Environmental Policy & Regulation

Anatomy of a Decision III: Evaluation of National Disposal at Sea Program Action Level Efficacy Considering 2 Chemical Action Levels

Sabine E Apitz,* † Chris Vivian, ‡ and Suzanne Agius§

†SEA Environmental Decisions, Hertfordshire, United Kingdom
‡Centre for Environment, Fisheries & Aquaculture Science, Suffolk, United Kingdom
§Environment and Climate Change Canada, Marine Protection Programs, Gatineau, Québec

ABSTRACT

The potential performance (i.e., ability to separate nontoxic from toxic sediments) of a range of international Disposal at Sea (DaS) chemical Action Levels (ALs) was compared using a sediment chemical and toxicological database. The use of chemistry alone (without the use of further lines of evidence) did not perform well at reducing costs and protecting the environment. Although some approaches for interpreting AL1 results are very effective at filtering out the majority of acutely toxic sediments, without subsequent toxicological assessment, a large proportion of nontoxic sediments would be unnecessarily subjected to treatment and containment, and a number of sublethally toxic sediments would be missed. Even the best tiered systems that collect and evaluate information sequentially resulted in the failure to catch at least some sublethally or acutely toxic sediments. None of the AL2s examined were particularly effective in distinguishing between non-, sublethally, or acutely toxic sediments. Thus, this review did not support the use of chemical AL2s to predict the degree to which sediments will be toxic.

Integr Environ Assess Manag 2017;13:1086–1099. © 2017 The Authors. Integrated Environmental Assessment and Management published by Wiley Periodicals, Inc. on behalf of Society of Environmental Toxicology & Chemistry (SETAC)

Keywords: Chemical action levels (AL) Contaminated sediments Dredged material management London Protocol Tiered assessment

INTRODUCTION

The objectives of many countries’ dredged material Disposal at Sea (DaS) Programs, and of related national and international legislation, mirror the 1996 Protocol to the London Convention (London Protocol) objective “to protect and preserve the marine environment from all sources of pollution and take effective measures, according to their scientific, technical, and economic capabilities, to prevent, reduce, and where practicable eliminate pollution caused by dumping or incineration at sea of wastes or other matter” (IMO 2016). The London Protocol, Article 1, paragraph 10 states that “…pollution means the introduction, directly or indirectly, by human activity, of wastes or other matter into the sea that results or is likely to result in such deleterious effects as harm to living resources and marine ecosystems, hazards to human health, hindrance to marine activities, including fishing and other legitimate uses of the sea, impairment of quality for use of seawater and reduction of amenities.” Thus, although the wording of the London Protocol suggests that its programmatic objectives focus on the prevention or elimination of pollution per se, the definitions of pollution imply that a substance only becomes a pollutant when and if it has unpleasant or harmful effects. Many potential pollutants have natural levels, harmless levels, or both, in sediment systems (Wenning et al. 2005). As such, although the objective of the protection of the marine environment from pollution may be considered a risk-based objective, the protection from “sources of pollution,” and the objective of eliminating pollution, appear much more absolute. National assessment frameworks for the evaluation of dredged material for its acceptability for DaS are informed to differing extents by risk-based and more absolute pollution prevention considerations.

For instance, many dredged material DaS frameworks use 2 chemical Action Levels (ALs), which can be lower levels (referred to as chemical Action Level 1 [AL1] in this article) or upper levels (referred to as chemical Action Level 2 [AL2]) that serve different purposes as explained in Box 1. Countries can set their ALs using only information about regional sediment chemistry, or they can use multiple lines of evidence in a risk-based approach (Apitz 2008, 2011) (Table 1). Regional chemical ALs are based on natural background levels, or, in the case of some constituents, on regional ambient levels...
also affected by ubiquitous human activities. This allows them to focus on the objective of elimination of pollution, in an absolute sense. Other tiered decision frameworks collect and consider information sequentially (usually starting with less complex or costly information and then progressing to more complex analyses). These tiered decision frameworks consider risk more directly and are specifically designed to evaluate lines of evidence (results from investigations of several sediment characteristics, e.g., physical, chemical, or biological) to determine whether constituents present in a sediment pose a risk to marine ecosystems and allow for both background-based evaluation (to compare contaminant levels with regional baseline conditions) and also risk- and bioavailability-based evaluation, to determine if constituents in place, whether natural or anthropogenic, pose risks to human health and the environment (Apitz and Power 2002; Chapman et al. 2002; Wenning et al. 2005; Barcelo and Petrovic 2007; Apitz et al. 2007; Agius and Porebski 2008; Apitz 2008, 2010a). The selection of the level and basis of ALs, and how they are used within the decision framework, are critical factors in how a decision framework supports program objectives (Apitz 2008, 2011); these differ within national programs to such an extent that detailed comparisons can be problematic (OSPAR 2008; Roper and Netzband 2011).

The Centre for Environment, Fisheries and Aquaculture Science (Cefas) trialed a tiered sediment assessment framework similar to those recommended by a range of international policies, reviews, or expert groups (Apitz and Power 2002; Chapman et al. 2002; Wenning et al. 2005; Barcelo and Petrovic 2007; Apitz et al. 2007; Agius and Porebski 2008; Apitz 2008, 2010a; CoA 2009) but adapted to United Kingdom and European priorities for the UK Environment Agency (Apitz, Crane et al. 2005) to assess its suitability on both a scientific and practical basis, based on scenarios of relevance to the Agency (Birchenough et al. 2006) based on a review of international best practice. They found that there were insufficient UK-specific colocated sediment chemical and toxicological data to sufficiently test the proposed framework and recommended extensive work, including the development of guidance on a UK approach for tiered bioassessment. Apitz, Crane et al. (2005) specified, in line with many international frameworks, that toxicity assessment should be applied to samples falling between AL1 and AL2 (and potentially high-cost sites falling above AL2), after considerations of factors controlling site-specific bioavailability, potentially followed by an evaluation of bioaccumulation potential.

If a tiered approach employing both AL1s and AL2s is applied to dredged material assessment decisions, it is often the case that sediments that have chemical levels below AL1 would be deemed to pose negligible risk, and marine licenses for unconfined disposal at sea are granted without further analysis (assuming all other DaS permit requirements are met). Between AL1 and AL2, a tiered assessment often examines lines of chemical, toxicological, and other evidence to determine whether contaminants present a risk. Above the AL2, sediments may go straight to a comparative risk assessment (CRA) to evaluate disposal options other than disposal at sea or may result in a refusal or withdrawal of a dredging license. Such frameworks allow for risk-based assessment and thus are designed to support risk-based (and not just background-based) decisions.

However, it should be noted that not all dredged material disposal at sea (DaS) frameworks, as they now stand, are explicitly tiered. Table 2 briefly summarizes some key aspects of a number of DaS frameworks from Europe and Canada. Although samples below AL1 are generally considered acceptable for unconfined disposal at sea, pending other considerations such as physical suitability for the disposal site and potential beneficial uses (with the exception of the Netherlands that uses a AL2 for that decision), and sediments above AL2 are considered unacceptable for unconfined disposal at sea (i.e., they require special handling and/or containment), procedures on how sediments that fall between AL1 and AL2 should be handled, or what LOEs should be considered to evaluate these samples are less consistent. In the United Kingdom and Canada, this is largely left to the scientific judgement of case officers. Although Cefas in the United Kingdom has carried out work to develop sediment toxicity testing (Cefas 1993, 2002), such approaches have not been formally adopted. Further assessment is required for these samples in Denmark, Finland, Ireland, and Norway, but what sort of assessment is required is not specified in review documents. Other
Table 1. Chemical action levels used in this study

| Action list | Units | BLG | DK | FIN | FR | GER | IRE | NL | NOR | POR | SP | UK | CAN | Consen |
|-------------|-------|-----|----|-----|----|-----|-----|-----|-----|-----|----|----|-----|--------|
| Action Level 1 | As mg/Kg | 20 | 20 | 15 | 25 | 40 | 9 | 52 | 50 | 80 | 20 | 16 |
| | Cd mg/Kg | 2.5 | 0.4 | 0.5 | 1.2 | 1.5 | 0.7 | 2.6 | 3.0 | 1.0 | 0.4 | 0.6 | 1.0 |
| | Cr mg/Kg | 60 | 50 | 65 | 90 | 120 | 120 | 560 | 100 | 200 | 40 | 63 |
| | Cu mg/Kg | 20 | 20 | 50 | 45 | 30 | 40 | 51 | 150 | 100 | 40 | 33 |
| | Hg mg/Kg | 0.30 | 0.25 | 0.10 | 0.40 | 0.70 | 0.20 | 0.63 | 1.50 | 0.60 | 0.30 | 1.0 |
| | Ni mg/Kg | 70 | 30 | 45 | 37 | 70 | 21 | 46 | 75 | 100 | 20 | 37 |
| | Pb mg/Kg | 70 | 40 | 40 | 100 | 90 | 60 | 83 | 150 | 120 | 50 | 60 |
| | Zn mg/Kg | 160 | 130 | 170 | 276 | 300 | 160 | 360 | 600 | 500 | 130 | 187 |
| | TBT μg/Kg | 3 | 7 | 3 | 3 | 100 | 20 | 100 | 5 | 100 | 31.07 |
| | PCB7 μg/Kg | 50 | 20 | 22 | 297 | 13 | 7 | 17 | 25 | 30 | 10 | 34 |
| | PAH16 μg/Kg | 2542 | 4789 | 5426 | 6700 | 1800 | 4000 | 2892 | 2000 | 3712 | 2500 | 1730 |
| | Lindane μg/Kg | 0.5 | 1.1 | 0.3 | 5.00 | 0.19 | |
| | Dieldrin μg/Kg | 1.8 | 0.3 | 17.0 | 2.5 | 0.3 | |
| | HCB μg/Kg | 10 | 4 | 20 | 1 | 3.95 |
| | DDT μg/Kg | 7 | 200 | 200 | 400 | 250 | 20 | 1000 | 59 |
| | 2,3,7,8-TCDD TEQ | 0.00 | |
| | Chlorodane | 0.32 | |
| | Aldrin | 0.06 | |
| Action Level 2 | As mg/Kg | 100 | 60 | 60 | 50 | 120 | 70 | 29 | 76 | 500 | 200 | 100 | 66 |
| | Cd mg/Kg | 7.0 | 2.5 | 2.5 | 2.4 | 4.5 | 4.2 | 4.0 | 15.0 | 10.0 | 5.0 | 5.0 | 5.1 |
| | Cr mg/Kg | 220 | 270 | 270 | 180 | 360 | 370 | 120 | 5900 | 1000 | 1000 | 400 | 255 |
| | Cu mg/Kg | 100 | 90 | 90 | 90 | 90 | 110 | 60 | 55 | 500 | 400 | 400 | 123 |
| | Hg mg/Kg | 1.50 | 1.00 | 1.00 | 0.80 | 2.10 | 0.70 | 1.20 | 0.86 | 10.00 | 3.00 | 3.00 | 1.50 |
| | Ni mg/Kg | 280 | 60 | 60 | 74 | 210 | 60 | 45 | 120 | 250 | 400 | 200 | 121 |
| | Pb mg/Kg | 350 | 200 | 200 | 200 | 270 | 218 | 110 | 100 | 600 | 500 | 238 |
| | Zn mg/Kg | 500 | 500 | 500 | 552 | 900 | 410 | 365 | 590 | 5000 | 3000 | 800 | 496 |
| | TBT μg/Kg | 7 | 200 | 200 | 400 | 300 | 500 | 250 | 20 | 1000 | 59 |
| | PCB7 μg/Kg | 50 | 200 | 210 | 605 | 40 | 1260 | 100 | 190 | 300 | 100 | 140 | 178 |
| | PAH16 μg/Kg | 6478 | 47 805 | 44 410 | 23 392 | 5500 | 11 481 | 8618 | 20 000 | 12 760 | 6610 |
| | Lindane μg/Kg | 1.50 | 2.20 | 1.0 | 4.0 |
| | Dieldrin μg/Kg | 5.5 | 1.0 | 20.0 | 61.0 | 50.0 | 1.0 |
| | HCB μg/Kg | 30 | 12 | 20 | 490 | 29.7 |
| | DDT μg/Kg | 7 | 200 | 200 | 400 | 300 | 500 | 250 | 20 | 1000 | 59 |
| | 2,3,7,8-TCDD | 3.6 | |
| | Chlordane | 0.32 | |
| | Aldrin | 0.06 | |

BLG = Belgium; DK = Denmark; FIN = Finland; FR = France; GER = Germany; IRE = Ireland; NL = the Netherlands; NOR = Norway; POR = Portugal; SP = Spain; UK = United Kingdom; CAN = Canada (including “corrected” PCB number as per Apitz and Agius [2013]); Consen = Consensus values developed in Apitz and Agius (2013). Basis of ALs: BK, background-based; BK’N, based on a multiple of the background-based value; ETX, ecotoxicology-based (may be mechanistic or empirical); Mixed, based on a mixture of approaches. DK AL1 values are based on a range of sediment guideline values used in Finland and the Netherlands, not only their AL1 values.

*CAN IPCB values are 100 ppb (arachlor); adjusted to ICES7 congener values as in Apitz and Agius (2013) and in Data S2. It should be noted that since this work was carried out, Spain has revised its Action Levels including moving from using a <63 μm sediment fraction to a <2 mm sediment fraction (OSPAR 2015).
countries specify bioassessment, monitoring, or impact assessment. Thus, other than identifying which sediments pose negligible risk, AL1 has differing impact on dredged material assessment decisions in various countries, and comparisons can be difficult.

The Canadian DaS framework, as it now stands, is tiered, but has only 1 chemical AL; an AL1. Although samples below AL1 are generally considered acceptable for unconfined ocean disposal (pending other considerations such as physical suitability for the disposal site and potential beneficial uses), upper action levels are based on other lines of evidence; there is no chemical AL2 above which sediments are rejected for ocean disposal without further assessment (rather, Canadian AL2s are based only on toxicity test results). One of the recommendations of a workshop to review the Canadian DaS approach recommended the consideration of chemical AL2 values (Agius and Porebski 2008). Subsequent studies (Apitz 2008, 2010a; Apitz and Agius 2013, this issue) have looked at various aspects of this recommendation; to evaluate potential effects of such a change, Apitz and Agius (2013) developed and examined a large database of North American coastal sediment data and determined that the choice of which constituents are included in a chemical action list was the dominant influence (of those evaluated) on AL decision rule outcomes. Apitz and Agius (this issue) combined this database with colocated sediment toxicological data and concluded that efficacy of AL1s in predicting acute and sublethal toxicity depended on the parameters comprising the chemical action list and chemical decision rules and found that the addition of a screening bioassay could enhance Tier 1 efficacy.

Following work in the United Kingdom and Canada suggesting a range of modifications to their DaS frameworks, this study used the database of colocated sediment toxicological data and concluded that efficacy of AL1s in predicting acute and sublethal toxicity depended on the parameters comprising the chemical action list and chemical decision rules and found that the addition of a screening bioassay could enhance Tier 1 efficacy.

Table 2. Summary of national DaS actions and decisions as a result of Action Level outcomes as summarized from Roper and Netzband (2011)

| Country | <AL1 | AL1-AL2 | >AL2 | Notes |
|---------|------|---------|------|-------|
| BLG     | DaS  | Further sampling; bioassessment | No DaS | Three AL2 failures for no DaS |
| DK      | DaS  | Further assessment; unspecified | Land disposal in most cases | |
| FIN     | DaS  | Case-by-case; unspecified | No DaS | |
| FR      | DaS  | Bioassessment (depending on level of exceedance) | Possibly no DaS; in-depth impact survey | Used in cases where AL2 is exceeded for one substance or in case of level between AL1 and AL2 when “GEODRISK” score exceeded 1 |
| GER     | DaS  | DaS with impact assessment | Containment and/or treatment; impact assessment; source ID | Even for lower levels, bioassay for placement decisions |
| IRE     | DaS  | Further assessment; unspecified | Containment and/or treatment | |
| NL      | N/A  | <AL2, DaS | Strict limit | Priority substances are strict limits with no exception; up to 50% permitted for others |
| NOR     | DaS  | Further assessment; unspecified | Containment and/or treatment and/or monitoring | |
| POR     | DaS  | DaS with monitoring | Landfill | 5 levels; equivalent 2 selected |
| SP      | DaS  | DaS with impact assessment and monitoring | Treatment and/or containment | |
| UK      | DaS  | Further assessment; unspecified | Containment and/or treatment; additional investigation | |
| CAN     | DaS  | >AL1 bioassessment; special handling if passes; containment and/or treatment if fails | N/A | |

BLG = Belgium; DK = Denmark; FIN = Finland; FR = France; GER = Germany; IRE = Ireland; NL = the Netherlands; NOR = Norway; POR = Portugal; SP = Spain; UK = United Kingdom; CAN = Canada.

It should be noted that information available differs for each country, and this simplified table cannot capture the detail of frameworks but only summarizes key points for a general comparison. For more detailed interpretation, specific national guidance should be examined.
assist Canada, the United Kingdom, and other countries implementing the London Protocol to evaluate regulatory decisions on the disposal of wastes at sea, and what these decisions mean to the environment and risks to human health and to permittees in terms of necessary and unnecessary costs.

The rest of the article is organized into Study Approach, Methods, Results & Discussion, and Conclusions. The study approach describes the overall approach and assumptions underlying the comparisons described in this study. The methods describe the development of: chemical action lists and the database of sediment data; specific AL values (or benchmarks) for contaminants on the chemical action list; result classifications as to the suitability for unconfined disposal at sea for specific chemical analytes; and overall, using a simplified, streamlined tiered decision making framework. Results are organized into comparisons of AL1 classifications and AL2 classifications. The conclusions summarize the findings and address limitations and further research opportunities.

STUDY APPROACH

This study seeks to evaluate and compare the efficacy of chemical-only ALs (both AL1 and AL2 when available) from a range of countries. Although, as described above, DaS frameworks can be focused on both pollution and risk prevention, in this study, efficacy of a AL was defined in terms of its ability to predict toxicity (both sublethal and acute). In terms of a tiered assessment framework, efficacy is measured in terms of framework outcomes, as will be described below. The evaluation of the efficacy and fitness for purpose of ALs and decision rules, in terms of their ability to identify potentially toxic sediments, requires the application of a tiered evaluation approach. Thus, this study seeks to evaluate the performance of a range of international ALs assuming a tiered approach, as applied in most OSPAR countries, with the proviso that the details of national approaches differ, as described above. However, the actual potential regulatory outcomes using the chemical action levels evaluated will be highly dependent on which and how other LOEs are applied within national frameworks when ALs do not provide unambiguous answers, particularly when samples fall between AL1 and AL2. Thus, various potential approaches and scenarios were evaluated in a comparable manner and are discussed in support of a high-level review, but the specific regulatory implications of various frameworks would require detailed, nation-specific analysis and data tailoring, which is outside the scope of this work.

The application of both AL1 and AL2 allows for a separate evaluation of AL selection at 2 levels in the decision process. As these 2 levels have different purposes and interpretive goals, the best choice of AL type for 1 level will not be the same as that for another. There have been extensive reviews and recommendations on the use of ALs and sediment quality guidelines, criteria, or values (SQGs, SQCs, or SQVs, respectively) for sediment and dredged material management internationally (Cefas 2003; 2005; Apitz, Crane et al. 2005; Wenning et al. 2005; Birchenough et al. 2006; Apitz et al. 2007; Apitz 2008, 2010b, 2011; Apitz and Agius 2013; OSPAR 2014). One of the best ways to evaluate ALs is to evaluate how effectively they perform (i.e., predict sediment toxicity) in a statistically significant number of sediments with a diverse range and level of contaminant mixes, sources, and mineralogies. Such an evaluation requires colocated chemical and toxicological data from a significant number of sites; such data should not be biased by having been collected due to specific regulatory aims, but it should be representative of the range of scenarios that may be encountered by a regulatory program.

Database development

Developing such a data set can be a very expensive prospect, so the most practical approach in developing a database is to use data mining. To evaluate Canadian DM frameworks, Apitz and Agius (2013, this issue) have developed such a database, with over 1000 records containing colocated chemical and toxicological data from coastal sediments. Using this database, they compared the efficacy of current Canadian ALs and both chemical and toxicological decision rules (that together form Canada’s Action Level guidance) and to potential modifications of the Canadian approach (Apitz and Agius 2013, this issue). They found that the efficacy of ALs depended upon a range of parameters, including: 1) the chemical action list (i.e., the list of chemical constituents, chemical action list), 2) the AL values, 3) the decision rules applied for cAL pass and/or fail, and 4) the relationship between AL and other LOEs in an overall decision framework. Database details, and how the database was adapted for this project, can be found in the Supplemental Data S1.

Simplified, tiered decision making framework

In a tiered decision framework that begins with chemical assessment, bioassessment (called Tier 2 in this work) should only be applied to reduce uncertainty (e.g., for samples falling between AL1 and AL2 and potentially high-cost sites falling above AL2) after considerations of factors controlling site-specific bioavailability and potentially followed by an evaluation of bioaccumulation potential. The data available for this review does not, however, allow for the consideration of regional background, mineralogical, or other factors controlling contaminant bioavailability, nor can it consider bioaccumulation potential. Thus, although the assumption of a tiered approach will underlie this review, it will only evaluate potential outcomes considering the combined use of ALs and selected approaches to toxicity assessment. Thus, to carry out this comparison, a simplified, tiered decision scheme was applied (Figure 1). In this approach, chemical data in the database were evaluated using national chemical action levels. Bioassessment data were evaluated using a uniform bioassessment decision scheme, and then regulatory outcomes were evaluated using the tiered decision scheme in Figure 1. Outcomes might differ if site-specific chemical
assessment were carried out to consider background conditions or potential confounding factors; these outcomes could be subject to significant regional variation, as sediment compositions and background levels can differ greatly between samples, sites, regions, and countries, as has been demonstrated in the United Kingdom (Cefas 2005) and elsewhere. Furthermore, specific national outcomes will differ when their specific bioassessment and tiered decision schemes are applied; this approach simply compares the efficacy of ALs in a uniform scheme.

The approach taken in this article is based on the assumption that the Tier 2 bioassessment is the baseline measure of “truth”—that toxicity measured in field sediments evaluate whether the chemical benchmarks are appropriate. This assumption is necessary to provide some basis of comparison, but it is acknowledged that the toxicity measures themselves contain uncertainty. Bioassessment can be subject to a range of confounding factors and may be sensitive to unmeasured contaminants in the sediments; errors or misclassifications can occur in any measurement (Wenning et al. 2005; Apitz 2011). The bioassays evaluated here were those for which data were available in the database and that matched, as far as possible, the battery of bioassays used in the Canadian disposal at sea program (this is discussed in more detail in Anatomy II (Apitz and Agius this issue)).

Representativeness of data set

Given the high degree of site-specific variability in sediment characteristics likely to be encountered by a regulatory program, the key assumption behind the use of any database to evaluate potential regulatory outcomes or fitness for purpose of a national DaS framework is that the characteristics (range and concentrations of chemicals, toxicity, mineralogy, etc.) are representative of those that might be encountered by that regulatory program. The large and diverse database used in this study is from around the coasts of the United States (a country that leads in dredging and sediment assessment, but that incidentally, has elected not to promulgate national ALs and thus could not be considered in this review—the United States relies on acute toxicity tests as their AL2 and other lines of evidence [e.g., bioaccumulation] to assess sublethal effects). This is assumed to be a reasonable proxy for the range of potential outcomes that may be encountered by other national DaS programs for the reasons outlined below. However, specific national outcomes cannot be predicted.

It is generally recognized that the toxicity of chemicals in sediments are affected by contaminant source, type, and postdepositional history, as well as by sediment mineralogy and organic matter content, geochemistry, and water chemistry (Wenning et al. 2005). Furthermore, toxicity of a given sediment varies as a function of toxicity test organism, exposure mechanism, test conditions, and pass and/or fail criteria (Wenning et al. 2005). Thus, an ideal approach is the development of national ALs using local species and validated with national sediments. However, even within 1 country or region, contaminant bioavailabilities or toxicities may be highly variable; thus, it is possible that a very small region-specific data set may be no more representative than a larger, broader, but less regionally specific data set. The size of the database used in this study, and the diversity of sites, contaminant sources, levels and bioavailabilities, and sediment types that it draws from make it a reasonable proxy for the range of sediment, contaminant, and toxicity combinations that might be encountered in other programs. However, the fact that it is based on another region and nonnative species means that, although broad conclusions can be drawn, small differences between approaches or scenarios should be viewed with caution, especially when applying conclusions to another region.

This high-level review should provide a reasonable comparative assessment of AL approaches and can provide a basis for recommendations for a path forward to refine existing dredged material decision frameworks, but ultimately, the actual effects of framework changes on national or regional DaS applications will need to be evaluated over a period of time with national data and approaches.

METHODS

Chemical action list and database development

This project pursued a “data mining” approach to evaluate potential refinements to DaS frameworks by compiling sediment chemistry and toxicity data sets and subjecting them to a series of decision protocols. The objective was to develop a data set of marine, coastal, and estuarine sediment chemical and toxicological analytical results that were representative of the range of sediment types and contaminant combinations and levels that might be encountered by a DaS Program. Only samples that had results, at a minimum, for some metals, PAHs, and PCBs, and data from as many other analytes and co-associated bioassays as possible were included in the data set. Data used in the database were from National Oceanic and Atmospheric Administration (NOAA)
Status and Trends (NS&T) and Mussel Watch data sets available online (NS&T 2012) and also an extensive data set of sediment chemistry and toxicity from Pearl Harbor, HI that had already been extracted from a report (Ogden Environmental and Energy Services Company 1997, 1998) and had, in part, been previously used for another project (Apitz et al. 2007). The database was developed as described in Apitz and Agius (2013, this issue); the same chemical action lists, toxicity data, and decision rules that were used in these articles are used here, although some constituent summations (and ALs, see below) have been adapted for this more international comparison. See Data S1 for additional details.

Development of AL1 and AL2 values for comparison

Environment and Climate Change Canada carried out a review of international DaS ALs and other sediment quality guidelines. This was revised and updated for Apitz et al. (2014) and Apitz (2014). Table 1 lists the international AL values used in this study. Details on these ALs and their adaptation for this study can be found in Data S1 and Data S2.

To examine the potential regulatory effects of a broader chemical action list in Canada, Apitz and Agius (2013) developed a set of hypothetical “consensus” ALs based on either international DaS ALs (where available) or risk-based sediment quality guidelines (for constituents that were not included in the DaS chemical action lists reviewed at the time). Given the fact that the current review provides a broader range of international AL values, it should be possible to generate new consensus values, but the ones used in Apitz and Agius (2013, this issue) will be used for continuity. It should be noted, however, that it is not the intent of this work or of previous studies to suggest that these values should be adopted; they have been used to illustrate the potential impacts of differences in DaS approaches in a consistent manner. The international and consensus ALs are compared in Data S2.

Classification of AL outcomes using database

To evaluate potential outcomes and “efficacy” of ALs, it is necessary to evaluate how effectively ALs predict measured sediment toxicity.

AL decision rules

International chemical action lists and ALs were reviewed and adapted to be as equivalent and comparable as possible (see Data S1 and S2 for details). AL quotients were calculated for all samples and constituents to be considered. For contaminant i, and sample j,

\[
\text{AL quotient}_{ij} = \left(\frac{C_{ij}^{\text{AL}}}{C^{\text{AL}_1}}\right)
\]

where \(C_{ij}^{\text{AL}}\) is the concentration of contaminant i in sample j and \(C^{\text{AL}_1}\) is the AL1 or AL2 of interest. If this quotient value is greater than 1, then the sample has “failed” for that chemical. For this study, a one out, all out rule is applied (Apitz 2011) such that if a sample fails for 1 chemical, it is considered to have failed that overall AL. There are a range of other decision rules that can be applied in a framework (Apitz 2008, 2011; Apitz and Agius this issue), all of which can be tested using the database, but this is the simplest approach and the one currently being applied in most OSPAR countries and in Canada. Chemical outcomes were then divided into 3 categories:

- AL1: all chemicals in the Action List are at levels below their AL1s;
- For AL test protocols with consensus ALs and screening Microtox, the AL1 pass requires a pass of all chemicals AND passing a screening-level Microtox bioassay (the same bioassay that is used in the Tier 2 bioassessment suite, but with a more conservative pass and/or fail criterion; see Apitz and Agius [this issue])
- AL1-AL2: at least 1 chemical in the Action List is above its AL1, but all are below their AL2; or
- AL2: at least 1 chemical in the Action List is above AL2.

Bioassay decision rules

As with chemical assessments, there are a range of approaches to interpreting toxicity data: in terms of which bioassays are used, pass and/or fail thresholds, and how the outcomes of multiple toxicity assessments are combined to generate an overall toxicity decision. For this high-level review, the approach developed in Apitz and Agius (this issue) was used. In this approach, toxicity data in the database were adapted and interpreted to be as similar as possible to the bioassay battery and decision rules used in the Canadian program, using 2 sublethal bioassays and an acute bioassay (see Supplemental Data S1). Three levels of toxicity are designated:

- Negligible toxicity, or nontoxic: the sample passes both sublethal bioassays and the acute bioassay
- Non-negligible toxicity, or sublethally toxic: the sample fails 1 sublethal bioassay
- Acutely toxic: the sample fails both sublethal bioassays and/or the acute bioassay

Within the Canadian DaS framework, these 2 designations of toxicity (sublethal and acute) result in different regulatory outcomes (see Apitz and Agius this issue, for a discussion). Toxicity LOEs and decision rules differ for each country compared here, but, for the purposes of this review, it can be considered that a designation of negligible toxicity (also called nontoxic here) suggests that a sediment poses no risk to the marine environment and is most likely suitable for unconfined ocean disposal (barring any other issues); nontoxic samples make up approximately 75% of the database. It can also be assumed that a designation of acutely toxic poses probable risk to the marine environment and would be barred from unconfined ocean disposal. Acutely toxic samples make up 5% of those in the database.
On the other hand, the regulatory outcome from designation of nonnegligible, or sublethal, toxicity is less clear-cut. Such sediments may pose some risk to the marine environment but less than acutely toxic sediments. Thus, the Canadian framework designates that sublethally toxic sediments can be disposed of in the open ocean but with special handling. Sublethally toxic samples are 20% of the database records, but it should be noted that samples that fail one sublethal bioassay are more likely to be responding to confounding factors (e.g., grain size, organic matter) than those that fail 2 sublethal or an acute bioassay (this is discussed briefly in the supplemental materials of Apitz and Agius this issue). Thus, although this analysis has the potential to be more nuanced, it can also be subject to some artifacts, some of which might be controlled for in a tiered or broader assessment, but are not addressed in this review.

In this study, overall outcomes with different action lists and ALs will be compared to toxicological outcomes, with the assumption that toxicological designations represent “truth”; the ability of ALs to correctly “predict” this toxicity will be considered a measure of their efficacy. It is fully recognized that toxicology tests are actually also subject to a range of errors and uncertainties (Wenning et al. 2005; Apitz 2011) and are often selected for regulatory use because they correlate well with chemistry. However, this assumption was used to allow for a comparative assessment of potential AL efficacy.

Clearly, the application of different toxicity assays, decision rules, and tiered approaches in different countries may change these outcomes or pass and/or fail thresholds. However, a country-by-country comparison would require country-specific tailoring of the database, which is outside the scope of this project. Rather, the effects of a range of chemical action lists and action levels in various countries, when applied to a consistent assessment of toxicity and a uniform, simplified, tiered framework will be examined.

**Overall chemical and/or biological classifications**

International AL outcomes and performance was examined using the database by evaluating framework chemical and/or toxicological classification rates (AL efficacy, based on sensitivity, efficiency and specificity are discussed briefly in Data S3). This approach looks at “tiered” classifications using the uniform simplified decision scheme described in Figure 1. Given the 3 chemical (<AL1, AL1-AL2, >AL2) and 3 toxicological (nontoxic, sublethally toxic, and acutely toxic) outcomes, there are 9 potential paired chemistry and/or bioassessment Classes I to IX. In the tiered decision framework, these classes can ultimately result in 3 “regulatory” outcomes—DaS permitted, special handling, or DaS refused. The 9 potential decision classes, their regulatory outcomes, and their meanings, are illustrated in Figure 2. These classifications, when applied to the database, provide insight into the relative rates at which each chemistry outcome results in each toxicity outcome. Details on chemical and bioassay pass and/or fail decision rules can be found in the Methods section above.

As noted above, the actual regulatory meaning of each decision class depends on the national approach applied by each country. Table 2 provides a brief summary of national DaS actions and decisions as a result of AL outcomes. The more general implications of a range of approaches are discussed for each decision class.

**Class I (nontoxic AL1 true negative).** In almost all cases, assuming that a national DaS approach would apply similar toxicity rules as in this case example, this classification reflects a successful AL1 application. The AL1 has successfully “predicted” a nontoxic sediment, which, in most cases, will be correctly designated as acceptable for unconfined open

---

Figure 2. Sediment classifications and regulatory outcomes as a result of paired chemical and bioassessment outcomes.
water disposal (assuming all other DaS permit requirements are met). Unnecessary further analysis is not required; a DaS framework should seek to maximize Class I occurrences (for nontoxic sediments), as they are both cost-effective and protective.

Class II (sublethally toxic AL2 false negative). The meaning of this classification depends on the national DaS protocol. If sediments passing AL1 are approved for unconfined open water disposal without further toxicological assessment (as is the case in many frameworks that are similar to that in Figure 1; see Table 2), then a sublethally toxic sediment will be approved for disposal without confinement. If a national protocol considers this bioassay failure an indicator for a sediment’s lack of suitability for such disposal, then this AL1 can be seen to be insufficiently protective. Thus, it can be argued that in many cases, DaS frameworks should seek to minimize Class II occurrences, although it can be claimed that such failures are less serious than Class III failures described below.

Class III (acutely toxic cAL1 false negative). If sediments passing AL1 are approved for unconfined open water disposal without further toxicological assessment (as is the case in many frameworks that are similar to that in Figure 1), then in this case acutely toxic sediment will be approved for disposal without control. Class III occurrences should be minimized as these pose the greatest risk to the marine environment.

Class IV (nontoxic AL1 false positive). In this classification, AL1s fail nontoxic sediments. In a tiered DaS framework such as the one in Figure 1, such sediments, after undergoing further toxicological assessment, may be passed, and thus will be approved for unconfined open water disposal. In such a case, the tiered assessment can be seen as successful as nontoxic sediment is not unnecessarily subject to treatment or containment, but costs are incurred during the application of toxicity assessments. On the other hand, if no further bioassessment is carried out (as in Portugal and Spain), or if protocols for sediments falling between AL1 and AL2 are not clearly laid out (as in the UK, Denmark, Finland, Norway, and Ireland), then it is possible that these nontoxic sediments will be unnecessarily subject to control, treatment or containment, or the refusal or withdrawal of a DaS permit.

Class V (sublethally toxic AL1 true positive). These sediments, which fail AL1, are sublethally toxic. If a DaS framework requires toxicity assessment for such AL1 failures, and if the toxicity thresholds are similar to those in this review, then these sediments can be seen to be correctly assigned. If toxicity thresholds are higher, then the outcome will be similar to that for Class IV. In any case, Class V outcomes can be seen as successful framework outcomes. For countries with an unspecified approach to samples falling between AL1 and AL2, outcomes for these samples are unclear.

Class VI (acutely toxic AL1 true positive). As with Class V, these acutely toxic sediments that correctly fail AL1 can be seen as successful designations in most frameworks. If bioassessment is carried out for these samples, uncontrolled DaS will be refused. For countries with an unspecified approach to samples falling between AL1 and AL2, outcomes for these samples are unclear.

Class VII (nontoxic AL2 false positive). These sediments have failed AL2 but are nontoxic according to the bioassay battery. If a DaS framework requires toxicity assessment for all sediments that fail AL2, it is possible that these sediments will ultimately be identified as nontoxic and permitted for unconfined open water disposal (or they may be subject to some restrictions due to elevated chemical levels even if nontoxic). On the other hand, in a framework that designates all sediments that fail AL2 as requiring special handling, treatment, or containment (as do most in Table 2), this classification will result in the potentially unnecessary expense of special handling, treatment, or containment of nontoxic sediments or disposal on land even when this is not the environmentally preferable option. Such an outcomes results not only in unnecessary expense and emissions but also in the loss of resources in the form of disposal space and clean sediments for beneficial use and a disruption of in-water sediment balance (Apitz 2010b).

Class VIII (sublethally toxic AL2 true positive). In this case, sublethally toxic sediments are correctly failed by AL2; the outcome of this classification depends on the broader DaS approach, but this can be seen as an indicator of a successful AL2, if sublethal toxicity is deemed to be of concern. On the other hand, this could also be seen as overprotective, as, depending on national approaches, it may result in sediments that could be disposed of at sea with monitoring (e.g., Norway) or special handling (e.g., Canada) being refused permits for disposal at sea.

Class IX (acutely toxic AL2 true positive). In this case, acutely toxic sediments are correctly failed by AL2; the outcome of this classification depends on the broader DaS approach, but this can be seen as an indicator of a successful AL2, as acutely toxic sediment is most likely prevented from being disposed of without treatment or containment.

RESULTS AND DISCUSSION

Figure 3 compares the sediment classifications using combined chemical and toxicological outcomes, as defined in Figure 2, for the range of national and hypothetical ALs considered in this review. In this scheme, Classes I to IV would be approved for DaS, Class V would require special handling, and Classes VI to IX would be refused DaS and require special handling, treatment, disposal, or containment. When considering the relative proportions of classifications reported in this study, and whether these rates are acceptable or not from a policy perspective, it should be noted that acutely toxic sediments are only approximately 5% of the samples in the
database; sublethally toxic sediments are approximately 20% of the samples in the database, and nontoxic sediments are approximately 75% of the samples in the database.

**Comparison AL1 classifications**

As described above, in the context of a 2-tiered assessment framework for dredged material, indicators of a “successful” AL1 are a minimization of Class II and III rates (that in most cases described in Figure 2 would result in sublethally or acutely toxic sediments being permitted for unconfined DaS), while balancing a secondary objective of minimizing Class IV classifications, which can result either in further assessment or can result in nontoxic sediment requiring control, monitoring, containment, or treatment. Based on these criteria, the relatively conservative (i.e., low) UK AL1 values have the lowest Class II levels of any national framework reviewed, with Class III levels being matched by Denmark and with Belgium, Finland, and Ireland having only slightly higher levels. The hypothetical consensus values of Apitz and Agius (2013), with their relatively conservative values and extended chemical action list, meet the UK Class III rate, and have a relatively low Class II rate. However, all of the AL1s reviewed in this study still miss over 10% of the acutely toxic samples and 25% or more of the sublethally toxic samples. Furthermore, these low Class II and III rates come at the expense of relatively high Class IV rates, especially in the UK framework. The frameworks of France, Germany, the Netherlands, Norway, Portugal, Spain, and Canada, either due to relatively unconservative (i.e., high) AL1s or short chemical action lists, are much less effective, missing up to 35% of the acutely toxic sediments and up to 83% of the sublethally toxic sediments. It should be noted, however, that the relative assignments of sublethally toxic sediments should be viewed with some caution, as these sediments (that have failed a single sublethal bioassay), may be more sensitive to confounding factors than to contaminants of interest. The inability to examine confounding factors (e.g., grain size and ammonia) in this database may have resulted in an inflated number of toxic results and a source of uncertainty in relative proportions of regulatory outcomes. If confounding factors could have been tested for, as is the case with many tiered frameworks, a proportion of these sublethally toxic sediments might be re-assigned as nontoxic, improving the perceived performance of the Class I and II assignments (this was briefly discussed in Apitz and Agius this issue). The Netherlands is a special case, as, with no cAL1, the cAL2 performs the role of an AL1 in the other frameworks. As such, the Dutch AL2 was the poorest performing AL, with the highest rates of toxic samples potentially being approved for unconfined DaS.

**Comparison cAL2 classifications**

If the objective of a AL2 is to identify toxic samples that should be refused DaS, and most likely be subject to containment, treatment, or upland disposal, potentially without further assessment, then an effective AL2 should have relatively high rates of Class VIII and IX samples (as these samples are sublethally or acutely toxic), and low rates of Class VII samples (as these nontoxic samples may require unnecessary treatment or containment). By these criteria, none of the AL2s can be deemed particularly effective, as they proved poor discriminators of toxic from nontoxic sediments. Less conservative AL2s (e.g., Portugal, Spain, and the UK) failed fewer nontoxic sediments, but also failed a considerably lower proportion of the acutely and sublethally toxic sediments. The most conservative AL2s (Germany, the
Netherlands, consensus) failed the highest proportion of toxic sediments, but at a cost of very high false positive rates (failed nontoxic sediments).

As was demonstrated by Apitz and Agius (this issue), more conservative AL1s, the expansion of the AL1 Action List (i.e., the short DaS list vs the longer list of “consensus” parameters), and the addition of a screening bioassay (i.e., consensus vs consensus plus screening Microtox) all improve the ability of the AL1 to identify sublethally and acutely toxic sediments, but at a cost of false positives. However, the AL2s are not very effective at identifying sediments that are most likely toxic. Although the bulk of samples that are acutely toxic also exceed cAL2, a much larger proportion of nontoxic samples fail cAL2. Thus, the use of a cAL2 to avoid unnecessary bioassessment (on the assumption that these samples are most probably toxic) is not supported, as the result would be a significant potential of nontoxic samples being refused unconfined ocean disposal.

Whether these samples should be refused ocean disposal because they have contaminant levels too high to allow for disposal at sea, regardless of toxicity, is a policy decision.

The consensus AL1s, which have the longest action list and are generally more conservative than the geometric mean of the other AL1s reviewed, perform significantly better than the previously 4-constituent Canadian DaS AL1 (as demonstrated above and in Apitz and Agius this issue). Class II rates are reduced to 6.7%, in the range of, but slightly higher than Ireland (6.3%) and the United Kingdom (5.0%). Class III rates (0.6%) meet the level of Denmark and the United Kingdom and are just below Ireland. The addition of a screening bioassay in the form of a more conservative Microtox decision rule in the Tier 1 assessment improves AL1 performance further—both Class II and III rates are the lowest in the review although at a cost of the highest rates of false positive (Class IV) assignments for any AL1 list barring that from the United Kingdom.

None of the AL2 lists perform particularly well at identifying sediments that are probably too contaminated for ocean disposal. Most frameworks have a higher level of false positives (negligibly toxic sediments that fail AL2) than sublethally or acutely toxic sediments. The only AL2 lists that result in relatively higher levels of toxic than nontoxic samples failing AL2 (Portugal, Spain, and the UK) have, relatively, the least conservative AL2s. Figure 4 illustrates the relative proportions of samples falling below AL1, between AL1 and AL2, and above cAL2 for all the cAL frameworks reviewed. These can be compared with the relative rates at which sediments in the database fall into each toxicity category, although Figure 4 clearly shows that the populations falling into the chemical and toxicological classifications do not completely overlap using any of the AL schemes, making clear that chemistry alone does not provide as protective or cost-effective an approach as a tiered approach including bioassessment, no matter how well-refined the ALs are.

As can be seen, the high AL2 levels result in the United Kingdom, Portugal, and Spain having the smallest proportion of samples falling above AL2. It should be noted, however, that Spain, and probably Portugal, along with Denmark and Germany, have ALs based on a fine-grained sediment fraction. If it were possible to correct for grain size in this data set, it is possible that the relative conservatism of the AL2s reviewed would be diminished. A possible approach to addressing this offset could be to “correct” ALs from countries addressing a different size fraction based on a hypothetical average grain size distribution. Although this would not correct for actual grain size distribution of samples in the database, it might allow for a more consistent

![Figure 4. Percent rates at which database samples fall into various AL classifications (or tiers). These are compared on the right to the relative proportion of sediments in the database that fall into each toxicity category.](image)
comparison of outcomes. The level of acceptable AL2 false positives (Class VII) is a policy decision. It is possible that AL2s can also represent contaminant concentrations deemed unacceptable for DaS, whether toxic or not, but such an assessment it outside the scope of this review.

CONCLUSIONS

National DaS CALs and chemical action lists differed greatly in their degree of protectiveness. Differences were a function of action lists and the relative conservatism of ALs. The efficacy and fitness for purpose of ALs in disposal at sea frameworks depend on a range of parameters, from values of chemical ALs and benchmarks (levels of AL1 and AL2), chemical decision rules, whether frameworks are tiered or not, and how other LOEs are used and interpreted in the framework. Thus, the question of whether the chemical action levels are appropriate and fit for purpose is not a straightforward one. The high-level review presented in this paper addressed a number of interpretation scenarios for a range of DaS ALs and reviewed and compared their potential performance using a database of colocated sediment chemical and toxicological data. It is not the intent of this review to make statements on the overall efficacy of any national framework, applied in its entirety, within a given region. Rather, the intent is to infer the potential effects of specific policy choices within a complex decision process. To allow for cross-national comparisons, a number of simplifications and assumptions needed to be made; specific conclusions about national outcomes would require more detailed analysis, nation-specific database adaptation, and, ideally, nation-specific data. Thus, the conclusions in this article are intended to contribute to regional, national and international discussions, rather than be seen to provide definitive answers or proposals. However, a number of general conclusions can be drawn as follows.

Screening Levels

- More conservative (low) AL1s tend to be better at identifying toxic sediments, but do this at a cost of higher false positive rates.
- AL1 strategies and chemical action lists can be designed that are very effective at filtering out the majority of toxic sediments. However, unless this is followed by subsequent toxicological assessment, a large proportion of nontoxic sediments will be unnecessarily subjected to treatment and containment or land-based disposal. Excessive false positive results can also have other unintended impacts—Canada has observed that, due to the perceived complexity and uncertainty of Tier 1 "fails," applicants often choose to forgo further toxicity testing (AL2 in Canada) and either not to dredge (potentially inhibiting development) or to go directly to land-based disposal, which falls into a different regulatory framework that may or may not have less overall ecological and economic impacts (Apitz and Agius this issue). False positive rates can be reduced with the appropriate use of bioassessment in a tiered framework.
- Screening or Tier 1 approaches are improved by using expanded chemical action lists and adding a screening level bioassay.
- As was shown in Apitz and Agius (2013) and Apitz and Agius (this issue), expanding the chemical action list increases the efficacy of AL1s in an overall screening or Tier 1 approach. The addition of a screening bioassay increased AL1 efficacy but not to the point where it could replace chemical measurements (Apitz and Agius this issue).

Use of Bioassessment

- Without bioassessment, a significant proportion of sediments would be misclassified at both the AL1 and the AL2 level.
- Increased proponent costs associated with adding bioassessment to Tier 2 are arguably justified given the resulting improvement in framework performance but require clearly delineated assessment protocols to deal with sediment that fall between ALs.

Assessment costs increase in tiered frameworks that subject AL1 failures to toxicological assessment, but this cost to proponents is arguably justified in that the majority of samples failing AL1 are likely to pass Tier 2 toxicity assessment, and thus, in many cases, be considered suitable for DaS, whereas at the same time, increased environmental protection through decreased Class III (false negative) rates is achieved. On the other hand, where AL failure is not followed by further assessment (e.g., the Netherlands) or where assessment protocols for AL1 failures are not clearly delineated (as in Denmark, Norway, Ireland, and the UK), it is unclear what will happen to these nontoxic AL1 failures. If they are refused for uncontrolled DaS, despite being nontoxic, this not only poses unnecessary expense on applicants, as treatment and containment have high costs (Bortone et al. 2004), but there are also environmental consequences in terms of land, energy and water use, transport and management risks, loss of disposal/storage capacity, and possibly, refusal or withdrawal of dredging permits or applications (Apitz and Black 2010). Nontoxic sediment not returned to the marine environment also has impacts on sediment balance within coastal systems (Apitz 2010b, 2012).

Use of chemical-only AL2s

- This review did not support the use of chemical AL2s to predict the degree to which sediments will be toxic. Based on the assumptions in this study, none of the AL2s examined were particularly effective in identifying toxic sediments that should be refused DaS without further
assessments nor did they help distinguish well between sediments that should be permitted for disposal at sea with some management or monitoring and those that should require treatment, containment, or upland disposal.

On the other hand, chemical AL2s may be used to define contaminant levels that would be refused based solely on their contaminant levels regardless of their toxicity. The efficacy of chemical AL2s for this narrative intent is outside the scope of this review.

- The main conclusion of this review was that chemistry alone, no matter how conservative or carefully designed, does not perform well at both minimizing bioassessment costs and protecting the environment. It appears clear that chemistry alone does not provide as protective or cost-effective an approach as an approach including bioassessment (whether tiered or not), no matter how well-refined the ALs are.

**Most Promising Framework Modifications**

- Modification to ALs and decision rules such as longer chemical action lists and the addition of a screening bioassessment can greatly improve performance. Expanded chemical action lists and the use of mean hazard quotients instead of one-out, all-out decision rules resulted in improved framework performance. However, no well-performing country (in terms of the prediction of toxicity) had uniformly long action lists, so the correct approach to “tuning” these still requires some work.

**Policy Choices**

- The acceptable balance between correctly classifying sediments as toxic or nontoxic is not a scientific decision, but a policy one. A well-designed, tiered system should be able to discriminate most toxic and nontoxic sediments, but even carefully designed tiered assessment will result in the failure to catch some sublethally or acutely toxic sediments. The acceptable balance between these objectives is a policy, not a scientific, decision.

**Acknowledgment**—This work was funded in part by Environment and Climate Change Canada’s Marine Protection Programs. The Coastal and Oceanographic Assessment, Status and Trends (COAST) Branch, part of NOAA’s National Centers for Coastal Ocean Science in the Center for Coastal Monitoring and Assessment (CCMA) is gratefully acknowledged for making its extensive data sets available online. This work also draws upon the research project “High level review of current UK action level guidance” (Project MMO1053) carried out by Cefas and SEA Environmental Decisions for the Marine Management Organisation in the United Kingdom.

**Disclaimer**—This article does not necessarily represent the views of the Environment and Climate Change Canada, the Centre for Environment, Fisheries & Aquaculture Science, or any affiliations represented by the authors. References to brand names and trademarks in this document are for information purposes only and do not constitute endorsements by Environment and Climate Change Canada, the Centre for Environment, Fisheries & Aquaculture Science, or the authors. It is not the intention of the authors to suggest conclusions on the potential ecological risk or regulatory status of the sediments from which the database was drawn; these samples were not collected for the assessment of ocean disposal and this review represents an analysis of only a small fraction of the data available. These data are only used to provide a data set that might realistically represent the range of sediment types that might be encountered by a country’s DaS program to evaluate the potential performance of a range of DM DaS decision rules.

**Data Accessibility**—The NOAA Status and Trends (NS&T) and Mussel Watch data sets are available online. See https://products.coastalscience.noaa.gov/collections/monitoring/ nsandt/ for NS&T and https://data.noaa.gov/data set/national-status-and-trends-mussel-watch-program for Mussel Watch.

**SUPPLEMENTAL DATA**

Data S1. Database and chemical Action Level (AL) background.

Data S2. Review of AL values.

Data S3. Evaluation of AL efficacy.

**REFERENCES**

Agius SJ, Porebski L. 2008. Towards the assessment and management of contaminated dredged materials. Integr Environ Assess Manag 4:255–260.

Apitz SE. 2008. The derivation and application of sediment quality guidelines (SQGs) and bioaccumulation models in the screening and first tier of the assessment of sediment quality. Report for the Canadian disposal at sea program. Little Hadham (UK): SEA Environmental Decisions Ltd. 90 p.

Apitz SE. 2010a. Review of the use of bioassays for Canada’s disposal at sea programme—Final Report to Environment Canada. Little Hadham (UK): SEA Environmental Decisions Ltd. 180 p.

Apitz SE. 2010b. Waste or resource? Classifying and scoring dredged material management strategies in terms of the waste hierarchy. J Soils Sediments 10:1657–1668.

Apitz SE. 2011. Integrated risk assessments for the management of contaminated sediments in estuaries and coastal systems. In: McLusky D, Wolanski E, editors. Treatise on estuarine and coastal science. Waltham (UK): Academic. p 311–338.

Apitz SE. 2012. Conceptualising the role of sediment in sustaining ecosystem services: Sediment-ecosystem regional assessment (SEcoRA). Sci Total Environ 415:9–30.

Apitz SE. 2014. Review of disposal at sea program action level guidance for Canada and other OSPAR nations: Evaluation considering two chemical action levels for Environment Canada, Draft Report, 7 March 2014. Little Hadham (UK): SEA Environmental Decisions Ltd. 51 p.

Apitz SE, Agius S. 2013. Anatomy of a decision: Potential regulatory outcomes from changes to chemistry protocols in the Canadian disposal at sea program. Mar Pollut Bull 69:76–90.
Apitz SE, Agius S. 2017. Anatomy of a decision II: Potential effectiveness of changes to tiered sediment assessment protocols in the Canadian disposal at sea program. *Integr Environ Assess Manag* 13:1072–1085.

Apitz SE, Barbanter A, Bernstein AG, Bocci M, Delaney E, Montobbio L. 2007. The assessment of sediment screening risk in Venice Lagoon and other coastal areas using international sediment quality guidelines. *J Soils Sediments* 7:326–341.

Apitz SE, Black K. 2010. ME1104 research and support for developing a UK strategy for managing contaminated sediments: An analysis of the project findings. Glasgow (UK): Partrac and SEA Environmental Decisions Ltd. 100 p. Report to Defra.

Apitz SE, Crane M, Power EA. 2005. Use of sediment quality values (SQVs) in the assessment of sediment quality. Final Report to the Environment Agency of England and Wales. Faringdon (UK): Crane Environmental.

Apitz SE, Power B. 2002. From risk assessment to sediment management: An international perspective. *J Soils Sediments* 2:61–68.

Apitz SE, Vivian C, Birchenough A. 2014. High level review of current UK action level guidance (IMO1053) for Marine Management Organisation, 28 March 2014. Lowestoft (UK): Cefas and SEA Environmental Decisions Ltd. 77 p.

Barcelo D, Petrovic M. 2007. Sediment quality and impact assessment of pollutants, sustainable management of sediment resources. Amsterdam (NL): Elsevier. 333 p.

Birchenough A, Mills A, Vivian C. 2006. R&D SC050066 - Trial of proposed tiered framework for the assessment of sediment quality. Bristol (UK): Environment Agency.

Bortone G, Arevalo E, Deibel I, Detzner H-D, Propris LD, Elskens F, Giordano A, Hakstge P, Hamer K, Harmsen J, et al. 2004. Sediment and dredged material treatment synthesis of the SedNet work package 4 outcomes. *J Soils Sediments* 4:225–232.

[Cefas] Centre for Environment, Fisheries and Aquaculture Science. 2003. Monitoring of the quality of the marine environment, 1999–2000. Lowestoft (UK): Centre for Environment, Fisheries and Aquaculture Science. Aquatic Environment Monitoring Report Number 54.

[Cefas] Centre for Environment, Fisheries and Aquaculture Science. 2005. Background levels and the anthropogenic component, of naturally-occurring elements in marine sediments subject to dredging and disposal. Burnham-on-Crouch (UK): Centre for Environment, Fisheries and Aquaculture Science. Report to Defra, Project AE0257.

Chapman PM, McDonald BG, Lawrence GS. 2002. Weight-of-evidence issues and frameworks for sediment quality (and other) assessments. *Hum Ecol Risk Assess* 8:1489–1515.

[CoA] Commonwealth of Australia. 2009. National Assessment Guidelines for Dredging. Canberra (AU): Commonwealth of Australia. 92 p.

[IMO] International Maritime Organization. 2016. London Convention and London Protocol, 2016 edition. London (UK): IMO Publishing.

[NS&T] National Status and Trends. 2012. National Status and Trends program download page. [cited 2012 February 2]. https://products.coastalscience.noaa.gov/collections/lmmonitoring/nsandt/

Ogden Environmental and Energy Services Company. 1997, 1998. Pearl Harbor sediment remedial investigation/feasibility study, Pearl Harbor, Hawaii, fact sheets. Pearl City (HI): Pearl City Public Library.

[OSPAR] Oslo-Paris Convention. 2008. Overview of contracting parties’ national action levels for dredged material (2008 Update). Biodiversity Series, No 363. OSPAR Commission Report.

[OSPAR] Oslo-Paris Convention. 2014. OSPAR guidelines for the management of dredged material; Agreement 2014-06. Convention for the Protection of the Marine Environment of the North-East Atlantic (OSPAR Convention).

[OSPAR] Oslo-Paris Convention. 2015. New Spanish guidelines for the management of dredged material and its relocation in waters of the maritime-terrestrial public domain. Meeting of the Environmental Impact of Human Activities Committee (EIHA); 2015 April 13–17 Santander, Spain. London (UK): OSPAR Convention for the Protection of the Marine Environment of the North-East Atlantic. EIHA 15/3/Info.1-E.

Roper H, Netzband A. 2011. Assessment criteria for dredged material with special focus on the North Sea region. Hamburg (DE): Hamburg Port Authority. 36 p.

Wenning RJ, Batley GE, Ingersoll CG, Moore DW. 2005. Use of sediment quality guidelines & related tools for the assessment of contaminated sediments (SQG). Pensacola (FL): SETAC Pr. 783 p.