, especially at high doses, also causes damage to cell structures such as membranes and proteins, which may lead to the cell death (Blatchley III et al., 2018). In this context, UV-based BWMSs may fulfill the IMO D-2 non-viability standards with lower doses compared to the non-vitality USCG standards.

A series of analytic techniques and instruments are available for determining compliance with the IMO and USCG standards for 10–50  $\mu\text{m}$  and  $> 50 \mu\text{m}$  organisms, which are divided into detailed and indicative analyses (Outinen and Lehtiniemi, 2017). Detailed analyses, FDA-CMFD fluorescence microscopy (FM) and Serial Dilution Culture-Most Probable Number (MPN) in vitro incubation, require laboratory examination by trained personnel, are more expensive, time consuming and, therefore, not appropriate for on board analysis. The two detailed analyses focus either on vitality (FM) or on viability (MPN) of the organisms and deliver organism concentrations. Indicative compliance instruments or Compliance Monitoring Devices (CMDs), are usually less accurate than detailed analyses because they mostly use biomass proxies such as chlorophyll fluorescence intensity or Adenosine TriPhosphate (ATP) concentration instead of organism counts. In addition, induced chlorophyll fluorescence can be used as a vitality proxy (Hansen et al., 2023; Peperzak and Brussaard, 2011; Stehouwer, 2016). Unfortunately, some CMDs can only detect phytoplankton, mostly 10–50  $\mu\text{m}$  organisms, and not zooplankton that is mostly  $>50 \mu\text{m}$ . Although indicative methods are less accurate, they have the advantage of quick on-board organism estimation and they can be used by technicians who are not trained as biologists.

In the present study, two detailed and four indicative methods for 10–50  $\mu\text{m}$  organisms were examined and compared (Table 1). In FM, a detailed method, a microscopist enumerates motile as well as vital 10–50  $\mu\text{m}$  organisms that have converted FDA and CMFDA

intracellularly to the green-fluorescent fluorescein. The method relies on FDA or CMFDA conversion in all living organisms and the detection of green fluorescence. However, not all vital organisms are stained or when they are stained, the fluorescence intensity may be too low to detect (Hansen et al., 2023; Peperzak and Brussaard, 2011). In the second detailed method, MPN motile heterotrophs (10–50  $\mu\text{m}$ ) are enumerated by microscopy, after which the sample is stepwise diluted and incubated in tubes for two weeks to detect autotroph growth by chlorophyll-fluorescence (IMO, 2019). Growth in the most diluted sample is a measure for the original viable autotroph concentration, to which the original microscopically measured heterotroph concentration is added (First and Drake, 2013; IMO, 2019). In MPN it is quintessential that all viable autotrophic organisms are able to reproduce in vitro, which is probably not the case (Drake et al., 2016; Peperzak, 2024). Because chlorophyll-fluorescence in the MPN may originate from autotrophs  $<10 \mu\text{m}$  or  $\geq 50 \mu\text{m}$ , the method is also less specific than microscope counting. By IMO-definition the result of a FM or MPN test is either compliant or non-compliant. The indicative methods may underestimate the risk because they are not as accurate as the FDA/CMFDA detailed method. Therefore, the manufacturers have raised their (non)-compliance limits relative to D-2 and thus CMDs tend to overestimate the risk.

Three of the four indicative methods that were tested in this study use biomass proxies (chlorophyll and ATP concentrations) or induced chlorophyll fluorescence ( $F_m$ ) as a vitality proxy. The fluorescence vitality proxies chosen,  $F_0/F_m$ , and  $F_v = (F_m - F_0)/F_m$ , where  $F_0$  is the dark-adapted chlorophyll fluorescence, is dependent on the manufacturer. The Ballast-Check 2 uses  $F_v/F_m$  and bulk chlorophyll fluorescence to estimate D-2 compliance. If a high risk of non-compliance is detected, the fluorescence of a  $< 10 \mu\text{m}$  sample needs to be subtracted from the original measurement so that the final conclusion is based on  $\geq 10 \mu\text{m}$  phytoplankton. The WALZ-PAM, not a CMD per se but extensively used in type approval testing by NIOZ, indicates (non)-compliance from  $F_v/F_m$  only from all phytoplankton size classes (Stehouwer et al., 2015). The B-QUA ATP method defines three size classes by 10 and 50  $\mu\text{m}$  filtration, measures the bulk ATP concentration in all three, and suggests 10–50  $\mu\text{m}$  D-2 (non)-compliance based on three risk levels that correspond with two ATP concentrations (Table 1) (Lo Curto et al., 2018). The fourth instrument tested, the BallastWISE, is a counting device that determines 10–50  $\mu\text{m}$  and  $>50 \mu\text{m}$  organism size and motility with two cameras and non-motile 10–50  $\mu\text{m}$  vitality by measuring a fluorescence vitality proxy ( $F_m/F_0$ ). Details on 10–50 organism sizing and counting is provided in Blackburn et al. (2022). Contrary to the other three indicative methods, the BallastWISE delivers organism concentrations and D-2 compliance is based on three risk levels, now corresponding to three combined auto- and heterotrophic organism concentrations as well as  $F_m/F_0$  (Table 1).

The specific CMDs have, besides a low-risk (compliant) limit, an intermediate and a high-risk limit (Table 1) that is meant to prevent false-negative conclusions. A false-negative is a test result that is declared negative (compliant) when in fact the sample is positive (non-

**Table 1**

Detailed (Fluorescence Microscopy (FM) and MPN) and four indicative methods with their respective compliance or risk limits. N.a. = not applicable. The Ballast-Check 2 limits are “very low risk”, “low risk” (both are Compliant) and “high risk” (non-compliant). If the concentration is higher than the Ballast-Check 2 upper detection limit ( $>2000 \text{ cells mL}^{-1}$ ), the risk is always “high”. Size differentiation in the indicative methods is by measuring (BallastWISE), filtration (Ballast-Check 2 and B-QUA ATP) or is not performed (WALZ-PAM). The BallastWISE adds autotrophs (with  $F_m/F_0 > 3.5$ ) and heterotrophs (motility).

Method	Size class	Heterotrophs included?	Compliant/low risk	Intermediate	Not compliant/high risk
FM	$\geq 10\text{--}50 \mu\text{m}$	Yes	$<10 \text{ mL}^{-1}$	n.a.	$\geq 10$
MPN	All sizes (if sieving is not applied)	Yes; in additional microscope count	$<10 \text{ mL}^{-1}\text{L}$	n.a.	$\geq 10$
BallastWise	$\geq 10\text{--}50 \mu\text{m}$	Yes	$<13 \text{ mL}^{-1}$ ( $F_m/F_0 > 3.5$ ) incl. heterotrophs	$13\text{--}30 \text{ mL}^{-1}$ ( $F_m/F_0 > 3.5$ ) incl. heterotrophs	$>30 \text{ mL}^{-1}$ ( $F_m/F_0 > 3.5$ ) incl. heterotrophs
Ballast-Check 2	$\geq 10 \mu\text{m}$	No	$<10 \text{ mL}^{-1}$ or $\geq 10 \text{ mL}^{-1}$ and $F_v/F_m < 0.25$	n.a.	$\geq 10 \text{ mL}^{-1}$ and $F_v/F_m \geq 0.25$
WALZ-PAM	All sizes	No	$F_v/F_m < 0.10$	n.a.	$F_v/F_m \geq 0.10$
B-QUA ATP	$\geq 10\text{--}50 \mu\text{m}$	Yes	Old: $<500 \text{ pg mL}^{-1}$ New: $<50 \text{ pg mL}^{-1}$	Old: $500\text{--}1500 \text{ pg mL}^{-1}$ New: $50\text{--}150 \text{ pg mL}^{-1}$	Old: $>1500 \text{ pg mL}^{-1}$ New: $>150 \text{ pg mL}^{-1}$

compliant). Vice versa, a false-positive is a test result that is declared positive (non-compliant) when the sample is in fact negative (compliant) (IMO, 2023). The Walz-PAM limit of 0.1 is based on extensive type approval experience at the NIOZ test facility (Stehouwer et al., 2015). The Ballast-Check 2 algorithm that relates chlorophyll-fluorescence to estimated cell concentrations is unknown. In addition, the origin of the (non)-compliance ATP-concentrations of 500 and 1500 pg mL<sup>-1</sup> are unknown and appear to be too high (Peperzak, 2023) and 90 %-reduced ATP (non)-compliance limits will be tested (Table 1). This new limit is based on a non-compliance limit derived from: 1) the D-2 non-compliance concentration (10 organisms mL<sup>-1</sup>), 2) which have the average size of coastal phytoplankton: 15 µm spherical equivalent diameter (Welschmeyer and Kuo, 2016), 3) not the average (0.6 mM) but the maximum ATP concentration (2 mM) in aquatic eukaryotic microorganisms (Bochdansky et al., 2021), 4) the potentially maximum increase (3×) in ATP extraction by the lysis method (Peperzak, 2023), resulting in 50 pg ATP mL<sup>-1</sup>, and 5), increasing the non-compliance limit by a factor of three, as in the original B-QUA limit, to 150 pg ATP mL<sup>-1</sup> (Table 1). Note that the BallastWISE also uses a factor three between compliance and non-compliance, i.e. from 13 organisms mL<sup>-1</sup> to >30 organisms mL<sup>-1</sup> as measured by this instrument.

The aim of this study is the comparison of two detailed techniques: FM and MPN with four indicative techniques: BallastWISE, Ballast-Check 2, Walz-PAM and B-QUA ATP by comparing their ability to detect compliance and non-compliance of the 10–50 µm size range, after UV-treatment of natural marine water. Different UV-doses were applied to induce increasing levels of stress to the 10–50 µm size fraction without the intention to simulate full ballast water treatment. The test results were compared by calculating Cohen's kappa as prescribed by ISO 3725 of the International Standard Organization (ISO, 2023).

## 2. Material and methods

### 2.1. Challenge water

#### 2.1.1. Challenge water pre-treatment

Ambient water was collected in August and September 2021 from the NIOZ harbor on Texel (The Netherlands; 53°00'22" N, 4°47'38" E). To obtain 10–50 µm organism abundances adequate for observing the effect of UV-treatment, water was pumped into a 3 m<sup>3</sup> open top cylindrical plastic container (Kemeling, Naaldwijk, The Netherlands). While pumping, water was filtered through a 50 µm mesh to remove large phytoplankton grazers. To promote phytoplankton growth, nutrient solutions with nitrogen (9 mL m<sup>-3</sup>), phosphorus (21 mL m<sup>-3</sup>) and trace metals (15 mL m<sup>-3</sup>) from Easy-Life® (Duiiven, The Netherlands), as well as a self-prepared Na<sub>2</sub>SiO<sub>3</sub> solution (10 µM Si final concentration) were added. The test water was incubated outdoors for 3–4 days with continuous aeration under natural light and temperature conditions. After this pre-treatment incubation, the challenge water was pumped into a 1 m<sup>3</sup> International Business Container (IBC) and then subjected to the UV-treatment.

#### 2.1.2. Salinity, temperature and UVT

Sample salinity and temperature were measured directly prior to the experiments with a GMH 3400 Digital Conductivity Meter with Pt sensor (Greisinger, Germany). UV transmittance (UVT) was measured in triplicate samples with a UV254 Go! Photometer (Photonic Measurements Ltd., Northern Ireland).

### 2.2. UV system and irradiation procedure

#### 2.2.1. Description of the UV system and UV dose calculation

Disinfection was applied using a flow-through UV-reactor (from now on the "UV-reactor") mounted into a test rig that was obtained from Alfa Laval (Tumba, Sweden). It contained a UV reactor with one medium-pressure 1 kW Hg lamp (117 W output) that was inserted in a

transversal position with respect to the water flow. The set-up was composed of a series of IBCs, a centrifugal pump (TYPE 3 M/I 40–160/4 IE3, EBARA Pumps Europe, Italy) connected to a frequency controller (ACS800, ABB, Switzerland) to control the flow rate, a flow meter (Endress+Hauser Promag 10D50 electromagnetic flow meter, Germany) and the UV-reactor (Fig. 1). The parts were connected with 51 mm PVC hoses. One IBC was filled with tap water to be recirculated while the lamp was warming up, and to be pumped into a waste IBC for flushing the system after each treatment. The target and treated water IBCs were swapped to apply repeated treatments (see Section 2.3 UV treatment).

An approximate UV-dose applied was estimated using a diagram with flow rate and UVT obtained from Alfa Laval (C. Hedberg, personal communication). We did not aim to apply exact doses but to apply incremental doses that would induce losses of living organisms that would be detectable by both detailed and indicative methods. Depending on the aimed UV-dose and the actual UVT a flow rate was set. The minimum flow rate that could be handled properly was 30 L min<sup>-1</sup>, corresponding to a reduction equivalent dose (RED) of ca. 165 mJ cm<sup>-2</sup>. The maximum flow rate used was 167 L min<sup>-1</sup>, corresponding to a RED of ca. 55 mJ cm<sup>-2</sup>.

#### 2.2.2. UV-treatment application

To obtain stepwise higher UV doses, the test water was pumped through the UV reactor multiple times, each time collecting the treated water in a clean IBC. In a separate test with the UV-lamp off, there was no effect by pumping on the total concentration of 10–50 µm organisms of up to four passes through the UV reactor (Fig. S1 and Table S1 in Supplementary Material). In one experiment with UV doses ranging from 0 to 160–320–480–640 mJ cm<sup>-2</sup>, UVT hardly changed (–2 %), meaning that UVT only needed to be measured at the start of the experiments and that flow rate is the only determinant of UV dose in one specific experiment. In the same experiment the water temperature increase was deemed negligible: only +1.2 °C in total. In conclusion, except for the direct effect of UV, the repeated pumping of the test water did not fundamentally alter water quality. In total, seven experiments were performed, in each 2–4 doses were applied. The total range of doses was 55 to 700 mJ cm<sup>-2</sup>. Due to time constraints, not all analysis techniques could be used simultaneously in each experiment.

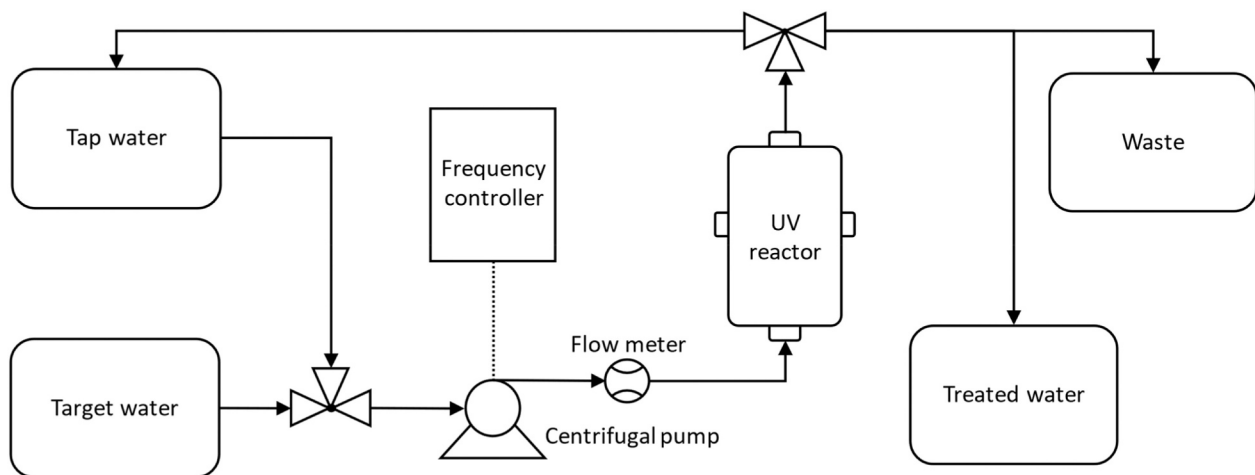
### 2.3. Compliance methods

#### 2.3.1. Fluorescence microscopy (FM)

FDA + CMFDA stock solution (500 µM + 250 µM in acetone, Invitrogen Inc., Carlsbad, CA, USA), stored at –20 °C, was added to 1.5 mL aliquots of well-mixed sample to reach final 5 µM and 2.5 µM concentrations. The samples were analyzed in duplicate. Samples were incubated in low light for 10 min at room temperature, then pipetted into a 1 mL graticules S52 Sedgewick-Rafter counting chamber (SPI Supplies, West Chester, PA, USA). The green fluorescence or motility of vital 10–50 µm plankton cells was observed with a Zeiss Axio Scope.A1 (Carl Zeiss AG, Jena, Germany) epifluorescence microscope, equipped with blue LED-light (470 nm) and a blue excitation filter set (BP 450–490 nm, FT 510 nm, LP 515 nm), at 100× or 200× magnification. Rows of 50 µL were analyzed until a minimum of 200 green fluorescent (vital) cells were scored. If <200 vital cells were present, the entire chamber was analyzed. The analysis was completed within 30 min of the FDA + CMFDA addition.

#### 2.3.2. MPN

Samples were analyzed according to the IMO Sub-Committee on Pollution Prevention and Response (PPR 7) protocol (IMO, 2019). A gentle gravity filtration step to remove <10 µm and >50 µm organisms was tested, but this option was discarded because it introduced a noticeable reduction on the >10 µm organisms while <10 µm were still present, comparable to previous results (de Castro and Veldhuis, 2019). Therefore, as by the PPR 7 protocol, non-filtered samples were diluted in



**Fig. 1.** Schematic of the experimental set-up. The tap, target, waste and treated water was held in IBCs. Flows were directed with three-way valves. Target or tap (rinse) water was pumped through the UV reactor using a centrifugal pump. Stepwise increasing UV doses were achieved by passing the treated water 1–4 times through the reactor without changing the flow rate.

ten-fold steps with 0.2  $\mu\text{m}$  filtered local seawater enriched with nutrients (Peperzak and Brussaard, 2011). The range of dilutions was initially selected according to the PPR 7 recommendations (experiments 5 to 7), however, due to the frequent sets of all positive tubes in control samples, the number of dilutions was increased up to 6 (i.e.,  $10^{-6}$ ) in experiments 9 to 12, to obtain a proper combination of positive and negative tubes (Annex 8.1). For each ten-fold dilution level, 1 mL of the diluted sample was added to five tubes (13 mm outer diameter, 11 mm inner and 100 mm length) with 5 mL of growth medium and covered with Parafilm® to prevent evaporation, contamination, and to ease mixing prior to fluorescence measurements. As part of the quality control a set of five blank tubes with only culture medium was also prepared. All tubes were incubated in a climate room at 15 °C, 120  $\mu\text{mol photons m}^{-2} \text{s}^{-1}$  and a 16 h:8 h L:D photoperiod. Chlorophyll fluorescence was measured by inserting the tubes into a 10-AU Fluorometer (Turner Designs, CA, USA), on day 0 and day 14 as required by the PPR 7 protocol. According to the PPR 7 protocol, tubes were scored positive if the fluorescence was greater than four times the standard deviation of the five initial measurements of the dilution. As an additional second check, following a recommendation by Cullen and MacIntyre (2016), a third fluorescence measurement was made on day 21 to detect the contribution of slow growing organisms to the cell concentration. There were some occasions in which a tube was positive after 14 days and negative after 21 days; in these cases, the score at day 21 was manually corrected into positive. The set of scores were introduced in an ISO spreadsheet for MPN calculation (<https://standards.iso.org/iso/7218/> (Jarvis et al., 2010)). Motile heterotrophs in the 10–50  $\mu\text{m}$  size class were no longer observed in the FM counts after UV treatment so their concentrations had not to be added to the viable autotroph MPN concentration.

Because the MPN method is solely based on bulk chlorophyll fluorescence, autotrophic organisms with sizes outside the 10–50  $\mu\text{m}$  range would contribute to false-positive readings. Therefore, an additional third check was performed in the final four (9–12) experiments by microscopically checking the organism composition, their size (minimum dimensions), and their ability to produce red fluorescence. From the set of tubes in the control, and the highest fluorescence-positive dilution, 100  $\mu\text{L}$  were examined in a 1 mL Sedgewick-Rafter chamber at 100 $\times$  to 200 $\times$  magnification with the same microscope and fluorescence settings to detect red autotrophic chlorophyll fluorescence as for the FM counts. In addition, >50  $\mu\text{m}$  predators on 10–50  $\mu\text{m}$  autotrophic organisms could lead to false-negative results. In experiment 9, 1 mL of the first negative tube was examined at 40 $\times$  magnification to check for any large predators that might have caused a negative fluorescence score by consuming red fluorescent autotrophs. This check was

abandoned in the following experiments because no predators were observed, and the chance of finding >50  $\mu\text{m}$  autotrophs at dilutions in the range of  $10^{-3}$ , in <6 mL, even in BWMS type approval D-2 compliant water ( $\geq 10^5 \text{ m}^{-3}$ ), which NIOZ harbor water usually is not in summer, was considered negligible.

### 2.3.3. BallastWISE

BallastWISE (MicroWISE, Denmark) samples were analyzed according to the manufacturer protocol version 5.5. In version 5.5 the  $F_0$  was set to 9 with a vitality proxy threshold  $F_m/F_0 = 3.5$ . In between BallastWISE runs all analysis bottles were thoroughly cleaned with tap water and the tubing and cuvette was flushed with new sample water before analysis. The 10–50  $\mu\text{m}$  cuvette volume is 24  $\mu\text{L}$  and, depending on organism concentration, a maximum of 40 cuvettes (960  $\mu\text{L}$ ) were analyzed per sample. In case of high organism concentrations, the BallastWISE system stops when 200 organisms are counted. All samples were analyzed in duplicate or triplicate if measurements were inconsistent. According to MicroWise, the counting accuracy is:  $10 \pm 50\%$ ,  $20 \pm 30\%$ ,  $30 \pm 20\%$ , calculated per mL at 95% confidence.

### 2.3.4. WALZ-PAM

Samples were analyzed in triplicate with a WALZ-PAM (Heinz Walz GmbH, Germany), consisting of a Water-ED, PAM-Control and a laptop with Win-Control version 3.29. Prior to the measurement, samples were stored dark for 20 min. The device was blanked with 0.2  $\mu\text{m}$  filtered seawater. To maintain linearity, the Photomultiplier (PM) gain setting was adjusted per sample in steps of 5 units such that the fluorescence ( $F_t$ ) was  $100 < F_t < 1000$ .

### 2.3.5. Ballast-check 2

Samples were measured in triplicate with a Ballast-Check 2 (Turner Designs, San Jose, USA) according to the manufacturer protocol (February 28, 2018) after being checked with the “Calibration Check Standard” supplied with the instrument. According to the protocol, “high risk” samples were filtered through a 10  $\mu\text{m}$  gauze to remove small (<10  $\mu\text{m}$ ) phytoplankton and analyzed again. Recorded were not only the “very low,” “low” and “high”-risk assessment of the instrument, but also the calculated organism abundance and  $F_v/f_m$ . If the concentration is higher than the upper detection limit ( $>2000 \text{ cells mL}^{-1}$ ), the results is “high risk” without measuring  $F_v/F_m$ .

### 2.3.6. BQUA PLUS ATP

Samples were prepared and the 10–50  $\mu\text{m}$  size fraction was measured according to the Luminultra® BQUA PLUS (2020) protocol using the

manufacturer's equipment in accordance with the manufacturer instructions. Due to time constraints samples were not analyzed in duplicate.

## 2.4. Data treatment

### 2.4.1. Ranking and interpreting “intermediate risk”

The result of detailed analyses (FM and MPN) is commonly interpreted as compliant ( $<10$  organisms  $\text{mL}^{-1}$ ) or non-compliant with the D-2 standard. A CMD is defined by IMO (PPR 10/7, Draft Protocol, Annex 2, 16-1-2023) as an instrument to determine compliance or non-compliance with a discharge standard. However, two CMDs tested here, BallastWISE and BQUA PLUS, also report an intermediate outcome which can be interpreted as an “intermediate risk”. The question is how to conclude compliance or non-compliance from the CMD test result (“low risk”, “intermediate risk” or “high risk”)?

- Should the discharge be considered compliant only when the CMD indicates a “low risk”? This means that the “intermediate risk” is a non-compliant outcome. This interpretation carries a lower chance that the discharge was in fact non-compliant. It is safer in terms of environmental protection.
- Or should the discharge be considered non-compliant only when the CMD indicates a “high risk”? This means that the “intermediate risk” is also a compliant outcome. This interpretation carries a higher chance that the discharge was in fact non-compliant. It is safer in terms of economy because the chance that a BWMS passes compliance testing increases (no need for retesting).

In ranking the results in terms of percentage compliance of all samples for all techniques, the environmental protection option (low risk only) was chosen for BallastWISE and B-QUA ATP.

### 2.4.2. Statistical analyses

Where applicable, duplicate and triplicate data were averaged. Linear regression was performed in SYSTAT 10.2, testing the null hypothesis that there is no correlation between the dependent (indicative) and the independent (detailed) variables.  $P$  is the chance that this hypothesis is rejected incorrectly and  $R^2$  is the proportion of variance that is explained by the regression model.

Cohen's  $\kappa$  method according to ISO 3725 (ISO, 2023) was used to determine the agreement between test results and a reference method for both treated and untreated samples. FM was chosen as the reference because it is a detailed method accepted by both the IMO and the USCG. Cohen's  $\kappa$  was calculated in Excel 365 in a matrix with the number of agreements in exceeding/meeting/indeterminate of the discharge standard. The “low risk”, “intermediate risk” and “high risk” (Table 1) were considered “Meets Discharge Standard”, “Indeterminate” and “Exceeds Discharge Standard”, respectively (ISO 3725). The method adjusts the overall observed agreement (PO) to the expected agreement (PE), which is the agreement due to chance. A value of  $\alpha = 0.05$  was used (95 % confidence intervals). The interpretation of  $\kappa$  from “no agreement,  $\kappa \leq 0$ ” in steps to “almost perfect agreement,  $0.81 \leq \kappa \leq 1.00$ ” was according to McHugh (2012). The number of BQUA ATP data was too low for statistical analysis. For the BallastWISE three compliance options were taken into account:  $<10$  c/mL (IMO),  $<13$  c/mL and  $\leq 30$  c/mL (Table 1).

## 3. Results

A total of 12 experiments were performed in August and September 2021. Experiments 1 to 4 were preliminary and used to set up the test rig and the analytical techniques. In experiment 8, the target organisms had an excessively low concentration and a poor condition to obtain conclusive results. Due to practical constraints, not all techniques could be used in all experiments.

### 3.1. Challenge water

The challenge water conditions were variable (Table 2). On average, the challenge water quality in terms of salinity and organism concentration and diversity complied with IMO and USCG rules for BWMS type approval testing in high salinity water. The average 10–50  $\mu\text{m}$  organism concentration ranged from 293 to 5150  $\text{cell mL}^{-1}$  ( $n = 7$ ) with an average of ca. 2000  $\text{cell mL}^{-1}$ . The 10–50  $\mu\text{m}$  organism diversity was not quantified in detail but was considered adequate due to the occurrence of centric and pennate diatoms, auto- and heterotrophic (dino)-flagellates, ciliates and occasionally cryptophytes, i.e.,  $>3$  phyla and  $>5$  species as required in type approval testing. Other chemical variables such as DOC, POC and TSS concentrations were not analyzed but based on past Control Union analyses of NIOZ harbor water these were below the IMO challenge water requirements. UVT was not manipulated and was 84 % on average.

### 3.2. Results as function of UV-dose

In general, the organism concentrations declined rapidly at low UV-doses ( $<200$   $\text{mJ cm}^{-2}$ ) to reach a more steady-state at higher doses (Fig. 3). Compared to the FM method, the MPN method yielded more D-2 compliant results (Fig. 3b). Details on MPN analyses and organism concentration results are in Annex 8.1. On the other hand, the Ballast-Check 2 and WALZ-PAM indicated compliance in (nearly) all UV-treated samples. The samples that were untreated by UV were all D-2 non-compliant according to both detailed and three out of four indicative methods. In a fourth method, BQUA ATP, one sample with 1785 organisms  $\text{mL}^{-1}$  would have been classified as compliant (1096  $\text{pg ATP mL}^{-1}$ ) and the other classified as non-compliant at 2480 organisms  $\text{mL}^{-1}$  (5264  $\text{pg ATP mL}^{-1}$ ) using the manufacturers' old limit of 1500  $\text{pg ATP mL}^{-1}$ . By using the 90 % reduced ‘new’ limit (150  $\text{pg ATP mL}^{-1}$ ), the two untreated samples were classified as D-2 non-compliant and only at UV-doses  $>200$   $\text{mJ cm}^{-2}$  samples became compliant (Fig. 3d).

### 3.3. MPN checks

In the MPN method three checks were made. In the first, microscopic analysis of the blank tubes, as expected no growth was observed. In the second check, incubating the tubes for an extra 7 days (21 incubation days in total), the calculated organism concentrations in the treated samples systematically increased (Fig. 4). Using the regression equation of Fig. 4, the concentration in the 1 to 100 organism/mL range increased by 142 %, i.e., more than doubled. In the third check, microscopic analyses showed a number of interesting observations (Table 3). In both the untreated control and the UV-treated sample tubes, and in both the  $<10$   $\mu\text{m}$  and  $>10$   $\mu\text{m}$  size fractions, pennate and centric diatoms were observed. Where the  $>10$   $\mu\text{m}$  size fraction only contained diatoms, the  $<10$   $\mu\text{m}$  fraction was much more diverse. Several  $>10$   $\mu\text{m}$  diatoms could be identified, but this was less simple for the phyla in the  $<10$   $\mu\text{m}$  fraction due to their small size and the use of the optically poor

**Table 2**

Challenge water conditions in August and September 2021. The organism concentration was determined after FM-staining and microscope counting.

Experiment	Temperature (°C)	Salinity (–)	UV (254 nm) transmittance (%)	10–50 $\mu\text{m}$ organisms ( $\text{mL}^{-1}$ )
5	26.1	20.1	78.6	920
6	27.5	20.2	89.7	2235
7	25.5	19.4	84.7	5150
9	19.0	23.0	84.4	293
10	17.6	23.0	86.1	1785
11	20.8	24.6	84.9	1930
12	17.0	27.0	80.9	2480
Average	21.9	22.5	84.2	2113



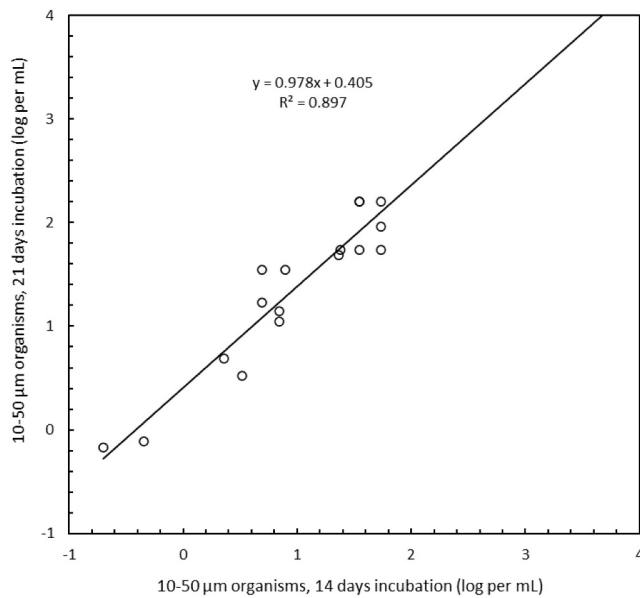


Fig. 4. Linear regression of log-transformed calculated MPN organism concentrations after 14 and 21 days of incubation.

WALZ-PAM (Fig. 5d).

### 3.5. Ranking compliance

In all untreated control samples, with the exception of ATP with the original BQUA non-compliance limit of 500 pg mL<sup>-1</sup>, the results were 100 % non-compliant (Table 4). Reducing the BQUA limit 10-fold (50 pg mL<sup>-1</sup>) also resulted in 100 % non-compliance in untreated control samples for this ATP method. In UV-treated samples, the lowest compliances are reached with the FM and BallastWISE techniques; the highest with the WALZ-PAM and the Ballast-Check 2.

### 3.6. Comparisons

Comparing MPN and the CMD techniques with FM using Cohen's κ, showed the highest, though moderate, agreement with the BallastWISE when the compliance limit was set at 9 c/mL (Table 5). At the manufacturer's compliance limits of 12 and 30 c/mL, the BallastWISE only scored a fair to slight/none agreement. The MPN scored a fair agreement at both incubation periods, but the Ballast-Check 2 and the WALZ-PAM had the least agreement with the FM method. Cohen's κ was not significant in all comparisons due to the unbalanced data matrix. In evaluating CMDs according to the ISO 3725 method (ISO, 2023), the samples used in the comparison with the reference method should be balanced in quantity and concentrations relative to the Discharge Standard.

## 4. Discussion

### 4.1. General

The aim of this study was to compare two IMO detailed techniques: FM and MPN with four indicative techniques: BallastWISE, Ballast-Check 2, Walz-PAM and BQUA ATP. IMO allows both detailed techniques, but the USCG only allows the FM method. Where the FM method is based on direct microscope observations and counts of the 10–50 µm organisms, comparable to the >50 µm motile organism microscope analysis that is prescribed by both IMO and USCG, the MPN method is based on the indirect measurement of autotrophic organisms by their chlorophyll fluorescence. For practical reasons, such as long and varying generation times, there is no MPN viability method for the >50 µm

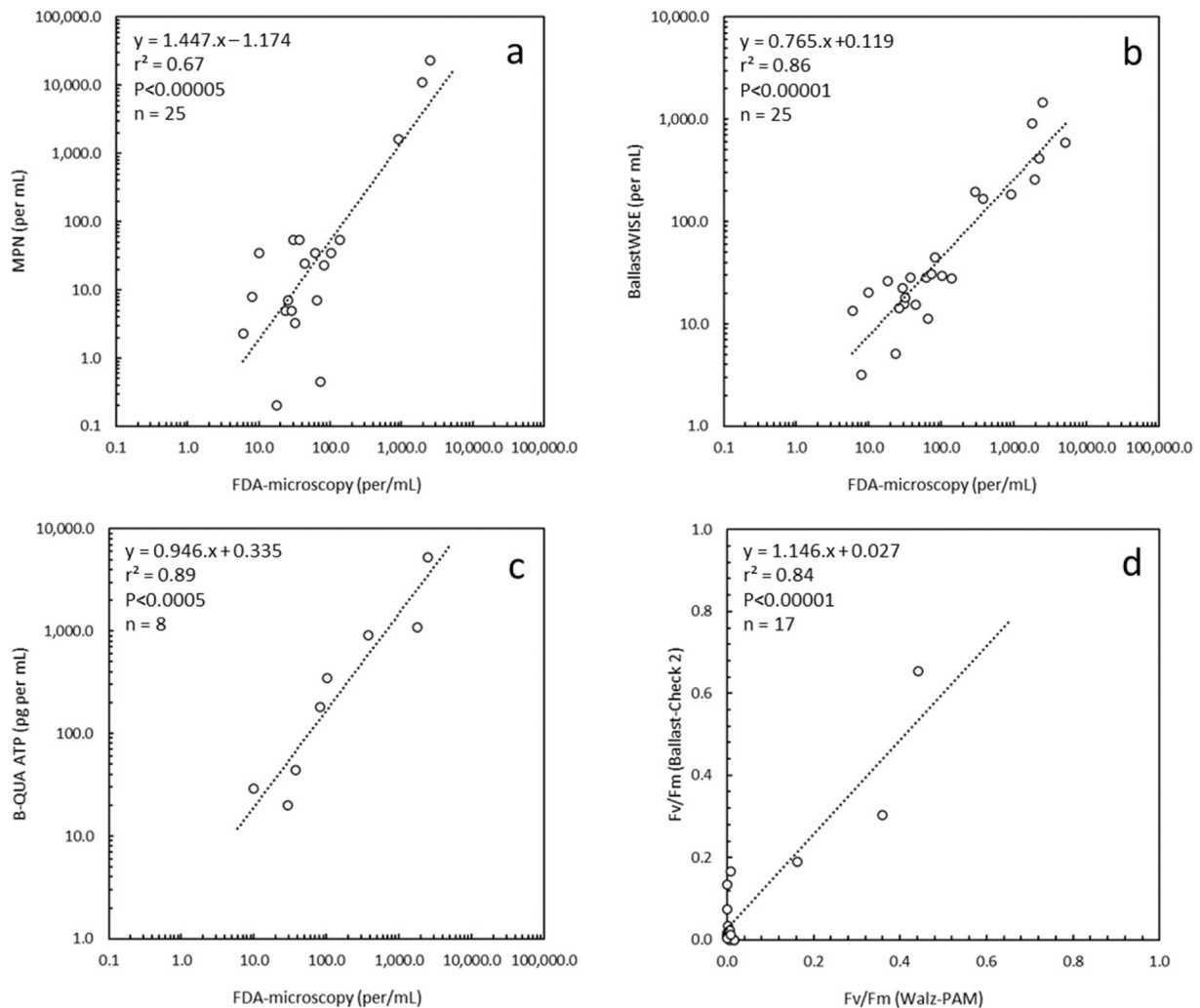
Table 3

Qualitative microscopic check of the MPN tube contents. Taxa: the number of taxa observed, including unidentified organisms. Because some taxa contain multiple species, the total number of taxa is a minimum. The relative frequency is the number of taxa divided by the number of tubes examined. Taxa corresponding in both size classes, but of different size, are in corresponding rows.

	>10 µm species		<10 µm species	
	Control	UV-treated	Control	UV-treated
<i>Asterionellopsis glacialis</i>			Centric diatoms	Centric diatoms
Centric diatoms		Centric diatoms (15 and 30 µm)		
<i>Chaetoceros</i> sp.			<i>Chaetoceros</i> sp.	<i>Chaetoceros</i> sp.
<i>Ditylum</i> sp.			Cyanobacteria, filamentous	Cyanobacteria, filamentous
<i>Leptocylindrus danicus</i>				cf. <i>Hemiselmis</i> sp.
<i>Lauderia</i> sp.			<i>Leptocylindrus minimus</i>	<i>Leptocylindrus minimus</i>
		<i>Navicula</i> sp.		<i>Navicula</i> spp.
		<i>Odontella</i> sp.		<i>Nitzschia</i> sp. (2 µm)
Pennate diatoms		Pennate diatoms	Pennate diatoms	Pennate diatoms
<i>Thalassiosira</i> sp.			<i>Phaeocystis</i> sp.	<i>Phaeocystis</i> sp.
				Unidentified cell (motile, 5 µm)
			Unidentified cell (non-motile, 2 µm)	Unidentified cell (non-motile, 2 µm)
				Unidentified cell (non-motile, 4 µm)
			Unidentified cell (non-motile, 5 µm)	Unidentified cell (non-motile, 5 µm)
				Unidentified cell (non-motile, 5–10 µm)
				Unidentified cell (non-motile, 8 µm)
	Unidentified diatoms		Unidentified diatoms	Unidentified diatoms
			Unidentified flagellates	Unidentified flagellates
Taxa	≥9	≥4	≥10	≥17
Relative frequency	0.27	0.05	0.30	0.21
Comment 1	All diatoms	All diatoms	Pro- and eukaryotes	Pro- and eukaryotes
Comment 2	Centric and pennate diatoms	Centric and pennate diatoms	Centric and pennate diatoms	Centric and pennate diatoms
Comment 3	See <10 µm	See <10 µm	Filamentous cyanobacteria	Filamentous cyanobacteria

organisms. This means that there are two analysis techniques, microscopy and MPN, for the 10–50 µm size fraction but only microscopy for the >50 µm fraction. Because the two detailed methods used in this study gave quite disparate results, these will be discussed first before reviewing the CMD data.

The two IMO-allowed 10–50 µm size fraction techniques have a different approach. The microscope method is based on measuring individual vitality, by green fluorescence (enzyme activity) or motility. On the other hand, MPN is based on measuring viability, the ability of 10–50 µm autotrophic organisms to multiply in test tubes and to pass the



**Fig. 5.** Linear regressions of MPN (a), BallastWISE (b), B-QUA ATP (c) on FM (Fluorescence Microscopy), and  $F_v/F_m$  of the Ballast-Check 2 on  $F_v/F_m$  of the WALZ-PAM (d). In a, b and c, the data were log-transformed. MPN data of  $\infty$  were omitted. The general equation (with  $y$  = dependent variable,  $x$  = independent variable,  $a$  = slope,  $b$  = intercept) is:  $y = ax + b$ , with  $n$  = number of data.

**Table 4**

Percentage of samples that were compliant or non-compliant in the control and UV-treated samples based on the compliance limits in Table 1. The data was ranked as % increasing compliance in UV-treated samples. The ATP limits are the original B-QUA limits as well as the reduced one (10 % of the original).  $n$  = number of samples.

	Limit for compliance ( $\leq$ )	Control			UV- treated		
		Compliant	Non-compliant	n	Compliant	Non-compliant	n
FM	9 mL <sup>-1</sup>	0 %	100 %	7	11 %	89 %	18
BallastWISE	12 mL <sup>-1</sup>	0 %	100 %	7	17 %	83 %	24
MPN (21 d)	9 mL <sup>-1</sup>	0 %	100 %	7	22 %	78 %	18
MPN (14 d)	9 mL <sup>-1</sup>	0 %	100 %	7	50 %	50 %	18
B-QUA ATP	50 pg mL <sup>-1</sup> (new)	0 %	100 %	2	50 %	50 %	6
B-QUA ATP	500 pg mL <sup>-1</sup> (old)	50 %	50 %	2	83 %	17 %	6
WALZ-PAM	$F_v/F_m = 0.09$	0 %	100 %	5	93 %	7 %	15
Ballast-Check 2	9 mL <sup>-1</sup> or $> 9$ mL <sup>-1</sup> and $F_v/F_m < 0.025$	0 %	100 %	7	100 %	0 %	24

chlorophyll fluorescence limit of detection. Here, vitality is defined as “manifesting life,” and viability as “the capability to successfully reproduce” Hence, cells can be (i) vital and viable, (ii) vital with reduced viability (e.g., as an effect of a toxic substance or UV-irradiance), (iii) vital but with no viability (e.g., after a nutrient has been depleted), and (iv) nonvital and nonviable (dead) (Peperzak and Brussaard, 2011). The quintessence of ballast water treatment is to reduce the risk of introducing viable organisms in a new environment. Therefore, MPN is considered a more relevant (viability) technique to verify ballast water

treatment than FM (vitality). However, guidance from IMO (IMO, 2022) accepts all analytical techniques if they are based on “essential characteristics of life” such as structural integrity (FM), metabolism (FM), reproduction (MPN), motility (FM, MPN,  $>50 \mu\text{m}$  organisms), or response to stimuli ( $>50 \mu\text{m}$  organisms) and accepts both detailed 10–50  $\mu\text{m}$  techniques.

Both FM and MPN have inadequacies. FM may underestimate the concentration of vital organisms in a number of cases because several species of living plankton do not stain with FDA/CMFDA (Hansen et al.,

**Table 5**

Ranking of instrument performance and MPN according to Cohen's  $\kappa$  (ISO 3725) by comparison with FM. BallastWISE (9), (12) and (30) are compliance limits at 9 (IMO), 12 and 30 (Table 1) organisms per mL respectively. The MPN data from 14- and 21-days incubation were both evaluated.

Instrument	Cohen's $\kappa$	Interpretation	Variance	95 % confidence interval	Significantly >0?
BallastWISE (9)	0.46	Moderate	0.37	0.61	No
BallastWISE (12)	0.34	Fair	0.36	0.59	No
MPN (14 d)	0.27	Fair	0.23	0.39	No
MPN (21 d)	0.25	Fair	0.34	0.56	No
BallastWISE (30)	0.10	None to slight agreement	0.16	0.26	No
Ballast-Check 2	0.07	None to slight agreement	0.14	0.23	No
WALZ-PAM	0.06	None to slight agreement	0.18	0.30	No

2024, 2023; Peperzak and Brussaard, 2011; Steinberg et al., 2011). However, a major disadvantage of FM in UV-treatment, potentially leading to an overestimation of “living” organisms, is that green-fluorescent 10–50  $\mu\text{m}$  organisms are classified as living, although their DNA may be damaged to the extent they are non-viable; for this reason, the UV-treatment efficacy is noticeably greater when measured with MPN with respect to FM (Lundgreen et al., 2018). Notably, this disadvantage is generally not considered in the microscopic analysis of >50  $\mu\text{m}$  size fraction where motile organisms that are classified as living, may also have been rendered non-viable due to UV-damage and, therefore, harmless to the environment.

On the other hand, major MPN disadvantages are that 99 % of the phytoplankton species has not been shown to grow in tubes leading to the major probability that this technique will underestimate the relatively large 10–50  $\mu\text{m}$  size group (Peperzak, 2024). Furthermore, MPN delivers a total size spectrum (including <10  $\mu\text{m}$  and theoretically >50  $\mu\text{m}$ ) overall viable organism concentration. This adds to the uncertainty in the estimation of the 10–50  $\mu\text{m}$  organism concentration because it means that MPN may also overestimate the overall organism concentration due to the more abundant <10  $\mu\text{m}$  organisms that are well known to thrive in small tubes. In this study the small (<10  $\mu\text{m}$ ) species appeared to be more frequent than large species in the MPN positive tubes, e.g. Table 3. In fact, the MPN method was developed to culture, identify and isolate relatively small species (Andersen and Throndsen, 2003; Throndsen, 1978).

The viable organism concentration end result of an MPN analyses is therefore an unknown balance of slow- or non-growing large organisms and fast-growing small organisms (Throndsen, 1978) (Peperzak, 2024). Furthermore, there is a misconception that the MPN outcome is the cell concentration of the originally most abundant species and that, according to Cullen (2019) competition in an assemblage of species is irrelevant as long as the winner is detectable. Rather, the MPN outcome is the total sum of all species present and even if the “winner” is a >10  $\mu\text{m}$  species, its concentration needs to be corrected for any <10  $\mu\text{m}$  autotrophs in the species mixture (Peperzak, 2024). At present, there is no independent MPN validation of ambient water with only natural autotrophs with a correction for the <10  $\mu\text{m}$  autotrophs available. In the present study, the overall viable organism concentration in the UV-treated samples may indicate compliance, but it is unknown how much is due to the inherent inability of 10–50  $\mu\text{m}$  organisms to grow in the small MPN tubes. In other words, both FM and MPN have their advantages and limitations. In conclusion, both the FM method and the MPN method suffer from inadequacies that can cause the incorrect estimation of “living” 10–50  $\mu\text{m}$  organisms (IMO, 2016).

A drawback that three of the indicative methods have in common is their use of biomass proxies. They solely rely on (active) chlorophyll fluorescence (Ballast-Check 2, WALZ-PAM) or on ATP concentration (BQUA) to calculate organism concentrations. The chlorophyll and ATP contents of organisms scale with three-dimensional size which makes it difficult to correlate such proxies with organism concentrations when these may have sizes anywhere between 10 and 50  $\mu\text{m}$  in linear minimum dimension. For instance, the volume of a spherical organism of 10  $\mu\text{m}$  diameter is 524  $\mu\text{m}^3$  and that of a 50  $\mu\text{m}$  diameter is 65,450  $\mu\text{m}^3$ , which is a factor 125 difference. Because the ATP content of aquatic

organisms is constant and relatively high (Bochdansky et al., 2021), the BQUA (non)-compliance limits can be reduced by a factor of 10 (Peperzak, 2023). In this context, past studies (Drillet et al., 2023; Peperzak et al., 2024) consider that ATP limit is over-precautionary. The BallastWISE is the only CMD that directly observes and counts organisms of the required size class, coupled with a fluorescence analysis of each individual cells separately in its field of view.

#### 4.2. Comparison of CMD techniques

The present experiment was not an IMO or USCG standard type approval, commissioning or ISO-validation test but was meant to compare the results of different CMDs in measuring compliance in natural water with ambient phytoplankton treated with UV. If the BQUA compliance limit was reduced from 500 to 50 pg ATP mL<sup>-1</sup>, all four CMDs as well as the two detailed methods indicated non-compliance of the untreated control samples. In the UV-treated samples, 89 % and 83 % non-compliance was reached by FM and the BallastWISE (both based on counting individual cells), shortly followed by MPN with 21 days incubation (78 %). This longer MPN incubation resulted in an extended growth period for the viable phytoplankton resulting in a higher percentage of non-compliant samples than in the standard 14 days MPN (50 %). The BQUA ATP results were also 50 % with the reduced (50 pg ATP mL<sup>-1</sup>) compliance limit. However, the two active chlorophyll fluorescence methods, Ballast-Check 2 and WALZ-PAM, showed compliance in nearly all samples. The different compliance thresholds, in Ballast-Check 2  $F_v/F_m = 0.5$  and  $F_v/F_m = 0.1$  (WALZ-PAM) did not considerably change the number of compliant samples, meaning that the Ballast-Check 2 algorithm for the conversion of chlorophyll fluorescence to organism concentration may need revision. The WALZ-PAM threshold that was used by NIOZ as an indicative method appeared too high for the UV-treated samples.

#### 4.3. Numeric analysis and compliance or non-compliance

In general, there was a good regression correlation between the FM with the MPN, BallastWISE and BQUA ATP methods. The FM data will be used in a comparison with the indicative methods. The MPN data are not taken into account due to the presence of <10  $\mu\text{m}$  organisms.

The excellent correlation between the FM method and the BQUA ATP enables a verification of the rationale to lower the original manufacturer's 1500 pg ATP mL<sup>-1</sup> non-compliance limit to 150 pg ATP mL<sup>-1</sup>. Using the slope ( $a = 0.946$ ) of the regression equation in Fig. 5c, which is the log ATP content of the log 10–50  $\mu\text{m}$  organism concentration, the ATP content of 1 organism (using  $\log(100 \text{ organisms}) = 2$ ) is calculated to be  $(10^{2 \cdot 0.946}/100 =)$  0.8 pg ATP. This means non-compliance at 10 cells mL<sup>-1</sup> in our tests is reached at 8 pg ATP mL<sup>-1</sup>. Assuming the average 0.6 mM ATP concentration in aquatic eukaryotic microorganisms (Bochdansky et al., 2021) and using a size-ATP model, this corresponds to a spherical organism with a sed = 17  $\mu\text{m}$  (Peperzak, 2023). On the other hand, assuming the maximum of 2 mM ATP concentration in aquatic eukaryotic microorganisms (Bochdansky et al., 2021), the calculated sed = 12  $\mu\text{m}$ . This size range, 12 to 17  $\mu\text{m}$ , corresponds to the Welschmeyer and Kuo (2016) model for the average size of coastal

phytoplankton (15  $\mu\text{m}$  sed) and partly validates the basic Peperzak (2023) size-ATP model. The calculation also implies that the worst-case scenario of  $\times 3.3$  ATP organism concentration (2 mM instead of 0.6 mM) and a supposedly  $3\times$  increase in extraction efficiency by the lysis method (Peperzak, 2023), resulted in a  $10\times$  overestimation of the ATP content. This means that the BQUA's non-compliance limit ( $3\times$  the compliance limit), using the  $10\times$  reduced new compliance limit of 50 pg ATP  $\text{mL}^{-1}$  could even be set to  $3 \times 5 = 15$  pg ATP  $\text{mL}^{-1}$ , instead of BQUA's 1500 pg ATP  $\text{mL}^{-1}$ . In that case, all ATP results in our tests would be non-compliant and in full agreement with the FM results. In the meantime, the ATP-model has been validated with both marine and freshwater data (Peperzak, 2024), suggesting that the BQUA (non)-compliance limits are indeed too high.

Guidance from IMO allows the use of two detailed methods in BWMS type approval testing, FM and MPN, that are based on different concepts: vitality-viability, direct-indirect observation, 10–50  $\mu\text{m}$ - all sizes, fast-slow analysis time, and that both have merits and deficiencies. Unfortunately, IMO has failed to undertake a rigorous validation between FM and MPN as it has prescribed for CMDs (IMO, 2023; ISO, 2023). Following a period of BWMS type approval tests at ballast water test facilities involving land-based and shipboard tests, the commissioning of installed BWMSs started in 2022. However, IMO again failed to prescribe that only validated CMDs should be allowed in commission testing. This means that at present (2024) the majority of ship-installed BWMSs has been commissioned with a variety of indirect and non-validated techniques.

In IMO type approval testing a BWMS manufacturer could deliberately opt for the MPN technique, which would probably lead to lower organism results compared to FM. In the next phase of implementation of the Ballast Water Convention, to test installation and working of a type-approved BWMS, a ship owner could opt for a CMD with an indirect technique and its accompanying unvalidated (non)-compliance limits, which would minimize the risk of failure to comply. Starting on September 8, 2024, the task for Port State Control is to choose a strategy for compliance monitoring in routine ballast water discharges. If this monitoring is based on indirect methods it would make sense to employ a technique that measures both the 10–50  $\mu\text{m}$  and  $> 50$   $\mu\text{m}$  size fractions, excluding those CMDs based on chlorophyll fluorescence only. However, because none of the CMDs have yet been formally validated, the risk is that, at present, non-compliance is either rarely or unduly concluded.

The environmentally safest option for PSC would be to use either detailed techniques or (once available) validated CMDs with results directly comparable to the D-2 standard (Loiacono et al., 2024). The major disadvantage of detailed techniques is that it is more complicated, expensive, time-consuming, and not likely to be executed by PSC officers. This choice is therefore likely to be made in cases where document inspection, visual inspection of ballast water treatment and inspection of the BWMS yield an unsatisfactory outcome. When executed by ISO 17025-accredited laboratories a verdict of non-compliant discharge would also provide PSC with a more solid legal basis for handing out

penalties.

## 5. Conclusions

As already concluded by Allen (1919), there is no single best approach that provides quantitative estimates of phytoplankton concentrations, so in ecological studies a variety of methods would have to be used to obtain a reliable estimate of cell abundance. Here, we showed that there a fundamental and large differences between two detailed methods, FM and MPN, and two biomass-based indicative methods (Ballast-Check 2 and BQUA ATP). For the other indicative methods, the WALZ-PAM merely indicated general fitness of the complete autotrophic community. The BallastWISE, although still considered an indicative method, is the only CMD providing direct counts of living organisms, with results directly comparable to the detailed FM method.

## CRedit authorship contribution statement

**Leonardo Romero-Martínez:** Writing – review & editing, Writing – original draft, Visualization, Validation, Methodology, Investigation, Formal analysis, Conceptualization. **Cees van Slooten:** Writing – review & editing, Methodology, Investigation, Conceptualization. **Michiel van Harten:** Methodology, Investigation, Conceptualization. **Enrique Nebot:** Project administration, Funding acquisition. **Louis Peperzak:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Conceptualization.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Acknowledgements

This work was supported by the Project RETOLASTRE (FEDER-UCA18-105151), funded by the Junta de Andalucía and the European Union. The research stay by L. Romero-Martínez at NIOZ was funded by the Plan Propio UCA - 2021-2022 (EST2021-017). The author L. Romero-Martínez is hired in the frame of the Programme Qualifica 2021 (Ref: QUAL21\_019\_UCA, Junta de Andalucía), funded by the “Plan Andaluz de Investigación, Desarrollo e Innovación (PAIDI 2020)” of the Regional Government of Andalusia.

Authors thank C. Hedberg, Alfa Laval (UV reactor), Robert Twijnstra, NIOZ (challenge water preparation), Anna Noordeloos and MMB staff, NIOZ (practical assistance), Dr. Erik Zettler, NIOZ (fluorimeter), Dr. R. Forster, University of Hull, UK (assistance in  $F_v/F_m$  interpretation) and Dr. H. Bolhuis, NIOZ (sequencing three cyanobacteria isolates) for their support in this research.

## Appendix A. Table A1. MPN: detailed analysis and results

The table represents the dilution performed for every sample and the MPN results, including the 95 % confidence interval (CI). It is important to keep in mind that MPN data include organisms  $< 10$   $\mu\text{m}$ . The 95 % CI depends on both the number of positive tubes and the dilutions performed. In the case that all tubes are positive, 95 % CI depends on the dilutions performed. The results depend on the initial cell concentration and the inactivating effect of the UV dose. In natural water samples with unknown organism concentrations, it is difficult to determine the combination of dilutions to obtain a numerical result. In experiments 11 and 12, the number of dilutions was increased as an effort to obtain numerical results. Increasing the number of dilutions implies a great amount of extra work, materials and space in the climate chamber.

Exp.	UV dose (mJ cm <sup>-2</sup> )	Dilutions performed	14 d		21 d	
			MPN (cells mL <sup>-1</sup> )	95 % CI	MPN (cells mL <sup>-1</sup> )	95 % CI
5	0	1, 2, 3	1600	530–4900	1600	530–4900
	290	0, 1, 2	0.5	0.11–1.9	0.8	0.24–2.5
	580	0, 1, 2	0.2	0.027–1.5	0.7	0.21–2.2
6	0	2, 3, 4	∞ (all tubes +)	6500 - ∞	∞ (all tubes +)	6500 - ∞
	130	0, 1, 2	7.0	2.3–22	14	5.5–36
	260	0, 1, 2	4.9	1.5–16	17	6.4–46
7	0	2, 3, 4	∞ (all tubes +)	6500 - ∞	∞ (all tubes +)	6500 - ∞
	110	0, 1, 2	54	16–190	160	53–490
	220	0, 1, 2	54	16–190	92	28–300
9	0	1, 2, 3	∞ (all tubes +)	650 - ∞	∞ (all tubes +)	650 - ∞
	90	0, 1, 2	24	7.8–74	54	16–190
	180	0, 1, 2	7.9	2.5–25	35	11–110
10	0	2, 3, 4	∞ (all tubes +)	6500 - ∞	∞ (all tubes +)	6500 - ∞
	55	0, 1, 2	∞ (all tubes +)	65 - ∞	∞ (all tubes +)	65 - ∞
	110	0, 1, 2	35	11–110	54	16–190
11	0	2, 3, 4, 5	11,000	3900–31,000	11,000	3900–31,000
	175	0, 1, 2	35	11–110	160	53–490
	350	0, 1, 2	3.3	1.1–10	3.3	1.1–10
	525	0, 1, 2	7.0	2.3–22	11	3.9–31
	700	0, 1, 2	2.3	0.76–7	4.9	1.5–16
12	0	3, 4, 5, 6	23,000	7600–70,000	41,000	18,000–93,000
	160	0, 1, 2, 3	23	7.6–70	49	15–160
	320	0, 1, 2	54	16–190	54	16–190
	480	0, 1, 2	4.9	1.5–16	35	11–110
	640	0, 1, 2	35	11–110	160	53–490

## Appendix B. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.marpolbul.2024.117193>.

## Data availability

Data will be made available on request.

## References

- Allen, E.J., 1919. A contribution to the quantitative study of plankton. *J. Mar. Biol. Assoc. U. K.* 12, 1–8. <https://doi.org/10.1017/S0025315400059889>.
- Andersen, P., Thronsen, J., 2003. Estimating cell numbers. In: *Man. Harmful Mar. Microalgae*. Paris UNESCO, pp. 99–129. <https://doi.org/10.25607/OBP-1370>.
- Blackburn, N., Haecky, P., Jurgensone, I., Griniene, E., Brugel, S., Andersson, A., Carstensen, J., 2022. The use of an automated organism tracking microscope in mesocosm experiments. *Limnol. Oceanogr. Methods* 20, 768–780. <https://doi.org/10.1002/lom3.10521>.
- Blatchley III, E.R., Cullen, J.J., Petri, B., Bircher, K., Welschmeyer, N., 2018. The biological basis for ballast water performance standards: “viable/non-viable” or “live/dead”? *Environ. Sci. Technol.* 52, 8075–8086. <https://doi.org/10.1021/acs.est.8b00341>.
- Bochdanský, A.B., Stouffer, A.N., Washington, N.J.N., 2021. Adenosine triphosphate (ATP) as a metric of microbial biomass in aquatic systems: new simplified protocols, laboratory validation, and a reflection on data from the literature. *Limnol. Oceanogr. Methods* 19, 115–131. <https://doi.org/10.1002/lom3.10409>.
- Cullen, J.J., 2019. The best available science describing type-approval testing methods and protocols for ballast water management systems that render nonviable organisms in ballast water. *Zenodo*. <https://doi.org/10.5281/zenodo.2656597>. Citation.
- Cullen, J.J., MacIntyre, H.L., 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. <https://doi.org/10.1007/s10811-015-0601-x>.
- de Castro, M.C.T., Veldhuis, M.J.W., 2019. Temporal changes in phytoplankton biomass and cellular properties; implications for the IMO ballast water convention. *Environ. Technol.* 40, 1455–1466. <https://doi.org/10.1080/09593330.2017.1423117>.
- Drake, L.A., Wier, T.P., Parson, E.W.J., Grant, J.F., First, M.R., 2016. Review of a request for approval of an alternative method for ballast water testing (46 CFR 162.060–10 (b)(1)): Trojan Marinex’s method for assessing organisms ≥ 10 µm and < 50 µm. *NRL Lett. Rep.* 6130/1622. *NRL Lett. Rep.* 6130 (6130–1622).
- Drillet, G., Gianoli, C., Gang, L., Zacharopoulou, A., Schneider, G., Stehouwer, P., Bonamin, V., Goldring, R., Drake, L.A., 2023. Improvement in compliance of ships’ ballast water discharges during commissioning tests. *Mar. Pollut. Bull.* 191, 114911. <https://doi.org/10.1016/j.marpolbul.2023.114911>.
- First, M.R., Drake, L.A., 2013. Approaches for determining the effects of UV radiation on microorganisms in ballast water. *Manage. Biol. Invasions* 4, 87–99.
- Gerhard, W.A., Lundgreen, K., Drillet, G., Baumler, R., Holbech, H., Gunsch, C.K., 2019. Installation and use of ballast water treatment systems – implications for compliance and enforcement. *Ocean Coast. Manag.* 181, 104907. <https://doi.org/10.1016/j.ocecoaman.2019.104907>.
- Hansen, G., Petersen, G., Maguet, R., Sim, E., Madsen, T., 2023. Inherent lack of CMFDA/FDA staining in certain algae and its implication for ballast water testing. *Mar. Pollut. Bull.* 194, 115312. <https://doi.org/10.1016/j.marpolbul.2023.115312>.
- Hansen, G., Tolstrup, L., Petersen, G., Maguet, R., Madsen, T., 2024. Viability assessed with the most probable number dilution culture method after chemical treatment of ballast water reveals the presence of false negatives from an approved vital stain method. *Mar. Pollut. Bull.* 205, 116586. <https://doi.org/10.1016/j.marpolbul.2024.116586>.
- IMO, 2004. International Convention for the Control and Management of Ships’ Ballast Water and Sediments. *Int. Marit. Organ. London, UK*, 13 Febr. 2004.
- IMO, 2016. 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. In: *Orig. Denmark Norway*. IMO PPR 4/7, 12 Oct. 2016, IMO PPR 4/7, 12 Oct. 2016.
- IMO, 2018. RESOLUTION MEPC.300(72). Code for Approval of Ballast Water Management Systems (BWMS Code).
- IMO, 2019. PPR 7/INF.10. Revised Guidance on Methodologies that may be Used for Enumerating Viable Organisms. Description of the MPN Serial Dilution Culture Method + Motility Method.
- IMO, 2020a. BWM.2/Circ.70/Rev.1. 2020 Guidance for the Commissioning Testing of Ballast Water Management Systems.
- IMO, 2020b. MEPC 174(58). Latest Information of Approval of Ballast Water Management System.
- IMO, 2022. BWM.2/Circ.61. Guidance on Methodologies that may be used for Enumerating Viable Organisms for Type Approval of Ballast Water Management Systems.
- IMO, 2023. BWM.2/Circ.78. Protocol for the Verification of Ballast Water Compliance Monitoring Devices.
- ISO, 2023. Ships and Marine Technology - Aquatic Nuisance Species - Methods for Evaluating the Performance of Compliance Monitoring Devices for Ballast Water Discharges. ISO 37252023, IDT 40.
- Jarvis, B., Wilrich, C., Wilrich, P.T., 2010. Reconsideration of the derivation of most probable numbers, their standard deviations, confidence bounds and rarity values. *J. Appl. Microbiol.* 109, 1660–1667. <https://doi.org/10.1111/j.1365-2672.2010.04792.x>.
- Khandeparker, L., Anil, A.C., 2017. Global concerns of ship’s ballast water mediated translocation of bacteria. In: Naik, M., Dubey, S. (Eds.), *Mar. Pollut. Microb.*

- Remediat. Springer, Singapore, pp. 255–262. <https://doi.org/10.1007/978-981-10-1044-6>.
- Li, S., Tao, Y., Zhan, X.M., Dao, G.H., Hu, H.Y., 2020. UV-C irradiation for harmful algal blooms control: a literature review on effectiveness, mechanisms, influencing factors and facilities. *Sci. Total Environ.* 723, 137986. <https://doi.org/10.1016/j.scitotenv.2020.137986>.
- Lo Curto, A., Stehouwer, P., Gianoli, C., Schneider, G., Raymond, M., Bonamin, V., 2018. Ballast water compliance monitoring: a new application for ATP. *J. Sea Res.* 133, 124–133. <https://doi.org/10.1016/j.seares.2017.04.014>.
- Loiacono, S., Ceballos-Osuna, L., Scianni, C., Brown, C.W., 2024. Technical considerations for sampling ballast water to determine compliance with discharge performance standards. *Front. Mar. Sci.* 11, 1323989. <https://doi.org/10.3389/fmars.2024.1323989>.
- Lundgreen, K., Holbech, H., Pedersen, K.L., Petersen, G.I., Andreasen, R.R., George, C., Drillet, G., Andersen, M., 2018. UV fluences required for compliance with ballast water discharge standards using two approved methods for algal viability assessment. *Mar. Pollut. Bull.* 135, 1090–1100. <https://doi.org/10.1016/j.marpolbul.2018.08.043>.
- McHugh, M.L., 2012. Lessons in biostatistics interrater reliability: the kappa statistic. *Biochem. Med.* 22, 276–282.
- Outinen, O., Lehtiniemi, M., 2017. Literature Review for the Indicative Ballast Water Analysis Methods. *Trafi Res. Reports* 10/2017. Finnish Transp. Saf. Agency.
- Peperzak, L., 2023. The critical adenosine triphosphate (ATP) concentration in treated ballast water. *Mar. Pollut. Bull.* 187, 114506. <https://doi.org/10.1016/j.marpolbul.2022.114506>.
- Peperzak, L., 2024. The Serial Dilution Culture-Most Probable Number (MPN) Assay to Estimate Phytoplankton Concentrations in Ballast Water: Comments and Improvements. *Submitt. to Front. Mar. Sci.*
- Peperzak, L., Brussaard, C.P.D., 2011. Flow cytometric applicability of fluorescent vitality probes on phytoplankton. *J. Phycol.* 47, 692–702. <https://doi.org/10.1111/j.1529-8817.2011.00991.x>.
- Peperzak, L., Casas-Monroy, O., Bailey, S.A., 2024. Validation of an adenosine triphosphate (ATP) model for 10–50 µm plankton. *Mar. Pollut. Bull.* 200, 116066. <https://doi.org/10.1016/j.marpolbul.2024.116066>.
- Pimentel, D., Zuniga, R., Morrison, D., 2005. Update on the environmental and economic costs associated with alien-invasive species in the United States. *Ecol. Econ.* 52, 273–288. <https://doi.org/10.1016/j.ecolecon.2004.10.002>.
- Ruiz, G.M., Rawlings, T.K., Dobbs, F.C., Drake, L.A., Mullady, T., Huq, A., Colwell, R.R., 2000. Global spread of microorganisms by ships. *Nature* 408, 49–50. <https://doi.org/10.1038/35040695>.
- Sinha, R.P., Häder, D.P., 2002. UV-induced DNA damage and repair: a review. *Photochem. Photobiol. Sci.* 1, 225–236. <https://doi.org/10.1039/b201230h>.
- Stehouwer, P.P.V., 2016. Effects of Various Ballast Water Treatment Methods on the Survival of Phytoplankton and Bacteria [Thesis fully Intern. (DIV), Univ. Groningen]. *Univ. Groningen*.
- Stehouwer, P.P., Buma, A., Peperzak, L., 2015. A comparison of six different ballast water treatment systems based on UV radiation, electrochlorination and chlorine dioxide. *Environ. Technol.* 36, 2094–2104. <https://doi.org/10.1080/09593330.2015.1021858>.
- Steinberg, M.K., Lemieux, E.J., Drake, L.A., 2011. Determining the viability of marine protists using a combination of vital, fluorescent stains. *Mar. Biol.* 158, 1431–1437. <https://doi.org/10.1007/s00227-011-1640-8>.
- Thronsen, J., 1978. The dilution-culture method. *Phytoplankt. Man. (Sournia, A.) UNESCO Monogr. Oceanogr. Methodol.* 218–224.
- USCG, 2012. Standards for living organisms in Ships' ballast water discharged in U.S. waters. *Fed. Regist.* 77, 17254–17320.
- Welschmeyer, N.A., Kuo, J., 2016. Analysis of Adenosine Triphosphate (ATP) as a Rapid, Quantitative Compliance Test for ships' Ballast Water. *Moss Land. Mar, Lab.*