MARINE ENVIRONMENT PROTECTION COMMITTEE
62nd session
Agenda item 2

MEPC 62/INF.31
6 May 2011
ENGLISH ONLY

HARMFUL AQUATIC ORGANISMS IN BALLAST WATER

Logistics of compliance assessment and enforcement of the Ballast Water Management Convention

Submitted by the Institute of Marine Engineering, Science and Technology (IMarEST)

SUMMARY

Executive summary:	This document provides information on the status of shipboard testing to the current D-2 standard for ballast water treatment, with implications for compliance assessment and enforcement when the Ballast Water Management Convention comes into force. We conclude that it will not be feasible to directly assess whether a vessel can meet all the numerical standards for compliance, and advocate a tiered approach to assessment, with addenda to the international standard that will facilitate this process
Strategic direction:	7
High-level action:	7.1.2
Planned output:	7.1.2.3
Action to be taken:	Paragraph 3
Related documents:	BLG 15/INF.6 and BLG 15/5/8

Introduction

1 Upon entry into force of the Ballast Water Management Convention, compliance assessment will form the basis of worldwide enforcement of the Convention. A potentially important component of this assessment will be shipboard compliance testing. The ability of the Administrations, through the port State control process, to effectively verify compliance with the Convention will depend on the precision and reliability of data obtained on use and performance of ballast water management systems. Any process that streamlines and normalizes compliance testing will result in a more even application of the Convention worldwide. Ship operators will also welcome a common and understandable approach by the port State control in verifying compliance.

2 Although performance (certification) testing has become increasingly standardized worldwide, several outstanding issues relating to test conditions, sampling strategies and endpoint determination remain to be resolved and refined. The annex to this document outlines those issues, describes the implications for compliance assessment and testing, and recommends approaches to facilitate meeting the D-2 standard.

Action requested of the Committee

3 The Committee is invited to note the information contained in this document.

***

ANNEX

LOGISTICS OF COMPLIANCE ASSESSMENT AND ENFORCEMENT OF THE BALLAST WATER MANAGEMENT CONVENTION

1 Introduction

1.1 The imminent entry into force of the 2004 IMO Ballast Water Management Convention[1] has focused urgent attention on the criteria that will be used to enforce the statute relating to effective ballast water treatment once it becomes law. The timetable indicates that ballast water exchange as a possible means of ballast water management will be completely phased out after 2016, and that all vessels must thereafter comply with the ballast water performance standard.

1.2 Certification testing to date has involved both land-based and shipboard components, entailing a matrix of replicated trials under different water conditions and having seasonal components. Once a certified system is operational on board a vessel, the question of its continuing efficacy in being able to process ballast water arises. This is coupled with the potential need for some form of periodic formal compliance testing by an external body to ensure that the vessel maintains the ability to meet the D-2 standard.

1.3 While land-based trials represent the most rigorous means of performance testing under controlled conditions, shipboard tests often may have constraints in terms of space, logistics, etc. While shipboard conditions may not always be optimal for testing, these trials have provided useful information about the effectiveness of Ballast Water Management Systems (BWMS) and whether the treatment system is being used appropriately. Shipboard testing for type approval has also revealed and informed numerous technical issues and constraints that may affect compliance testing. As such, it seems likely that shipboard testing of some kind will remain a component of compliance assessment once the above time-table has been activated.

2 Issues concerning international standards

2.1 Although performance (certification) testing has become increasingly standardized worldwide, several outstanding issues relating to test conditions, sampling strategies and endpoint determination remain to be resolved and refined. As certification testing may be seen as a 'blue-print' for eventual compliance testing, at least in its most comprehensive form, any residual ambiguities, particularly as they relate to interpretation of the worldwide standard, could have implications for eventual compliance testing. Priority should, therefore, be given to any process that streamlines the compliance testing process and minimizes ambiguities, while retaining the degree of rigour required for a legally defensible enforcement of the international standard. Article 7 of the Convention states that a vessel should not be unduly delayed by the application of an extended survey process and it is thus of importance that any compliance testing regime follow some form of recognisable common standard with an expected time scale. The lack of such formalised protocols may result in the application of local or regional testing methodologies with disputes arising from unwarranted delays while compliance testing is carried out or results awaited.

[1] http://www.imo.org

Sampling strategies

2.2 Section 6.2.2 of Guidelines (G2) states that **"the sampling protocol should result in samples that are representative of the whole discharge of ballast water from any single tank or any combination of tanks being discharged"**. A suitable sampling scheme is required to obtain a "representative sample", and considerable effort has been devoted to the definition of this term within the context of compliance testing. It is generally assumed that smaller, more numerous organisms, particularly bacteria, will have a more homogeneous distribution than larger, more sparsely distributed organisms, such as zooplankton. Because of the rarity of larger (>50 µm) organisms, the largest possible volume of water must be filtered in order to obtain an accurate estimate of their number. In the 2010 EMSA report, S. Gollasch and M. David[2] suggested that a representative sample for the <50 µm size category could be satisfactorily obtained from an integrated low-volume 'split' collected over all or part of a ballasting/de-ballasting cycle. However, within-tank collections designed to examine patchiness have shown that the densities of even relatively small organisms, i.e., in the 10-50 µm size range, may differ more than 10-fold among samples from the same tank.[3]

2.3 The recent publication by A. W. Miller *et al*. (2011)[4] demonstrates that continuously time-integrated sampling of ballast water throughout a discharge event can provide statistically sound and representative estimates of organism abundances, if conducted appropriately. Analysis of organisms is usually considered to conform to a Poisson distribution, with the consequence that the variance associated with any count is proportional to the total number or organisms counted. However, while continuous sampling of a ballast stream throughout a complete ballasting cycle best represents **"the whole discharge"** (above), logistical and time constraints, particularly in the context of examining vessels during a short port visit, may dictate that counts would have to come from smaller samples.

2.4 While the D-2 standard currently makes no mention of the error(s) associated with threshold limits, important questions facing regulators are: (a) should the error accompanying any count or set of counts be reported?, and (b) should the error be added to the mean organism count to create a standard that takes variance into account? A corollary of this relates to results from replicated tests, and poses the question: "should **every** replicate count be at or below the D-2 standard, or can a BWMS be considered compliant if the **mean** of the replicates meets the standard?" The 2010 ICES document[5] deals comprehensively with this problem, but illustrates that sampling protocols satisfying statistical requirements may be at odds with the realities of compliance testing.

---

2 Gollasch, S. and David, M. (2010). Testing Sample Representativeness of a Ballast Water Discharge and Developing Methods for Indicative Analysis. Report No. 4. Research Study. European Maritime Safety Agency, EMSA (2010). EMSA, Lisboa, Portugal.

3 Wright, D.A., R.W. Gensemer, C.L. Mitchelmore, W.A. Stubblefield, E. van Genderen, R. Dawson, C.E. Orano-Dawson, J.S. Bearr, R.A. Mueller and William J. Cooper (2010). Shipboard Trials of an Ozone-Based Ballast Water Treatment System. Mar. Pollut. Bull. 60, 1571-1583.

4 Miller, A.W., M. Frazier, G.E. Smith, E.S. Perry, G.M. Ruiz, and M.N. Tamburri, 2011. Enumerating Sparse Organisms in Ships' Ballast Water: Why Counting to 10 is Difficult?*Environ. Sci. Tech 45: 3530-3546.*

5 ICES. 2010. Harmful Aquatic Organisms in Ballast Water – Overview of statistical methods that could be used to verify compliance with the D-2 standard. ICES/IOC/IMO Working Group on Ballast and Other Ship Vectors (WGBOSV). ICES CM 2010/ACOM:65.

Endpoint determination

2.5 While the IMO D-2 standard will be the international standard when the Convention comes into force, it is not without problems or differences in interpretation. Usually, the >50 μm size category comprises mostly motile zooplankton, allowing viability to be determined from movement by all or part of the organism with or without physical stimulation using a probe. Not all organisms in this size category are motile, however; fish and crustacean eggs are cases in point, and several biologists involved with sample examination have employed a variety of vital stains to enhance live/dead determination. However, not all organisms take up the stain.

2.6 Vital stains have also been increasingly used in examining viability of the 10-50 μm size category which primarily comprises phytoplankton and protists, many of which are non-motile. Methods for determining live phytoplankton have included staining with such vital stains as fluorescein diacetate (FDA) and 5-chloromethyl fluorescein diacetate (CMFDA), as well as grow-out techniques which use growth potential as a determinant of viability. While vital stains have demonstrated utility in specific geographical areas[6] and have the potential of providing phytoplankton numbers that are compatible with the D-2 standard, overall, results have been mixed. The technique therefore remains speculative, particularly in view of the prospect of region-specific protocols for compliance testing with possible adverse legal consequences.

2.7 In view of the length of time required for grow-out, any analytical method that eliminates cell culture or grow-out as a means of determining viability merits examination. For phytoplankton, pulse-amplitude modulated (PAM) fluorometry holds this promise. In measuring the photochemical efficiency of photosystem II (F_v/F_m), it provides a rapid measure of photosynthetic activity as an indicator of cell viability, although results cannot directly be translated into cell numbers and are therefore incompatible with the D-2 standard. However, S. Gollasch and M. David[6] found a correlation of PAM readings with organism numbers, which may be used for the indicative sample analysis, as this delivers prompt results. Similar problems of interpretation are also associated with several assays used to determine viable indicator bacteria. Likewise the determination of biologically important molecules such as nucleic acids and adenosine triphosphate (ATP) are incompatible with D-2 standard endpoints and may indicate false positives, although such methods may be used for an indicative sample analysis. A comprehensive assessment of the methodologies involved, with a variety of biological endpoint determinations, is documented by S. Gollasch and M. David.[7]

2.8 Further complications result from the rigid size ranges defining published standards and the enormous quantitative differences associated with those size categories. In localities where large dinoflagellates comprise a significant proportion of the plankton flora, phytoplankton cell numbers in this (>50 µm) size range may exceed 10^7 per m^3.[8] At this density, a mortality rate of 99.9999999% would need to be recorded to comply with the >50 µm standard. It is clearly impossible to reach this degree of precision.

[6] Reavie, E.D., Cangelosi, A.A. and Allinger, L.E. (2010). Assessing Ballast Water Treatments: Evaluation of Viability Methods for Ambient Freshwater Microplankton Assemblages J. Gr. Lakes Res. 36:540-547. Steinberg, M. K., Drake, L.A. and Lemieux, E.J. (2011). Determining the viability of marine protists using a combination of vital fluorescent stains Marine Biology. DOI 10.1007/s00227-011-1640-8.

[7] Gollasch, S. and David, M. (2010). Testing Sample Representativeness of a Ballast Water Discharge and Developing Methods for Indicative Analysis. Report No. 4. Research Study. European Maritime Safety Agency, EMSA (2010). EMSA, Lisboa, Portugal.

[8] Wright, D.A. (2007) Problems associated with performance and compliance testing for ballast water treatment. Proc. Inst. Mar. Eng. Sci. Technol. J. Mar. Design Ops. (B12): 25-38.

3 Implications for compliance assessment and testing

3.1 Given the eventual need to evaluate compliance of many thousands of vessels in hundreds and perhaps thousands of ports, testing choke-points fall into two major categories: precision of sampling to D-2 standard and precision of live/dead determination to D-2 standard. Given limited worldwide analytical resources, two "end-members", the United States and Singapore, illustrate the nature of the challenge. In the United States, the primary logistical problem would relate to the coverage required to serve their 49 major ports separated by hundreds or thousands of miles, including the Great Lakes and Hawaii, given that analyses would need to begin within a few hours of collection, if microscopic examination of fresh samples is part of compliance assessment. In contrast to the United States, Singapore receives over 70,000 commercial vessels per year in a single port, not including barges, tugs, ferries and passenger vessels[9]; this is an average of more than 190 vessels per day. Turnaround times also vary enormously among different ports. The average turnaround time for ships in Singapore is between 6-8h, whereas in other ports, it may take more than 10 times as long for a similar type of vessel. Unlike the United States, port and presumably testing facilities in Singapore are relatively centralized, although in this case the impediment to full compliance testing relates to the sheer volume of traffic and impossibility of mounting the sampling and analytical effort required. In this regard, the problems facing testing facilities in other parts of the world fall between those illustrated by the United States and Singapore. As it now stands, even if only a very small percentage of vessels are selected for testing, resources available for full compliance testing fall well short of those required.

3.2 In view of the rapidly approaching need for standardized, worldwide compliance assessment, it is critical to examine the feasibility of available alternative options to complete compliance testing. Where penalties for non-compliance are likely to be very high, it is reasonable to assume that legal challenges are likely to ensue from ambiguities in the international standard(s) and their interpretation. D. M. King and M. N. Tamburri (2010)[10] and S. Gollasch and M. David (2010)[11] conclude that it will be extremely difficult and costly to directly assess whether a vessel can meet all the numerical standards for viable organisms, as published in the current D-2 standard, and they advocate a tiered approach to assessment, e.g., to start the compliance control sample processing with the "easiest-to-prove" organism group. This strategy recognizes that it will not be feasible to test more than a very small fraction of the world fleet at any given time. The proposed solution entails the use of reporting, inspections, and testing, involving a phased series of steps that increase the likelihood of detecting non-compliance but also increase cost and logistical challenges. Such a strategy is under development in the IMO Flag State Implementation (FSI) Sub-Committee. The rationale for such an approach is illustrated by the cost-effectiveness curve (Figure 1, King and Tamburri 2010[10]).

9 http://www.mpa.gov.sg/sites/pdf/vessel-arrivals.pdf

10 King, D.M. and Tamburri, M.N. (2010). Verifying Compliance with Ballast Water Discharge Regulations. Ocean Development & International Law, 41: 152–165.

11 Gollasch, S. and David, M. (2010). Testing Sample Representativeness of a Ballast Water Discharge and Developing Methods for Indicative Analysis. Report No. 4. Research Study. European Maritime Safety Agency, EMSA (2010). EMSA, Lisboa, Portugal.

4 Conclusions

4.1 Upon entry into force of the Ballast Water Management Convention, compliance assessment will form the basis of worldwide enforcement of the Convention. A potentially important component of this assessment will be shipboard compliance testing.

4.2 The ability of the Administrations, through the port State control process, to effectively verify compliance with the Convention will depend on the precision and reliability of data obtained on use and performance of BWMS. While any penalties for non compliance will come under the purview of the laws and requirements of individual countries, any process that streamlines and normalizes compliance testing will result in a more even application of the Convention worldwide. Ship operators will also welcome a common and understandable approach by port State control (PSC) in verifying compliance with the Convention.

4.3 Performance testing to date has related exclusively to certification against international standards, primarily the IMO D-2 standard. Trials have been conducted by relatively few testing centers worldwide. In order to reach the requisite degree of precision in demonstrating compliance with the standard, the level of sampling and analysis required is long and expensive.

4.4 It is estimated that it will not be possible to effectively use current performance testing for compliance monitoring worldwide. This is due to:

.1 lack of available, qualified testing personnel for real-time testing in hundreds/thousands of ports worldwide. Coverage would not be possible for full compliance testing of representative samples to the current standard; and

.2 time constraints relating to test procedures (few-several days) and to vessel turnaround time in port (few hours-few days). In many cases, full compliance information would not be available before the vessel left port.

5 Recommendations

5.1 We support a tiered approach to assessment, recognizing that it will not be feasible to test more than a fraction of the world fleet at any given time. The proposed solution entails the use of reporting, inspections, and testing, involving a phased series of steps that increase the likelihood of detecting non-compliance, but also increase cost and logistical challenges. Such a strategy would identify the most obvious cases of non-compliance, based on the rationale that it is much easier and cheaper to detect clear non-compliance than it is to identify full compliance, bearing in mind the sampling and analytical effort involved. The rationale for such an approach is illustrated by the cost-effectiveness curve (Figure 1, from D. M. King and M. N. Tamburri 2010[12]).

5.2 The most rudimentary step is the onboard port State control inspections by an enforcement official, in line with normal PSC practices, who would verify the certified treatment system's use, appropriate operation and maintenance of reports. This could be reinforced by pre-arrival reports submitted by vessel operators on the type of certified treatment system onboard and documentation indicating appropriate use and record of maintenance, which could help PSC target specific vessels for detailed inspection.

[12] King, D.M. and Tamburri, M.N. (2010). Verifying Compliance with Ballast Water Discharge Regulations. Ocean Development & International Law, 41: 152–165.

5.3 Verification of systems based on mandatory reporting and inspections of BWMS alone will not achieve acceptable levels of confidence that ballast water regulations are meeting their goals.[12] Verification of systems based on direct measurement (ballast water biological sampling) that are not comprehensive in terms of being both intensive (high volumes of ballast water sampled per vessel) and extensive (many vessels sampled) will not provide acceptable levels of confidence (Figure 1[12]). Those that are comprehensive enough to provide high levels of confidence have been estimated to be very expensive.[12]

5.4 In order to provide rigorous, legally defensible criteria for compliance verification, the international standard requires clarification through sampling guidelines to eliminate ambiguities relating to biological endpoints, including identification of zooplankton/protist groups, and to allow the use of autonomous devices to provide evidence of BWS usage and for determining viability of non-motile organisms. Autonomous measurements could include: particulate profile analysis and surrogate indicators of disinfection efficacy, e.g., total residual oxidant [TRO] and/or oxidation reduction potential [ORP] sensors for chlorine and ozone treatments; dissolved oxygen and/or pH sensors for deoxygenation treatments; and radiometers or measures of power output + water transmittance for UV treatments.[13] Indirect or indicative measures of abundances of live organisms may also be collected autonomously, or by inspectors, for indications of clear non-compliance (e.g., vital stains + flow cytometry, ATP kits, *in situ* PAM fluorometry).

5.5 Verification of systems based on indirect monitoring of ballast water using sensors has the potential to provide a high level of confidence at a cost that is far lower than even the lowest cost and least reliable biological sampling strategies. The success of a verification system based on sensors will depend on the development of accurate, reliable sensors that generate data that can withstand technical, statistical, and legal challenges.[12]

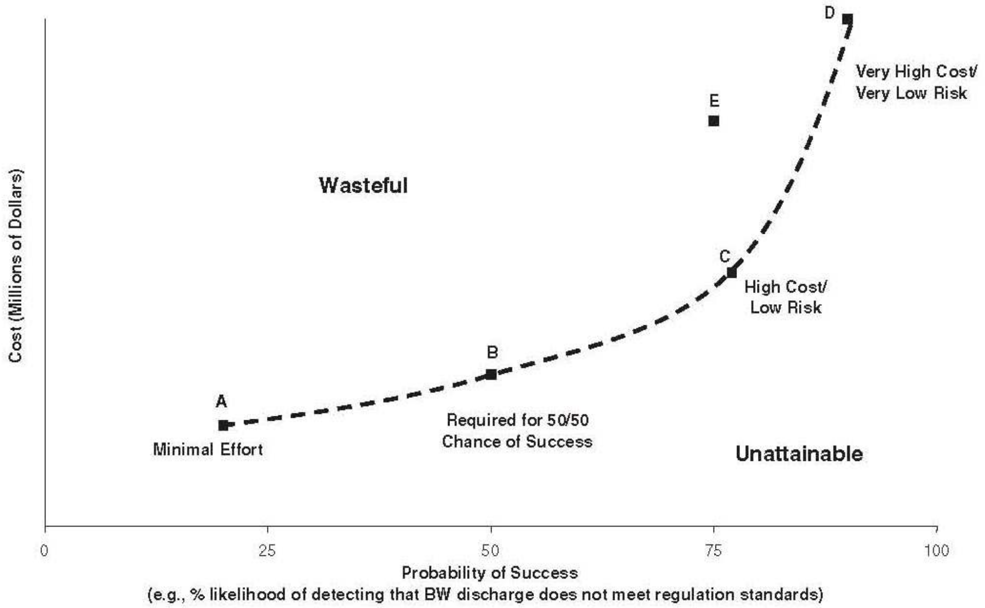

Figure 1: Cost effectiveness curve

(From D. M. King and M. N. Tamburri 2010[12] with permission)

[13] M.N. Tamburri personal communications.