PERFORMANCE VERIFICATION STATEMENT
For the 10Cells (BBE Moldaenke)

TECHNOLOGY TYPE:	Ballast water compliance tools
APPLICATION:	Shipboard analysis of ballast water
PARAMETERS EVALUATED:	Response linearity, accuracy, and precision
TYPE OF EVALUATION:	Laboratory and field performance verifications
DATE OF EVALUATION:	Testing conducted from July to September 2015
EVALUATION PERSONNEL:	M.R. First, S.C. Riley, S.H. Robbins-Wamsley, V. Molina, T. Johengen, H. Purcell, G.J. Smith, E. Reavie, K. Carney, C.S. Moser, E.N. Buckley M.N. Tamburri, and L.A. Drake

TABLE OF CONTENTS

BACKGROUND AND OBJECTIVES

In an effort to mitigate the risk of transporting aquatic nuisance species, the United States Coast Guard (USCG) has finalized a rule limiting the concentrations of organisms in ships' ballast water discharged into US ports (US Coast Guard 2012). The specified concentrations are nearly identical (with the exception of not including limits for *Vibrio cholerae* in zooplankton samples) to those in the International Maritime Organization's (IMO) convention (IMO 2004). Further, the limits are consistent with those in the US Environmental Protection Agency's Vessel General Permit (VGP)—regulations on a suite of vessel operations, including the discharge of ballast water (US EPA 2013). In order to meet these limits, most ships will use a ballast water management system (BWMS). These systems incorporate a variety of technologies (including filtration, UV radiation, electrolytic chlorination, and deoxygenation) to ensure that the discharge water meets the specifications.

Determining concentrations of living organisms can require extensive effort and sensitive equipment, especially for sparse populations. For example, direct counts of living organisms ≥10 and <50 µm according to the method stipulated in the US Environmental Technology Verification (ETV) Program Protocol for land-based testing of BWMS requires (1) labeling organisms within a sample with a set of vital fluorophores and (2) tallying the organisms via epifluorescence microscopy (EPA 2010; Steinberg et al. 2011). Direct counts of living organisms yield concentrations comparable to the numerical standard. While this rigorous, complex, and time-consuming analysis is appropriate for verification testing of BWMS, it is typically not feasible to perform this analysis during routine shipboard inspections. Rather, simple, hand-held, field instruments ("compliance tools")—with the ability to rapidly assess that the ballast water *clearly* exceeds the discharge limits—will be of much greater value to the ship owner, the BWMS vendor, and the compliance officer. Compliance tools should immediately produce results that are reliable indicators of the concentrations of living organisms within a regulated size class and predict whether a sample meets or exceeds the discharge standard.

New or refined compliance tools require carefully considered test protocols for evaluating and verifying their performance. The overall goal of this *technology verification* was to evaluate the performance of potential compliance tools designed to rapidly assess ballast water discharge. The outputs of the compliance tools were compared to the standard, validated approach (i.e. epifluorescence microscopy; EPA 2010) used to quantify organisms ≥10 and <50 µm in size during verification testing of BWMS. The objectives outlined below support this goal:

- In a series of laboratory trials to be conducted at the Naval Research Laboratory in Key West, FL (NRL), determine **linearity**, **precision** and **accuracy** of the compliance tool with samples of algal monocultures over a range of concentrations, including concentrations below, equal to, and above the IMO and US discharge standard.

- Evaluate the relationship between numerical concentrations of living organisms ≥10 and <50 µm and the accuracy and precision of the instrument using ambient organisms

collected from natural waters at three various locations (Key West, Chesapeake Bay, and Lake Superior).

INSTRUMENT TECHNOLOGY TESTED

This report describes the test of the **BBE Moldaenke 10Cells** (hereafter, 10Cells). The instrument employs variable fluorescence fluorometry, an approach that measures chlorophyll *a* fluorescence at variable illumination intensities and intervals. These measurements are used to estimate concentrations of living organisms within an aliquot of water. As photosynthetic algae are abundant in the ≥10 and <50 µm size class, the instruments may provide a reasonable determination that a sample meets the discharge limit of 10 living organisms mL^{-1} in that size class. Upon completion of sample analysis, 10Cells displays the estimated cell concentration (mL^{-1}) based upon the fluorescence measurements and a conversion factor (the "cell factor"). Further details of the operation of the 10Cells are available in the test plan (**Appendix A**).

PERFORMANCE EVALUATION TEST PLAN

The test protocol for this performance verification was developed at a conference with NRL and the Alliance for Coastal Technologies (ACT) personnel, the participating instrument manufacturers, and a technical advisory committee. The verification of the instrument included both laboratory and field experiments: these tests are summarized briefly in this document and in detail in the test protocol. Experiments were designed to challenge the compliance tool by analyzing ranges of concentrations—spanning from zero to well above the discharge standard. Measurements reported by the instrument were compared to the results of the standard technique, described below. The critical comparison was the agreement on the disposition of the sample: if both the compliance tool and the microscope count indicate concentrations ≥10 mL^{-1}, the methods agree. Likewise, if both methods determine concentrations are <10 mL^{-1}, the methods agree.

Laboratory Experiments

Laboratory tests examined the agreement between cell concentrations measured via microscopy and the compliance tool using two cultured microalgae: *Tetraselmis marina* (cell dimensions: 9-15 µm) and *Prorocentrum micans* (25-50 µm). The organisms represented cell dimensions towards the extremes of the ≥10 and <50 µm size class. For the laboratory experiments with cultured algae, all living cells were counted, even though some individuals may have been slightly larger or smaller than the size limits. Samples with either *T. marina* or *P. micans* were prepared by diluting stock cultures with 0.22-µm filtered seawater (FSW) to yield concentrations of 0, 5, 10, 20, 50, and 100 mL^{-1}. Additionally, two samples were prepared to examine interferences from (1) dissolved and particulate materials and (2) disinfection byproducts (DBP). These samples contained 10 mL^{-1} of either *T. marina* or *P. micans*.

Field Experiments

Instrument performance was also tested in field experiments using ambient water samples collected from three locations representing a range of water temperatures, salinities, and community compositions: The Naval Research Laboratory (NRL; latitude 24.58°N; Longitude: 81.79°W) in Key West, FL represented offshore, high salinity, waters (temperature: 27°C; salinity: 36 psu). The Great Ships Initiative (GSI) in Superior, WI (46.71°N; 92.05°W) represented the Great Lakes (20°C; 0 psu). The Smithsonian Environmental Research Center (SERC; 38.89°N; 76.54°W) in Edgewater, MD, located on the Chesapeake Bay, represented estuarine waters (25°C; 9 psu). Samples with a mixed assemblage of ambient organisms were prepared by either diluting or concentrating natural water from the location: dilution was performed by mixing the sample with FSW (or at GSI, 0.22-µm filtered *lake* water, FLW). Cells were concentrated by screening water through a sieve with mesh netting to retain organisms ≥10 µm. Following these procedures, four samples were generated with different target concentrations:

- 0 mL^{-1}, the 0.22-µm filtered water to be used as a control or blank for fluorescence,
- $5 – 20 \text{ mL}^{-1}$, representing concentrations near the discharge standard (DS),
- $30 – 50 \text{ mL}^{-1}$, representing concentrations above the DS, and
- $≥50 \text{ mL}^{-1}$, representing concentrations well above the DS.

Determining Concentrations of Microalgae by Epifluorescence Microscopy

Organisms ≥10 and <50 µm were quantified using the approach in the Environmental Technology Verification (ETV) Program protocol (EPA, 2010), namely, labeling organisms with a set of vital, fluorescing probes and manually counting fluorescent organisms via microscopy. This is the standard method used in land-based verification of ballast water management systems, and test participants designated this as the reference method for evaluating compliance tools. Fluorophores—chloromethylfluorescein diacetate (CMFDA) and fluorescein diacetate (FDA)—are added to a water sample. After a brief (10-min) incubation period, the sample is transferred into a gridded counting chamber, and a portion of the chamber is scanned for organisms moving, fluorescing, or both. Fluorescing organisms encountered were identified to general taxonomic group (e.g., dinoflagellates, diatoms, etc.) and manually tallied on a datasheet. At GSI, a validation study demonstrated that a single fluorophore (FDA) yielded equivalent counts of organisms as the dual set, so at this site, only FDA was used to label organisms. The detailed protocol for this approach is in **Appendix A**.

Measuring Cell Concentration using 10Cells

The instrument was contained in a water-resistant case; consumable materials (disposable syringes and 10-µm mesh filter disks) and reusable materials (a 50-µm sieve and filter disk cartridges) were supplied for the testing. The analytical method included an optional step of pre-

screening water through a 50-µm mesh (to eliminate organisms larger than 50 µm). The sample water, whether pre-screened or not, was aspirated into a 10-mL syringe, and that water was then passed through a 10-µm filter affixed to the syringe tip. The 10-µm filter was held into place on a cartridge that was inserted into the instrument so that the organisms and particles ≥10 µm on the filter were positioned for light exposure and fluorescence detection. Measurements of fluorescence yield were converted to cell concentrations (mL^{-1}) using a cell factor of 89, which was set prior to the first trial by the vendor.

RESULTS
Linearity

The linear response of the10Cells was measured by the change in reported concentrations relative to the measured concentration of organisms ≥10 and <50 µm. Results of the laboratory and field trials are shown in Figures 1 and 2, respectively. For both laboratory and field trials, linear regression was used to generate a line-of-best-fit describing the relationship between concentration and abundance. A linear relationship indicates the compliance tool's measurements will vary in proportion to the number of organisms in the sample. The strength of that relationship is measured by the coefficient of determination (R^2), which ranges from 0 to 1 and indicates how well the measurement conform to the line-of-best fit. Linear regression was performed on data from all trials for each organism or field site as well as the combined data set (from both organisms and all field sites). Results of linear regression analyses are shown in Table 1.
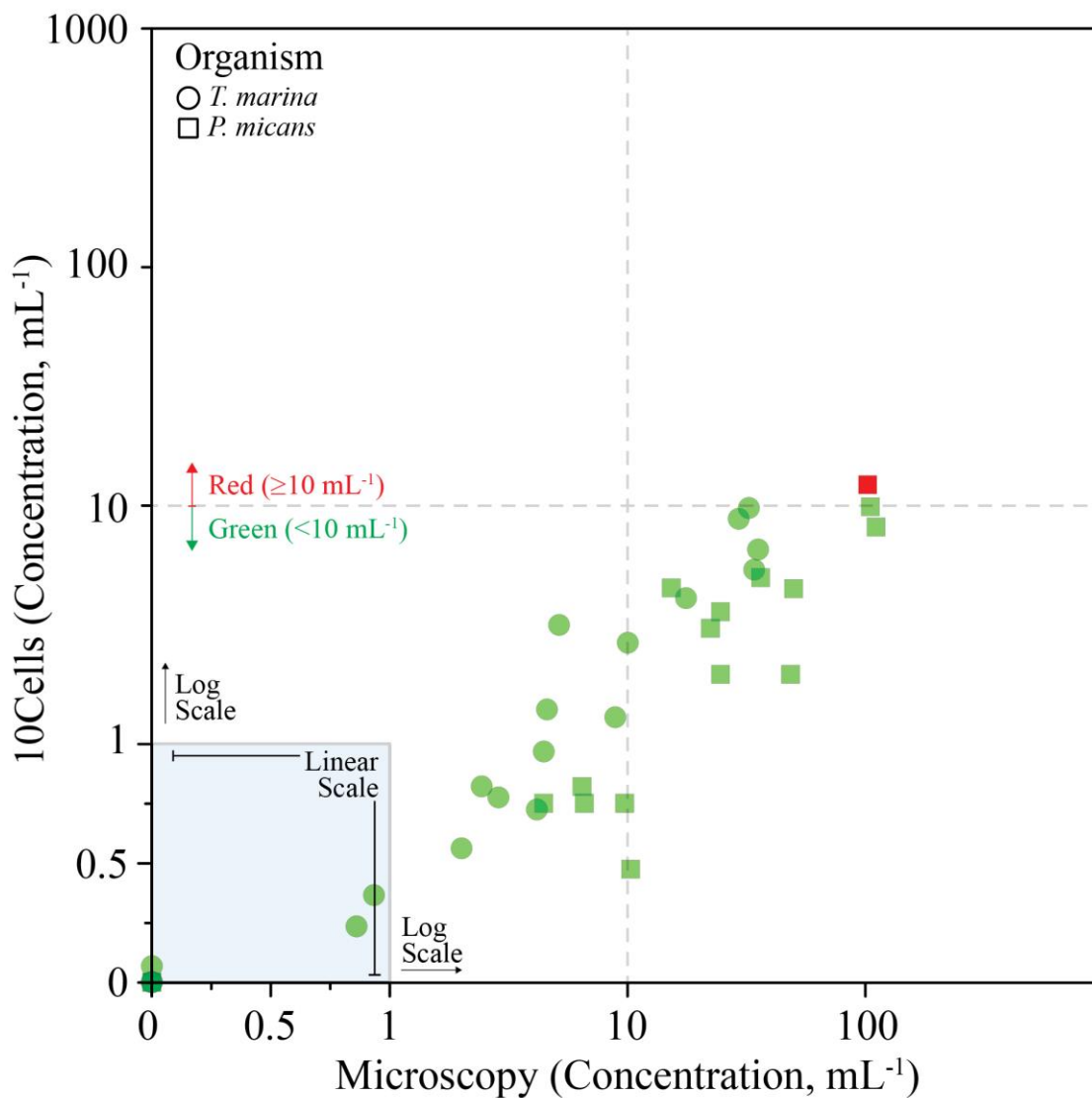
Figure 1. Results of the laboratory experiments. Measurements from the 10Cells are compared to concentrations of *Tetraselmis marina* or *Prorocentrum micans*. Symbols mark the mean organism concentration and colors show whether the reported value was ≥10 mL^{-1} (Red) or <10 mL^{-1} (Green). The figure inset, with axis scaled from 0 to 1, has a linear scale. The rest of the figure displays data on a logarithmic scale.
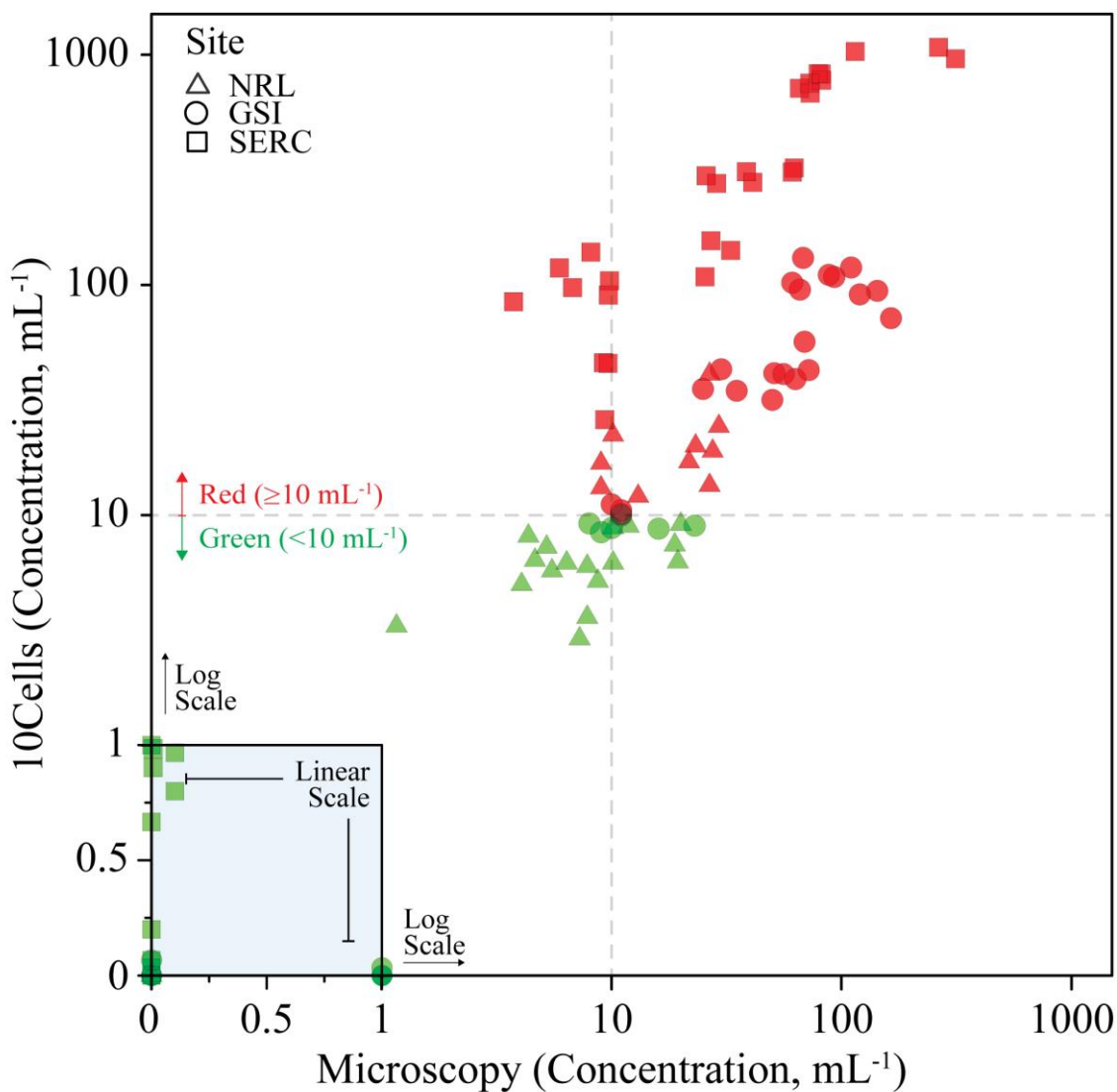
Figure 2. Results of the field experiments. Measurements from the 10Cells are compared to concentrations of ambient organisms ≥10 and <50 μm at the three test sites. Symbols mark the mean organism concentration and colors show whether the reported value was ≥10 mL^{-1} (Red) or <10 mL^{-1} (Green). The figure inset, with axis scaled from 0 to 1, has a linear scale. The rest of the figure displays data on a logarithmic scale.

Table 1. Results of linear regression analyses for both laboratory and field trials. Values indicate the adjusted (Adj.) R^2 value, the standard error (SE) of the estimates, *F*-values, slopes and y-intercepts (int.) of the relationship between estimates of cell concentrations from microscopy and 10Cells.

| Data Set	Adj. R^2	R^2 SE	*F*-Value	Slope (±SE)	y-int. (±SE)	n
Laboratory Trials						
Both organisms	0.683	1.90	$F_{1,34}= 76.3$	0.10±0.01	0.86±0.39	36
T. marina	0.852	1.20	$F_{1,16}= 99.1$	0.23±0.02	0.17±0.37	18
P. micans	0.844	1.45	$F_{1,16}= 92.8$	0.09±0.01	0.28±0.46	18
Field Trials						
All Sites	0.484	170.9	$F_{1,106} = 101.3$	3.37±0.34	7.79±19.52	108
NRL	0.612	5.5	$F_{1,34} = 56.2$	0.76±0.1	1.11±1.35	36
GSI	0.687	23.3	$F_{1,34} = 77.8$	0.79±0.09	5.82±5.34	36
SERC	0.679	196	$F_{1,34} = 5.54$	4.21±0.49	111.9±38.94	36

All p-values for regressions <0.001

In laboratory trials, readings of cell concentration were strongly related to cell concentrations of both *T. marina* and *P. micans* (Adj. R^2 > 0.8; Figure 1 and Table 1); however, combining data sets from both organisms reduced the R^2 value to 0.68 (Table 1). For field trials, R^2 values ranged from 0.48 to 0.68 (Table 1 and Figure 2).

Precision

Precision is a measure of the variation among repeated analyses. The precision of the instrument was determined by calculating the coefficient of variation (CV, %), a relative measure of the variation among replicate readings. CV is sensitive to small mean values (e.g., mean cell concentration <10): as mean approaches 0, CV approaches infinity. Because of this, the CV of mean concentrations <10 mL^{-1} were reported, but only CV from samples ≥10 mL^{-1} were used to summarize of values of CV. In laboratory trials, with the exception of one sample of *P. micans*, mean concentrations were <10 mL^{-1} (Table 2). Therefore, the measurements were not high enough to report a CV mean and range. For field trials, most samples with target concentrations above or well above the discharge standard had mean concentrations ≥10 mL^{-1} (Table 3). From these samples, the CV of three subsamples (each with three readings) ranged from 6 to 52% (24% and 22%, mean and median CV, respectively, n = 21).

Table 2. Mean, standard deviation (SD), and coefficient of variation (CV) of 10Cells concentrations in laboratory trials (n = 3 for each sample). The rows show the target cell concentrations: 0, 5, 10, 20, 50, and 100 mL^{-1}. Mean values were rounded to the nearest integer, and a zero mean yielded an undefined CV. The black circle marks the sample with concentrations ≥10 mL; the gray circles are shown as 10 mL^{-1} upon rounding to the nearest integer, but values were <10 mL^{-1}, so they are not colored red in Figure 1. Only one measurement exceeded 10 mL^{-1}.

Target Concentration	Organism	Trial ID	Concentration (mL^{-1}) Mean ± SD		CV
0 mL^{-1}	*T. marina*	LAB-1	0 ± 0		Undefined
		LAB-2	0 ± 0		Undefined
		LAB-3	0 ± 0		Undefined
	P. micans	LAB-1	0 ± 0		Undefined
		LAB-2	0 ± 0		Undefined
		LAB-3	0 ± 0		Undefined
5 mL^{-1}	*T. marina*	LAB-1	0 ± 0.1		25%
		LAB-2	1 ± 0.2		33%
		LAB-3	0 ± 0.1		16%
	P. micans	LAB-1	1 ± 0.4		67%
		LAB-2	1 ± 0.3		57%
		LAB-3	1 ± 0.2		23%
10 mL^{-1}	*T. marina*	LAB-1	1 ± 0.1		9%
		LAB-2	1 ± 0.2		29%
		LAB-3	1 ± 0.2		16%
	P. micans	LAB-1	1 ± 0.2		27%
		LAB-2	0 ± 0.2		58%
		LAB-3	5 ± 4.3		95%
20 mL^{-1}	*T. marina*	LAB-1	1 ± 0.3		21%
		LAB-2	3 ± 1.0		38%
		LAB-3	1 ± 0.1		8%
	P. micans	LAB-1	2 ± 1.3		67%
		LAB-2	4 ± 2.3		63%
		LAB-3	3 ± 1.9		62%
50 mL^{-1}	*T. marina*	LAB-1	3 ± 0.7		22%
		LAB-2	5 ± 1.4		25%
		LAB-3	4 ± 1.6		40%
	P. micans	LAB-1	5 ± 4.2		83%
		LAB-2	2 ± 0.5		23%
		LAB-3	5 ± 1.6		34%
100 mL^{-1}	*T. marina*	LAB-1	7 ± 1.7		26%
		LAB-2	10 ± 1.1	●	11%
		LAB-3	9 ± 2.8		32%
	P. micans	LAB-1	12 ± 10	●	84%
		LAB-2	10 ± 7.0	●	71%
		LAB-3	8 ± 1.6		19%

Table 3. Mean, standard deviation (SD), and coefficient of variation (CV) of 10Cells concentrations in field trials (n=9 for each sample). The rows show the general sample concentration ranges: Control (target cell concentration = 0 mL^{-1}), near the discharge standard (Near DS, 5 – 10 mL^{-1}), Above DS (30 – 50 mL^{-1}), or Well Above the DS (>50 mL^{-1}). Mean values were rounded to the nearest integer, and SD values were reported with at least two significant figures. Note that a zero mean value yielded an undefined CV. Black circles mark samples with concentrations ≥10 mL^{-1}.

Sample	Trial ID	Concentration (mL^{-1}) Mean ± SD		CV
Control	NRL-1	0 ± 0		Undefined
	NRL-2	0 ± 0		Undefined
	NRL-3	0 ± 0		Undefined
	GSI-1	0 ± 0		Undefined
	GSI-2	0 ± 0		Undefined
	GSI-3	0 ± 0		Undefined
	SERC-1	0 ± 0		Undefined
	SERC-2	0 ± 0		Undefined
	SERC-3	1 ± 0.2		26%
Near DS	NRL-1	13 ± 6.3	●	48%
	NRL-2	6 ± 0.7		13%
	NRL-3	3 ± 1.5		46%
	GSI-1	9 ± 1		11%
	GSI-2	10 ± 2	●	19%
	GSI-3	10 ± 1.5	●	16%
	SERC-1	39 ± 10	●	26%
	SERC-2	97 ± 16	●	17%
	SERC-3	114 ± 25	●	21%
Above DS	NRL-1	18 ± 7.1	●	39%
	NRL-2	8 ± 1.9		25%
	NRL-3	6 ± 1.9		33%
	GSI-1	39 ± 10	●	26%
	GSI-2	44 ± 11	●	25%
	GSI-3	39 ± 5.3	●	14%
	SERC-1	135 ± 32	●	23%
	SERC-2	314 ± 53	●	17%
	SERC-3	284 ± 16	●	6%
Well Above DS	NRL-1	28 ± 15	●	52%
	NRL-2	13 ± 6.2	●	47%
	NRL-3	7 ± 2.1	●	29%
	GSI-1	95 ± 23	●	24%
	GSI-2	96 ± 14	●	14%
	GSI-3	117 ± 17	●	15%
	SERC-1	1024 ± 128	●	12%
	SERC-2	716 ± 48	●	7%
	SERC-3	809 ± 182	●	22%

Accuracy

Accuracy of the instrument is a measure of the difference between a measurement and the actual or expected value from a recognized standard for measuring organisms ≥10 and <50 µm. (Note: from the Test Protocols "Accuracy is measured as the proportion of samples that correctly assess whether a sample meets the discharge standard"). For each sample examined, the instrument reports cell concentration, and values ≥10 mL^{-1} were considered to have not met the discharge standard (DS). A logistical regression analysis was used to determine the probability that the instrument correctly identifies exceedances of the DS as cell concentrations diverge from the DS, whether below the DS (e.g., 0 to 9 mL) or above the DS. Results of the logistical regression analyses are shown in Table 4.

Table 4. Logistic regression results for the field trials. Only one sample in the laboratory trials ≥10 mL^{-1}, so there was insufficient variability in outcomes to perform logistic regression analysis and results are not available (N/A).

		Constant (*C*)			Coefficient (*x*)			
		Value	**SE**	**p-Value**	**Value**	**SE**	**p-Value**	**n**
Laboratory Trials	Both organisms	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	T. marina	N/A	N/A	N/A	N/A	N/A	N/A	N/A
	P. micans	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Field Trials	All Sites	-0.50	0.30	0.097	0.183	0.041	0.001	108
	NRL	-1.39	0.55	0.012	0.211	0.067	0.002	36
	GSI	-1.22	0.66	0.063	0.176	0.075	0.019	36
	SERC	64.5	3325	0.985	8.104	413	0.984	36

To visualize the results of this analysis, the resulting values—the constant (*C*) and the coefficient (*x*)—were used to calculate the probability (*ρ*) of a High *Risk* (*H*) outcome across a range of cell concentrations (*P*):

EQ. 1
$$\rho(H) = \frac{1}{(1+e^{(-C+xP)})}$$

Resulting *ρ(H)* values across a range of cell concentrations are shown in Figure 3. At an organism concentration of 30 mL^{-1}, which is three times the DS exceedance, the probability of High *Risk* (*ρ(H)*) was 0.99, 0.98, 1.00, and 0.99 for NRL, GSI, SERC, and All Field Sites, respectively (Figure 3).
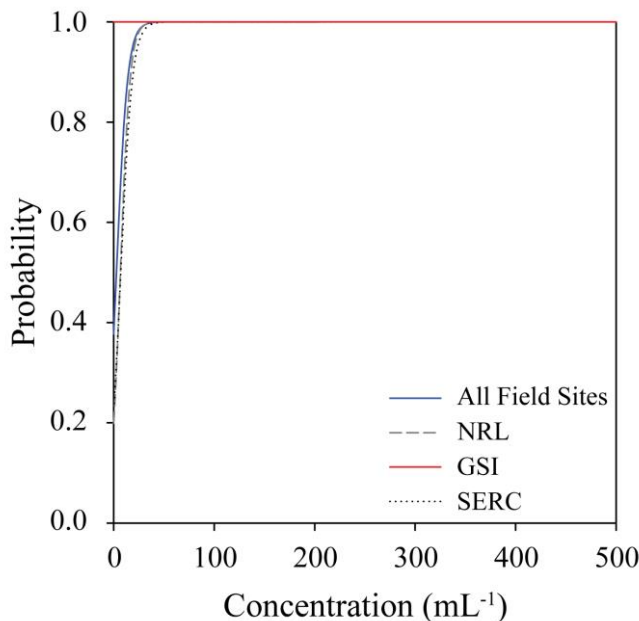
Figure 3. Probability of indicating a sample is above the discharge limit (≥ 10 mL^{-1}) based upon cell concentrations in field trials. There were not enough measurements ≥ 10 mL^{-1} in the laboratory trials to perform logistical regressions and generate equation parameters.

QUALITY MANAGEMENT

All technical activities conducted by ACT and NRL comply with their respective Quality Management System (QMS), which includes the policies, objectives, procedures, authority, and accountability needed to ensure quality in work processes, products, and services. A QMS provides the framework for quality assurance (QA) functions, which cover planning, implementation, and review of data collection activities and the use of data in decision-making, and quality control. The QMS also ensures that all data collection and processing activities are carried out in a consistent manner, to produce data of known and documented quality that can be used with a high degree of certainty by the intended user to support specific decisions or actions regarding technology performance. Both organizations' QMS meet U.S. Environmental Protection Agency quality standards for environmental data collection, production, and use. The QMS also meets the requirements of General Requirements for the Competence of Testing and Calibration Laboratories (ISO/IEC 17025:2005[E]).

An effective assessment program is an integral part of a quality system. The ACT Quality Assurance (QA) Manager independently conducted six Technical Systems Audits (TSA, described below) and data quality assessments of all reference data sets for the evaluation.

Technical System Audits

A TSA is a thorough, systematic, on-site qualitative audit of sampling and measurement processes and procedures associated with a specific technology evaluation. The objectives of the

TSAs conducted during this evaluation were to assess and document the conformance of on-site testing procedures with the requirements of the Test Protocols and associated Standard Operating Procedures (SOPs).

The TSAs were conducted in accordance with the procedures described in EPA's Guidance on Technical Audits and Related Assessments for Environmental Data Operations (EPA QA/G-7) and ISO 19011, Guidelines for Quality and/or Environmental Management Systems Auditing. A TSA checklist based on the Test Protocols was prepared prior to each audit and reviewed by the respective laboratory's personnel. The TSA assessed the respective laboratories' personnel, the test and analytical facilities, equipment maintenance and calibration procedures, sample collection, analytical activities, record keeping, and QC procedures. The audits were conducted for all field trials and laboratory trials.

During each audit, the auditor met with each person involved in testing and asked that person to describe the procedures. All procedures were observed, and logbooks, data forms, and other records were reviewed.

Key components of each audit included assessments of the following:
Quality Assurance/Quality Control:
- Adequacy of procedures and adherence to procedures
- Chain of command regarding description of assignments and specific duties

Sample System:
- Sample collection
- Analytical procedures
- Analytical equipment maintenance and calibration
- Documentation.

Data and Document Control:
- Chain of custody
- Validation and processing procedures
- Documentation

The findings of the TSA for the four field tests and two laboratory tests were positive. All of these tests were being implemented consistent with the Test Protocols and SOPs. Minor deviations were documented in laboratory records. None of the deviations had an effect on data quality for the evaluation Test Instruments. Failures were due to mechanical problems with the instrument. All phases of the implementation of the test reviewed during the TSAs were acceptable and performed in a manner consistent with ACT/NRL data quality goals. The overall quality assurance objectives of the test were met.

ACT and NRL personnel are well qualified to implement the evaluation and demonstrated expertise in pertinent procedures. Communication and coordination among all personnel was frequent and effective. Internal record keeping and document control was well organized. The

ACT and NRL staff understands the need for QC, as shown in the conscientious development and implementation of a variety of QC procedures.

All samples and instrument measurements were collected, analyzed and cataloged as described in the Test Protocols and SOPs. Examination of maintenance and calibration logs provided evidence of recent and suitable calibration of sampling and analytical equipment.

Data Assessments

Data review was conducted to ensure that only sound data that are of known and documented quality and meet quality objectives were used in making decisions about technology performance. Data review processes are based in part on two EPA guidance documents: Guidance on Environmental Data Verification and Data Validation (QA/G-8) (EPA, 2002) and Guidance on Technical Audits and Related Assessments for Environmental Data Operations (QA/G-7) (EPA, 2000).

At the outset of the evaluation, data were verified and validated to evaluate whether data were generated according to the Test Protocols, satisfied acceptance criteria, and were appropriate for their intended use of evaluating the performance of the test instruments. Data verification evaluates the completeness, correctness, and consistency of data sets against the requirements specified in the Test Protocols, measurement quality objectives, and any other analytical process requirements contained in SOPs. The ACT QA Manager reviewed the reference (microscopy) data sets from all field and laboratory tests. Thirty-six (36) reference samples were counted for each field test (total 216 microscopy counts); fifty-six (56) reference samples were counted for each laboratory test (total 112 microscopy counts). The overall reference data set included 328 microscopy counts. The data review verified that the sampling and analysis protocols specified in the Test Protocols were followed, and that the ACT/NRL measurement and analytical systems performed in accordance with approved methods, based on the following criteria:

- The raw data records were complete, understandable, well-labeled, and traceable
- All data identified in the Test Protocols were collected
- QC criteria were achieved
- Data calculations were accurate

Data validation uses the outputs from data verification and included inspection of the verified field and laboratory data to determine the analytical quality of data set. A representative set of approximately 10% of the reference data was traced in detail from 1) raw data from field and laboratory logs, 2) data transcription, 3) data reduction and calculations, to 4) final reported data. Validation of the referenced data set established:

- Required sampling methods were used
- Sampling procedures and field measurements met performance criteria
- Required analytical methods were used

The data validation also confirmed that data were accumulated, transferred, summarized, and reported correctly. There is sufficient documentation of all procedures used in data collection and analysis to validate that data were collected in accordance with the evaluation's quality objectives.

A Data Quality Assessment (DQA) is the third and final process of the overall data assessment. It is a scientific and statistical evaluation of validated data to determine if data are of the right type, quality, and quantity to support conclusions on the performance of the test instruments. The DQA determined that the evaluation's data quality objectives, described in the Test Protocols (Appendix A) were achieved.

REFERENCES

International Maritime Organization; IMO (2004) International convention for the control and management of ships' ballast water and sediments. http://www.imo.org/OurWork/Environment/BallastWaterManagement/Pages/BWMConventionandGuidelines.aspx

Steinberg MK, EJ Lemieux and LA Drake (2011) Determining the viability of marine protists using a combination of vital, fluorescent stains. Marine Biology 158: 1431–1437

US Coast Guard (2012) Standards for living organisms in ships' ballast water discharged in U.S. waters. Federal Register 77:17254–17320

US Environmental Protection Agency (2000) Guidance on technical audits and related assessments for environmental data operations. Washington, DC Report number EPA/600/R-99/080, 101 pp, https://www.epa.gov/sites/production/files/2015-07/documents/g7-final.pdf

US Environmental Protection Agency (2002) Guidance on environmental data verification and validation. Washington, DC Report number EPA/240/R-02/004, 96 pp, https://www.epa.gov/sites/production/files/2015-06/documents/g8-final.pdf

US Environmental Protection Agency Environmental Technology Verification Program (2010) Generic protocol for the verification of ballast water treatment technology. Washington, DC, Report number EPA/600/R-10/146, 156 pp, http://www.uscg.mil/hq/cg5/cg522/cg5224/docs/600r10146.pdf

The vendor's formal response letter is appended as a separate document.

APPENDIX A: TEST PLAN

Available for download at www.act-us.info/evaluations.

APPENDIX B: RAW DATA

Table 1. Summary of trials conducted.

Location	Trial Name	Trial Date	Trial Replicate
Naval Research Laboratory (NRL; Key West, FL)	NRL-1	7/14/2015	1 of 3
	NRL-2	7/15/2015	2 of 3
	NRL-3	7/16/2015	3 of 3
Laboratory Trial (LAB; Key West, FL)	LAB-1	6/2/2015	1 of 3
	LAB-2	6/3/2015	2 of 3
	LAB-3	6/4/2015	3 of 3
Smithsonian Environmental Research Center (SERC; Edgewater, MD)	SERC-1	8/7/2015	1 of 3
	SERC-2	8/8/2015	2 of 3
	SERC-3	8/10/2015	3 of 3
Great Ships Initiative (GSI; Superior, WI)	GSI-1	9/1/2015	1 of 3
	GSI-2	9/2/2015	2 of 3
	GSI-3	9/3/2015	3 of 3

Table 2. Concentrations of living organisms ≥10 and <50 µm in samples from field trials. Target concentrations were Control (0 mL^{-1}), near the discharge standard (DS, 5 – 20 mL^{-1}), above the DS (30 – 50 mL^{-1}), and well above the DS (>50 mL^{-1}).

Trial Number	Sample		Concentration (mL^{-1})		
			NRL	GSI	SERC
1 of 3	Control	A	0	5	0
		B	0	1	0
		C	0	1	0
	Near DS	A	9	9	10
		B	9	23	9
		C	20	16	9
	Above DS	A	23	30	26
		B	13	56	27
		C	10	50	33
	Well Above DS	A	27	110	313
		B	29	143	264
		C	28	164	115
2 of 3	Control	A	0	0	0
		B	0	1	0
		C	0	1	0
	Near DS	A	8	11	10
		B	9	11	7
		C	6	11	10
	Above DS	A	4	69	62
		B	19	25	39
		C	5	51	61
	Well Above DS	A	12	61	66
		B	27	120	73
		C	22	66	73
3 of 3	Control	A	0	0	0
		B	0	0	0
		C	0	0	0
	Near DS	A	7	10	6
		B	1	10	8
		C	8	8	4
	Above DS	A	10	63	26
		B	6	35	29
		C	4	72	41
	Well Above DS	A	5	68	82
		B	10	93	82
		C	19	88	79

Table 3. Concentrations of cultured organisms in samples from laboratory experiments. In two samples, the cultured organisms—*Tetraselmis marina* and *Prorocentrum micans*—were amended with dissolved and particulate materials or disinfection byproducts (DBP). Target concentrations ranged from 0 to 100 mL^{-1}. These are the actual concentrations used for comparisons with the 10Cells.

| Trial | Sample | Concentration (mL^{-1}) | |
		T. marina	*P. micans*
LAB-1	0 mL^{-1}	0	0
	5 mL^{-1}	0	7
	10 mL^{-1}	2	10
	20 mL^{-1}	5	25
	50 mL^{-1}	5	36
	100 mL^{-1}	35	102
	10 mL^{-1} (Amended)	2	12
	10 mL^{-1} (DBP)	2	9
LAB-2	0 mL^{-1}	0	0
	5 mL^{-1}	3	4
	10 mL^{-1}	4	10
	20 mL^{-1}	10	25
	50 mL^{-1}	34	48
	100 mL^{-1}	32	105
	10 mL^{-1} (Amended)	3	10
	10 mL^{-1} (DBP)	0	6
LAB-3	0 mL^{-1}	0	0
	5 mL^{-1}	2	6
	10 mL^{-1}	4	15
	20 mL^{-1}	9	22
	50 mL^{-1}	18	50
	100 mL^{-1}	29	111
	10 mL^{-1} (Amended)	4	10
	10 mL^{-1} (DBP)	0	6

Table 4. 10Cells cell concentration (mL^{-1}) of samples from field trials at NRL. Red symbols (●) indicate a measured concentration exceeding the discharge standard (DS).

Trial	Sample		Replicate Reading 1 of 3	2 of 3	3 of 3	Mean	SD
NRL-1	Control	A	0	0	0	0.0	0.1
		B	0	0	0	-0.1	0.1
		C	0	0	0	0.0	0.1
	Near DS	A	25 ●	16 ●	10	16.8	7.8
		B	7	12 ●	21 ●	13.2	7.3
		C	9	10	9	9.1	0.4
	Above DS	A	11 ●	24 ●	25 ●	20.0	7.8
		B	13 ●	9	14 ●	12.1	2.5
		C	23 ●	29 ●	15 ●	22.3	7.0
	Well Above DS	A	25 ●	41 ●	57 ●	41.0	16.0
		B	15 ●	41 ●	17 ●	24.3	14.5
		C	16 ●	19 ●	22 ●	19.0	3.0
NRL-2	Control	A	0	0	0	0.0	0.0
		B	0	0	0	0.0	0.0
		C	0	-1	0	-0.4	0.7
	Near DS	A	5	7	6	6.0	0.7
		B	6	5	5	5.2	0.8
		C	6	7	6	6.2	0.4
	Above DS	A	7	7	11 ●	8.1	2.5
		B	8	8	7	7.4	0.7
		C	4	8	9	7.3	2.6
	Well Above DS	A	9	6	12 ●	9.0	2.9
		B	7	19 ●	15 ●	13.5	6.4
		C	9	24 ●	18 ●	17.1	7.4
NRL-3	Control	A	0	0	0	0.0	0.0
		B	0	0	0	0.0	0.1
		C	0	0	0	0.0	0.0
	Near DS	A	3	5	0	2.9	2.4
		B	2	3	5	3.3	1.5
		C	4	4	3	3.6	0.8
	Above DS	A	4	9	5	6.2	2.8
		B	4	7	7	5.7	1.8
		C	5	7	4	5.0	1.5
	Well Above DS	A	7	7	5	6.4	1.1
		B	6	12 ●	9	8.8	3.0
		C	5	6	7	6.3	0.9

Table 5. 10Cells cell concentration (mL^{-1}) of samples from field trials at GSI. Red symbols (●) indicate a measured concentration exceeding the discharge standard (DS).

Trial	Sample		Replicate Reading			Mean	SD
			1 of 3	2 of 3	3 of 3		
GSI-1	Control	A	0	0	0	0.0	0.0
		B	0	0	0	0.0	0.1
		C	0	0	0	0.0	0.0
	Near DS	A	9	8	9	8.4	0.3
		B	10 ●	10 ●	7	9.0	1.5
		C	8	10 ●	8	8.7	1.1
	Above DS	A	43 ●	51 ●	35 ●	43.0	8.0
		B	57 ●	32 ●	34 ●	41.0	13.9
		C	32 ●	25 ●	38 ●	31.7	6.5
	Well Above DS	A	112 ●	129 ●	116 ●	119.0	8.9
		B	107 ●	101 ●	75 ●	94.3	17.0
		C	66 ●	82 ●	67 ●	71.7	9.0
GSI-2	Control	A	0	0	0	0.1	0.1
		B	0	0	0	0.0	0.0
		C	0	0	0	0.0	0.0
	Near DS	A	9	9	13 ●	10.1	2.5
		B	11 ●	9	10 ●	10.0	1.1
		C	11 ●	13 ●	8	10.5	2.8
	Above DS	A	57 ●	67 ●	46 ●	56.7	10.5
		B	35 ●	33 ●	38 ●	35.3	2.5
		C	47 ●	40 ●	37 ●	41.3	5.1
	Well Above DS	A	124 ●	86 ●	97 ●	102.3	19.6
		B	80 ●	98 ●	95 ●	91.0	9.6
		C	107 ●	98 ●	81 ●	95.3	13.2
GSI-3	Control	A	0	0	0	0.0	0.0
		B	0	0	0	0.0	0.0
		C	0	0	0	0.0	0.0
	Near DS	A	9	10	8	8.8	0.7
		B	9	11 ●	13 ●	11.1	1.8
		C	10	8 ●	9	9.2	0.8
	Above DS	A	39 ●	36 ●	42 ●	39.0	3.0
		B	33 ●	36 ●	35 ●	34.7	1.5
		C	38 ●	39 ●	51 ●	42.7	7.2
	Well Above DS	A	104 ●	134 ●	155 ●	131.0	25.6
		B	106 ●	109 ●	110 ●	108.3	2.1
		C	111 ●	116 ●	105 ●	110.7	5.5

*Value rounded to 10 mL^{-1}.

Table 6. 10Cells cell concentration (mL^{-1}) of samples from field trials at SERC. Red symbols (●) indicate a High (or *Fail*) *Risk* of exceeding the discharge standard (DS).

Trial	Sample		1 of 3	2 of 3	3 of 3	Mean	SD
SERC-1	Control	A	0	0	0	0.2	0.1
		B	0	0	0	0.0	0.1
		C	0	0	0	0.1	0.1
	Near DS	A	45 ●	46 ●	46 ●	45.7	0.6
		B	42 ●	47 ●	49 ●	46.0	3.6
		C	21 ●	26 ●	31 ●	26.0	5.0
	Above DS	A	79 ●	85 ●	162 ●	108.7	46.3
		B	150 ●	157 ●	160 ●	155.7	5.1
		C	151 ●	141 ●	132 ●	141.3	9.5
	Well Above DS	A	725 ●	1140 ●	1020 ●	961.7	213.6
		B	1080 ●	1030 ●	1120 ●	1076.7	45.1
		C	964 ●	1130 ●	1010 ●	1034.7	85.7
SERC-2	Control	A	0	0	2	0.7	1.0
		B	0	0	0	0.0	0.0
		C	0	0	0	0.0	0.0
	Near DS	A	76 ●	121 ●	74 ●	90.3	26.6
		B	97 ●	87 ●	108 ●	97.3	10.5
		C	103 ●	115 ●	95 ●	104.3	10.1
	Above DS	A	336 ●	367 ●	264 ●	322.3	52.8
		B	312 ●	284 ●	335 ●	310.3	25.5
		C	258 ●	261 ●	409 ●	309.3	86.3
	Well Above DS	A	696 ●	749 ●	699 ●	714.7	29.8
		B	663 ●	750 ●	628 ●	680.3	62.8
		C	754 ●	772 ●	734 ●	753.3	19.0
SERC-3	Control	A	1	1	1	0.8	0.1
		B	1	1	1	1.0	0.4
		C	1	1	1	1.0	0.2
	Near DS	A	118 ●	125 ●	113 ●	118.7	6.0
		B	144 ●	133 ●	140 ●	139.0	5.6
		C	92 ●	76 ●	86 ●	84.7	8.1
	Above DS	A	297 ●	309 ●	288 ●	298.0	10.5
		B	299 ●	266 ●	263 ●	276.0	20.0
		C	281 ●	271 ●	286 ●	279.3	7.6
	Well Above DS	A	670 ●	1090 ●	716 ●	825.3	230.4
		B	829 ●	722 ●	775 ●	775.3	53.5
		C	658 ●	1140 ●	685 ●	827.7	270.8

Table 7. 10Cells cell concentration (mL^{-1}) of *T. marina* in samples from laboratory trials. In two samples, cultures were amended with dissolved and particulate materials or disinfection byproducts (DBP). Red symbols (●) indicate a measured concentration exceeding the discharge standard (DS).

Trial	*T. marina* sample	Replicate Reading			Mean	SD
		1 of 3	2 of 3	3 of 3		
LAB-1	0 mL^{-1}	0	0	0	0.0	0.0
	5 mL^{-1}	0	0	0	0.2	0.1
	10 mL^{-1}	1	1	1	0.7	0.1
	20 mL^{-1}	1	2	1	1.4	0.3
	50 mL^{-1}	2	3	4	3.2	0.7
	100 mL^{-1}	8	5	8	6.6	1.7
	10 mL^{-1} (Amended)	1	1	1	0.5	0.0
	10 mL^{-1} (DBP)	1	1	0	0.5	0.5
LAB-2	0 mL^{-1}	0	0	0	0.0	0.0
	5 mL^{-1}	1	0	1	0.6	0.2
	10 mL^{-1}	1	0	1	0.5	0.2
	20 mL^{-1}	2	3	3	2.7	1.0
	50 mL^{-1}	4	6	7	5.4	1.4
	100 mL^{-1}	9	11 ●	10 ●	9.8	1.1
	10 mL^{-1} (Amended)	1	1	-4	-0.8	2.6
	10 mL^{-1} (DBP)	0	0	0	0.0	0.0
LAB-3	0 mL^{-1}	0	0	0	0.1	0.1
	5 mL^{-1}	0	0	0	0.4	0.1
	10 mL^{-1}	1	1	1	0.9	0.2
	20 mL^{-1}	1	1	1	1.3	0.1
	50 mL^{-1}	3	3	6	4.1	1.6
	100 mL^{-1}	7	12 ●	8	8.8	2.8
	10 mL^{-1} (Amended)	1	1	1	0.7	0.1
	10 mL^{-1} (DBP)	0	0	0	0.0	0.1

*Value rounded to 10 mL^{-1}

Table 8. 10Cells cell concentration (mL^{-1}) of *P. micans* in samples from laboratory trials. In two samples, cultures were amended with dissolved and particulate materials or disinfection byproducts (DBP). Red symbols (●) indicate a measured concentration exceeding the discharge standard (DS).

Trial	*P. micans* sample	Replicate Reading			Mean	SD
		1 of 3	2 of 3	3 of 3		
LAB-1	0 mL^{-1}	1	0	0	0.6	0.4
	5 mL^{-1}	1	0	1	0.6	0.2
	10 mL^{-1}	1	2	3	2.0	1.3
	20 mL^{-1}	2	4	10 ●	5.0	4.2
	50 mL^{-1}	8	24 ●	5	12.3	10.3
	100 mL^{-1}	1	2	1	1.2	0.5
	10 mL^{-1} (Amended)	0	1	2	0.9	0.6
	10 mL^{-1} (DBP)	0	0	0	0.0	0.0
LAB-2	0 mL^{-1}	1	0	1	0.6	0.3
	5 mL^{-1}	0	0	1	0.3	0.2
	10 mL^{-1}	3	2	6	3.6	2.3
	20 mL^{-1}	2	2	2	2.0	0.5
	50 mL^{-1}	18 ●	5	7	9.9	7.0
	100 mL^{-1}	1	2	1	1.3	0.4
	10 mL^{-1} (Amended)	1	2	3	2.0	1.2
	10 mL^{-1} (DBP)	0	0	0	0.0	0.0
LAB-3	0 mL^{-1}	1	1	1	0.7	0.2
	5 mL^{-1}	2	10 ●	2	4.5	4.3
	10 mL^{-1}	5	3	1	3.1	1.9
	20 mL^{-1}	3	6	4	4.5	1.6
	50 mL^{-1}	8	7	10 ●	8.2	1.6
	100 mL^{-1}	3	1	3	2.3	1.3
	10 mL^{-1} (Amended)	0	0	0	0.1	0.1
	10 mL^{-1} (DBP)	1	0	0	0.6	0.4

*Value rounded to 10 mL^{-1}

Comment on the findings of the evaluation of the bbe 10cells

Dear ACT-Team, dear reader!

Many thanks to the evaluation team for their effort in evaluating the bbe 10cells. We saw how much effort went into carrying out all these tests.

The tests related to examining the parameters: linearity, precision and accuracy which determine the equipment behaviour both in the lab and in the field. To this end comprehensive quality standards were established and largely met.

The bbe 10cells could demonstrate its potential both in lab and in field tests even under diverse conditions.

While 10Cells performed very well, we feel there is an unavoidable limitation in this evaluation study. This depends greatly on the method for estimating the true or actual number of live cells in a sample and on the preparation of dilutions. Unfortunately, there was no true reference standard to compare instrument performance against at the time being, so it was agreed that comparisons would be made between the accepted ballast water management system certification test method (which has its associated uncertainties) and instrument generated values (which also have associated uncertainties). In principle, you can see that bbe 10cells measures with pure water with an accuracy and precision well below a single cell but is apparently subject to significant noise for low cell number below the range of e.g. 50 cells/ml, and this has a great contribution by the chosen way to estimate the 'true values'.

The diagrams indicate that the counted or estimated values entered on the x-axis represent the desired values (measured directly using the stain method) for the 10cells. However, this cannot be possible, because the desired values themselves are also subject to distribution fluctuations. Even when assuming that the desired value was determined correctly with exactly e.g. 10 cells/ml, this Poisson distribution calculates that it is improbably that these 10 cells are present in the device. The probability to find exactly 10 cells is 12.5%. To find seven or 14 cells still has half the probability of finding exactly 10 cells! If the average value of 10 cells was incorrect during the calculation, the probability reduces further.

Given the determination of pure water with 0 cells, which the 10cells does very well (in the assessment mostly less than 0.2 cells/ml, table 3), we can initially assume that the 10cells measurements are very stable for low cell numbers. This aspect is supported by the fact that the microscopic counts are carried out at 1 ml volumes, whereas 10cells analyses at least 10ml (this reduces the distribution by the root of 3), if not 50 ml (the signal/noise ratio is improved by a factor of 7, the probability to find the exact 10 cells mentioned here approx. 50%) during a test. This means that the accuracy of the measurements rises significantly with the volume examined.

Already with 10ml the measurement is clearly better for reasons of probability than the count in a volume of 1 ml (as agreed to for this evaluation except the low number of repetitions).

Thus the spread is caused by both 2 different samples having been measured and the two different methods measuring two different end points (stain counts vs fluorescent signal). It is highly improbable that these are identical based on the Poisson distribution. Of these 10 ml only 1 ml is determined in the counting chamber.

This can also be noted in the spreading results of the natural samples or samples in which the same concentration is practically measured repeatedly, as can also be seen in figure 2 and the tables.

Therefore, it is obviously not very meaningful to calculate correlations which are decisively determined by the noise due to the selection of different samples for 10cells and microscopy for the comparison trial.

While these results clearly demonstrate that the 10Cells can identify non-compliance with ballast water discharge standards, bbe cannot really utilise this comprehensive work further because of the limitations described above. It should be noted that this statistical problem cannot explain all phenomena of deviations between 10cell and the counts or counting estimates.

We also found that the conversion factor of the bbe 10cells seems to be too low by a factor of five. This factor was adjusted using the results of the Meteor cruise. This is a bit odd because similar species and conditions were used to compare the methods. Maybe a round robin test also in counting should be performed as well.

Looking at the counts collected during the field trials in Key West, the evaluation shows that the problem of small cell numbers had to be addressed. It partially solves the problem by concentrating the sample by a factor of 10. Here, this step is well and clearly commented and communicated. bbe also recognized the problem of the variability of small numbers and tackled it by concentration on the filter strip. Now, the evaluation claims this advantageous step (chapter 5.5.1 of the test protocol) for its own procedures, but does not allow a clear indication of this greatly improved evaluation in the bbe device. I feel that the bbe 10cells approach has significant advantages over the stain and microscope counting approach. For example, in Chapter 6.1, there is an explicit statement:

As the ambient concentrations of organisms vary among locations (and at a location over time), the ambient concentrations should be first measured by direct counts using the method described above. Because of the site-to-site variation in concentrations of ambient organisms, it will be necessary to sample and process water differently among locations. These differences, however, should not affect the integrity of the comparison between concentration measurements by direct counts and variable fluorescence. However, I feel the integrity could be compromised in many of the trials that were carried out as part of this evaluation. Again, I hold that bbe is much stricter by always examining 10 ml to 50 ml. I feel strongly that bbe prescribes the most appropriate approach in cases of concentrations let's say lower than 50 cells ml^{-1} and thus produces the most consistent results.

Christian Moldaenke